The Computation of Surface Lightness in Simple and Complex Scenes

Alexandra C. Schmid

School of Psychology

The University of Sydney

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

2016

Declaration of Originality

This thesis is submitted to the University of Sydney in fulfilment of the requirements for the Degree of Doctor of Philosophy. The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.

Alexandra C. Schmid

SYDNEY, October 2015

Abstract

The present thesis examined how reflectance properties and the complexity of surface mesostructure (small-scale surface relief) influence perceived lightness in centre-surround displays. Chapters 2 and 3 evaluated the role of surface relief, gloss, and interreflections on lightness constancy, which was examined across changes in background albedo and illumination level. For surfaces with visible mesostructure ("rocky" surfaces), lightness constancy across changes in background albedo was better for targets embedded in glossy versus matte surfaces. However, this improved lightness constancy for gloss was not observed when illumination varied. Control experiments compared the matte and glossy rocky surrounds to two control displays, which matched either pixel histograms or a phase-scrambled power spectrum. Lightness constancy was improved for rocky glossy displays over the histogram-matched displays, but not compared to phase-scrambled variants of these images with equated power spectrums. The results were similar for surfaces rendered with 1, 2, 3 and 4 interreflections. These results suggest that lightness perception in complex centre-surround displays can be explained by the distribution of contrast across space and scale, independently of explicit information about surface shading or specularity. The results for surfaces without surface relief ("homogeneous" surfaces) differed qualitatively to rocky surfaces, exhibiting abrupt steps in perceived lightness at points at which the targets transitioned from being increments to decrements. Chapter 4 examined whether homogeneous displays evoke more complex mid-level representations similar to conditions of transparency. Matching target lightness in a homogeneous display to that in a textured or rocky display required varying both lightness and transmittance of the test patch on the textured display to obtain the most satisfactory matches. However, transmittance was only varied to match the contrast of targets against homogeneous surrounds, and not to explicitly match the amount of transparency perceived in the displays. The results suggest perceived target-surround edge contrast differs between homogeneous and textured displays. Varying the mid-level property of transparency in textured displays provides a natural means for equating both target lightness and the unique appearance of the edge contrast in homogeneous displays.

Acknowledgements

First and foremost I would like to thank my supervisor Professor Bart Anderson for his guidance and support throughout my PhD candidature. I have been lucky to have a supervisor who responds to emails almost instantaneously at any time of day, and provides exceptionally quick feedback on chapter drafts. I would also like to thank Bart for inspiring me to become a better thinker and writer, and giving me the opportunity to learn many valuable skills working as an RA in the years leading up to my PhD. Above all, you have imparted your extensive knowledge and passion for the field of material perception, which I hope to carry into my future research career.

Thank you to past and present members of the Anderson lab for creating a great working environment. I would particularly like to thank the post-docs in the lab Phil Marlow and Juno Kim for helping me whenever I was stuck with a programming problem. Thank you Kai for being one of the few people I could talk to about the lightness literature, and whiz kid Scott for learning to program Blender in his first week of honours, and inadvertently motivating me to do the same. Thank you Siva for your company in the lab towards the end.

I would like to thank my family and friends, especially my parents Rennie and Robyn for their constant encouragement and support of my educational pursuits. You have taught me that as long as I try hard I can achieve anything I put my mind to. Thank you Danielle for being a great and supportive friend, and willing research participant.

Finally, I would like to express my love and gratitude to my partner Marios for going through both my honours and PhD theses with me, while simultaneously finishing his own PhD thesis. Your integrity as a researcher and constant enthusiasm for learning is an inspiration to myself and others.

List of Publications

Peer-Reviewed Articles

Schmid, A. C., & Anderson, B. L. (2014). Do surface reflectance properties and 3-D mesostructure influence the perception of lightness? *Journal of Vision, 14*(8):24, 1-24.

Oral Presentations

Schmid, A. S., & Anderson, B. L. (2015). Surface lightness in complex centre-surround displays: The important role of interreflections. 42nd Australasian Experimental Psychology Conference (EPC), Sydney, NSW, Australia.

Schmid, A. S., & Anderson, B. L. (2014). How does the brain represent surface lightness for simple versus complex centre-surround displays? 41st Australasian Experimental Psychology Conference (EPC), Brisbane, QLD, Australia.

Schmid, A. S., & Anderson, B. L. (2013). How does the brain represent surface lightness? Psychology Postgraduate Conference, Sydney, NSW, Australia.

Schmid, A. S., & Anderson, B. L. (2013). Lightness perception in scenes that are 3D and naturalistic versus 2D and impoverished. University of Sydney Faculty Heats for the Three Minute Thesis Competition (3MT), Sydney, NSW, Australia.

Schmid, A. S., & Anderson, B. L. (2013). We are living in a material world: Lightness perception is more consistent in natural scenes that have real material properties such as rockiness and gloss. 1st Inaugural University of Sydney's School of Psychology Day of Research (PsychFest), Sydney, NSW, Australia.

Conference Proceedings

Schmid, A. S., & Anderson, B. L. (2014). Linking low-level and mid-level accounts of lightness perception. 14th Annual Vision Sciences Society Conference (VSS), St. Pete Beach, FL, USA.

Contents

Abstract i

Acknowledgements ii

List of Publications iii

List of Figures vi

List of Tables viii

Chapter 1. Determining how the Visual System computes Surface Lightness in Complex Scenes 1

Classes of Lightness Theories 4

Brightness Models 5

Framework models 10

Decomposition models 11

Relating Existing Lightness Theories to Real-World Scenes 17

Potential Information about Lightness Available in Real-World Scenes 19

Shading, cast shadows and specular highlights 19

Texture caused by complex mesostructure 24

Unanswered Questions 26

Chapter 2. Do Surface Reflectance Properties and 3D Mesostructure influence the Perception of Lightness? 29

Experiment 1A and 1B: Varying Surface Relief and Gloss Level of the Surround 30

Methods 30

Results and discussion 41

Experiment 2A and 2B: Variegated Surrounds 50

Methods 51

Results and discussion 54

Experiment 3: Flat, Matte Adjustable Surface 57

Methods 57

Experiment 4: Phase-Scrambled Surrounds 62

Methods 62

Results and discussion 64

Concluding Thoughts 66

Chapter 3. The Role of Interreflections on Lightness Perception under Changing Illumination Level 69

Experiment 5A and 5B: Lightness Perception of Complex Surfaces under Changing Illumination Level and Number of Interreflections 69
Methods 70
Results and discussion 76
Chapter 4. Perceptual Mechanisms underlying Lightness Perception in Homogeneous Centre-Surround Displays 89
Experiment 6: Matching the Texture of the Central Patch and the Surround 89
Methods 90
Results and discussion 93
Experiment 7: Matching Surfaces on Two Dimensions 101
Methods 103
Results and discussion 106
Experiment 8: Which Dimensions lead to Better Matches? 118
Methods 118
Results and discussion 121
Experiment 9A and 9B: Investigating the Increment-Decrement Asymmetry 125
Methods 125
Results and discussion 126
Experiment 10: Low Contrast Surrounds 132
Methods 132
Results and discussion 135
Chapter 5: General Discussion 138
Summary of Experimental Results 138
Relation to Previous Work 140
Lightness computations in complex displays 140
Lightness computations in simple displays 143
Relation to the simultaneous contrast effect 148
Conclusions 150
References 151

List of Figures

Figure 1.1. Examples of simple stimuli used in many lightness studies. 4

Figure 1.2. Workshop metaphor taken from Adelson and Pentland (1996). 13

Figure 1.3. Stimuli used in Koenderink et al.'s (2007a) experiment. 21

Figure 1.4. The cue conditions in Boyaci et al.'s (2006a) study. 23

Figure 1.5. Setup of Snyder et al.'s (2005) experiment. 23

Figure 2.1. Computer rendered centre-surround stimuli used in the experiments. 31

Figure 2.2. Paper texture used to deform low-relief (flat) surfaces. 33

Figure 2.3. The "grove" light field used to illuminate surfaces used in the experiments. 35

Figure 2.4. Transformation of HDR luminance values (x-axis) to tone-mapped luminance values (y-axis). 37

Figure 2.5. Test patch albedos for surround reflectance 19.8%. 39

Figure 2.6. Average data for Experiment 1A (left panels) and 1B (right panels). 43

Figure 2.7. Increment minus decrement settings for Experiment 1A (A) and 1B (B). 43

Figure 2.8. Average difference scores for the rocky conditions of Experiment 1A (A) and 1B (B). 48

Figure 2.9. Examples of variegated centre-surround stimuli used in Experiment 2. 50

Figure 2.10. Results of Experiment 2. 52

Figure 2.11. Difference scores from Experiment 2. 53

Figure 2.12. Average data for Experiment 3. 59

Figure 2.13. Increment minus decrement settings for Experiment 3. 59

Figure 2.14. Average difference scores (left panel) and standard deviation scores (right panel) for Experiment 3. 61

Figure 2.15. Examples of phase-scrambled centre-surround stimuli used in Experiment 4...63

Figure 2.16. Average data for Experiment 4. 65

Figure 2.17. Average difference scores (left panel) and standard deviation scores (right panel) for Experiment 4. 66

Figure 3.1. Examples of difference maps between images rendered with 1, 2, 3 and 4 interreflections, for the matte white (A) and grey (B) surrounds rendered under bright illumination. 72

Figure 3.2. Comparison of matte white and grey displays rendered with 1 and 4 interreflections. 74

Figure 3.3. Examples of phase-scrambled stimuli for each interreflection condition. 75

Figure 3.4. Average data for Experiment 5A, plotted in Munsell values. 77
Figure 3.5. Graphs of interaction effects in Experiment 5A. 80
Figure 3.6. Average data for Experiment 5A, plotted in log luminance. 85
Figure 3.7. Average surround Munsell settings from Experiment 5B, plotted next to test patch settings from Experiment 5A. 85
Figure 4.1. Stimuli used in Experiment 6. 93
Figure 4.2. Average results for Experiment 6. 94
Figure 4.3. Average data from the continuous and discontinuous texture conditions plotted together, for matte (left panel) and glossy (right panel) displays. 98
Figure 4.4. Illustration of perceptual decomposition (scission) of homogeneous centre-surround displays. 103
Figure 4.5. Adjustable displays used in Experiments 7 and 9. 105
Figure 4.6. Lightness settings for Experiment 7, plotted in Munsell values. 110
Figure 4.7. Increment minus decrement settings for the matte (A) and glossy (B) adjustable surfaces in Experiment 7. 110
Figure 4.8. Transmittance settings from Experiment 7, for the matte (A) and glossy (B) condition. 112
Figure 4.9. Representation of the layout of a trial in Experiment 8. 120
Figure 4.10. Results of Experiment 8 for the matte condition (A) and the glossy condition (B). 123
Figure 4.11. Results of Experiment 8, averaged across test patch Munsell. 123
Figure 4.12. Test patch lightness settings for Experiment 9A (A) and Experiment 9B (B). 127
Figure 4.13. Transmittance settings for Experiment 9A. 129
Figure 4.14. Transformed transmittance data from Experiment 9A. 130
Figure 4.15. Stimuli used in Experiment 10. 134
Figure 4.16. Results of Experiment 10. 136
Figure 4.17. Immediate increment and decrement settings for each display type condition. 136

List of Tables

Table 2.1. Modifiers used to create the effects in the low-relief (flat) and high-relief (rocky) surrounds. 33

Table 2.2. Surround and centre patch reflectance and luminance values of the test surfaces used in the experiments. 40

Table 2.3. *F* values and *p* values for the increment–decrement difference scores of Experiment 1A........ 45

Table 2.4. *t* values, *p* values, and *df* comparing the increment-decrement difference scores between conditions in Experiment 1B........ 45

Table 2.5. *t* values, *p* values, and *df* for the increment-decrement difference scores of Experiment 3. 60

Table 4.1. *t* values and *p* values comparing lightness settings between the surround-albedo test patch to the six lowest contrast test patches against the surround. 96

Table 4.2. *t* values, *p* values, and *df* comparing the lightness settings between the continuous and discontinuous texture conditions, for each test patch Munsell value. 99

Table 4.3. Matte matching display: *t* values and *p* values comparing transmittance settings between the contrast and transparency instruction conditions, for each surround albedo and test patch albedo condition. 116

Table 4.4. Glossy matching display: *t* values and *p* values comparing transmittance settings between the contrast and transparency instruction conditions, for each surround albedo and test patch albedo condition. 117

Table 4.5. *t* values and *p* values comparing preferences in Experiment 8 for the contrast instructions condition over the no transmittance control condition, for each surround Munsell condition and gloss level. 124

Table 4.6. The slope and y-intercept for each function fitted to increments in Figure 4.14.131

Chapter 1. Determining how the Visual System computes Surface Lightness in Complex Scenes

When we look at a scene, the image projected onto the retina is perceptually parsed into different causal sources. We get distinct impressions of object shape, illumination, and surface reflectance properties (lightness, colour, gloss, and transparency). The idea that our visual experience corresponds to the physical properties of an object in the world (distal stimulus) rather than retinal stimulation (proximal stimulus) can be traced back to Helmholtz (1866/1962). Helmholtz recognised that a white piece of paper placed in shadow still looks white, and can be distinguished from a black piece of paper in sunlight that projects the same *luminance* on the retina (local retinal stimulation is the same). In this example the two papers have identical luminance, but the perception of surface *lightness* (how light or dark the pigment is perceived) correlates with the different physical *albedos* of the papers (the proportion of light reflected diffusely, also called *diffuse reflectance*). This classic example highlights the extensively studied problem of lightness perception: how does the visual system recover the lightness of surfaces? The computation of lightness is typically regarded as an underconstrained problem because, as demonstrated by Helmholtz's example with paper, any particular luminance could be generated by an infinite combination of surface albedos, illuminations, and surface pose. While still unresolved, a number of theories have been developed that propose how the visual system constrains estimates of surface lightness.

The earliest theories of lightness perception were put forward by Helmholtz (1866/1962) and Hering (1874/1964), whose ideas have underpinned most subsequent theories. Helmholtz posited that to constrain estimates of surface lightness, the visual system first infers the illumination. Hering argued that to infer the illumination, reflectance must be known, suggesting that Helmholtz's views were circular. He claimed that lightness could be determined through peripheral sensory mechanisms such as pupil size, adaptation, and reciprocal interactions in the neural image. Hering instigated an approach to lightness perception that considers how surface lightness is affected by the initial processing of the image. Helmholtz talked about cognitive influences on lightness, but he later inspired an approach to lightness perception that explores how feedback from higher visual areas influences the way luminance values are processed and mapped to lightness.

Many subsequent theories of lightness perception have incorporated Helmholtz and Hering's ideas, and have proposed different perceptual mechanisms that might be involved in lightness computations. These mechanisms are often referred to as low-, mid- and high-level mechanisms. These terms have been used to refer to different levels of processing in the visual stream, or more generally how (or whether) scene layout and cognitive factors influence perception. In the present thesis, we use the term "low-level" to refer to how surface lightness is affected by the transduction of light in the retina and early visual processes such as adaptation (Helson, 1943), contrast (Wallach, 1948), and lateral interactions of cells at the retina (Cornsweet, 1970; Hurvich & Jameson, 1957; Jameson & Hurvich, 1964). We also use the term low-level to refer to theories of lightness perception that are primarily concerned with these processes. Although low-level mechanisms may account for some lightness phenomena (low-level processing occurs first by definition), an important question is whether low-level explanations are sufficient to account for all lightness phenomena. It is possible that feedback from higher visual areas might influence the initial processing of the image. "High-level" mechanisms refer to top-down cognitive factors that might influence lightness, and are not a focus of the present thesis. "Mid-level" is often defined as an ill-specified region between low- and high-level processing (Adelson, 2000). Proposed low-level mechanisms often have connections to physiology, but the physiology associated with mid-level processes is less clear. We refer to mid-level processes as those that involve representations of surfaces, objects, 3D structure, and/or illumination. We also use the term mid-level to refer to theories that postulate that lightness computations are influenced by scene organisation and the segmentation of surfaces into separate objects (e.g. Anderson, 1997; Anderson & Winawer, 2005, 2008; Barrow and Tenenbaum, 1978), or those that invoke estimates of the illumination (e.g. Boyaci et al., 2006a; Snyder et al., 2005). The aim of the present thesis is to tease apart low- and mid-level contributions to lightness perception.

A second aim of this thesis is to explore how lightness is computed for complex natural surfaces. Early theories of lightness perception adopted an experimental approach that still persists in modern lightness studies. This approach uses extremely simplistic stimuli with only a few luminance values arising from smooth, matte, flat surfaces (e.g. Annan & Gilchrist, 2004; Arend & Spehar, 1993a, 1993b; Bressan, 2006; Economou et al., 2007; Gilchrist et al., 1999; Gilchrist, 1977, 1979; Kingdom, 2011). Examples include the simultaneous contrast display (Figure 1.1A), and displays called Mondrians (Land & McCann, 1971), so named because they bear a vague resemblance to the work of the painter Piet Mondrian (Figure 1.1B). These *simple scenes* are vastly different than the scenes that people normally encounter, but continue to dominate the lightness literature because they allow rigorous experimental control over extraneous variables. In such impoverished scenes, disentangling the contributions of lightness and illumination is a maximally ill-posed problem (e.g. see Anderson et al., 2014). However, most natural surfaces are made of materials that are not characterised by purely diffuse reflectance but by a complex reflectance function, and often contain medium-scale surface relief (mesostructure). A surface's microstructure typically generates both diffuse and specular reflections whereas a surface's mesostructure interacts with the light field to create shading, shadows, and illuminance flow. These features potentially provide the visual system with information not available in simple scenes that may be involved in computations of surface lightness. This will be discussed in detail in a later section.

In the following sections we outline the major theories of lightness perception, then evaluate the extent to which these theories can be generalised to information-rich complex scenes normally encountered by humans. This discussion is followed by evidence that the visual system has access to extra image information produced by complex scenes that has not been incorporated into any major theory of lightness perception. Finally, it will be hypothesised that the visual system uses this extra information when computing surface lightness.

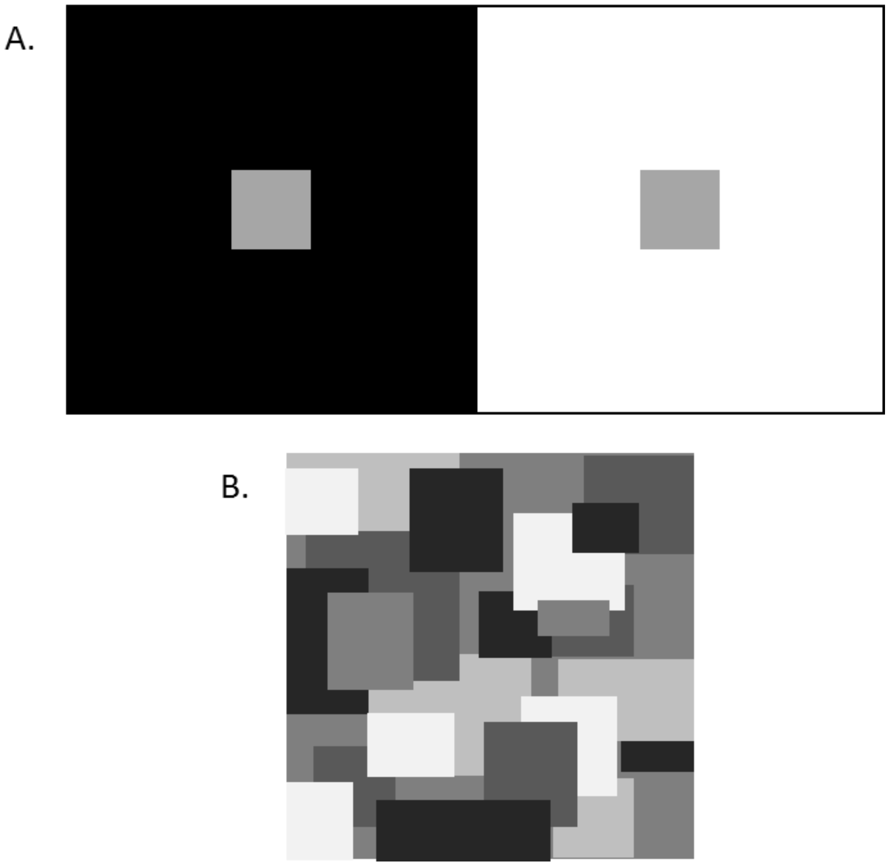


Figure 1.1. Examples of simple stimuli used in many lightness studies. (A) Simultaneous contrast display. The patch on the black background looks lighter than the patch on the white background even though they are equiluminant. (B) Example of a Mondrian display.

Classes of Lightness Theories

Theories of lightness perception, most of which are based on Helmholtz's and Hering's ideas, can be categorised into three broad classes: (1) brightness models, which do not distinguish lightness from brightness (perceived luminance) and do not require an explicit representation of the illuminant; (2) framework models, which divide scenes into frameworks of approximately constant illumination but also do not advocate an explicit illuminant representation; and (3) decomposition models, which posit that the retinal image is perceptually divided (*decomposed*) into representations of illumination and surface reflectance. Brightness models are mostly based on low-level perceptual mechanisms, whereas framework and decomposition models propose that mid-level processing of scene layout influences surface lightness.

Brightness Models

Hering, along with Mach (1865), anticipated the discovery of lateral inhibition, which entails the opposing centre-surround organisation of the receptive fields of lower order visual neurons (i.e. retinal ganglion cells through to cells in the lateral geniculate nucleus; Hartline, 1940; Hartline & Graham, 1932; Kuffler, 1953, 1973; Wiesel & Hubel, 1966). The discovery of lateral inhibition led to the development of low-level theories of brightness perception based on contrast (Cornsweet, 1970; Helson, 1943; Hurvich & Jameson, 1957; Jameson & Hurvich, 1964; Wallach, 1948). These early contrast theories are referred to as brightness models because they are insensitive to the various sources of image structure (e.g. reflectance, illumination, and surface pose), and thus model brightness perception without dissociating it from lightness.

Early contrast models

Two influential early contrast theories were put forward by Wallach (1948) and Helson (1943, 1964). Wallach (1948) proposed that lightness is determined by the luminance ratio between a surface and its adjacent surround. He developed a ratio principle from experimental results that showed that a disk surrounded by a brighter annulus appears the same shade of grey as another disk surrounded by an annulus that has the same luminance ratio, regardless of the absolute luminance values of the elements in the display. Helson's (1943, 1964) adaptation level theory was also based on luminance ratios. In this model, the lightness of a patch is determined by the luminance ratio between the patch and the average luminance (adaptation level) of a collection of surfaces called an "adaptive window". The average luminance, or adaptation level, is weighted toward surfaces closer to the patch. According to Helson, the boundaries of the adaptive window should coincide with the boundaries of illumination fields. However, neither Helson's nor Wallach's models articulated any principled way of identifying illumination boundaries.

Other early contrast theories were based on lateral inhibition, such as Jameson and Hurvich's opponent-process theory (Hurvich & Jameson, 1957; Jameson & Hurvich, 1964) and Cornsweet's (1970) theory. These theories asserted that the brightness of a target patch is the net result of excitation and inhibition of centre-surround cells in the retina. The perception of brightness was considered to be isomorphic to the response profiles of these neurons, i.e. lightness was proposed to directly correspond to the net level of excitation and inhibition at a given location. In these models, the level of excitation of the target patch is determined by the luminance of the patch itself and the level of inhibition is determined by the average luminance of the target's immediate surround. For example, the simultaneous contrast effect in Figure 1.1A would be explained in the following way: the excitation of the patch on the white background is lowered due to inhibition from the background, while the patch on the black background is not affected by such inhibition. This leads to the patch on the black background appearing brighter than the one on the white background.

Retinex

In the early 1970s, Edwin Land described the concept of constancy, where objects and surfaces appear to have a consistent colour or lightness regardless of temporal or spatial changes in the intensity of the illumination. He formulated retinex theory to explain colour and lightness constancy (Land & McCann, 1971). "Retinex" is a combination of the words "retina" and "cortex", which suggests that the brain is also involved in the processing of lightness and colour, not just the retina as was suggested by early contrast theories. Land recognised that the visual system would have to rely on pre-existing assumptions when estimating lightness, which he incorporated into his model by assuming that rapid changes in luminance (i.e. edges) correspond to reflectance changes, while shallow gradients of luminance correspond to spatial variations in illumination. Illumination gradients were assumed to be "removed" by the visual system, or at least sensitivity to them was reduced. Although retinex can account for illumination that varies gradually across a scene, it is not capable of recognising edges that correspond to rapid changes in illumination, e.g. shadow boundaries.

Land and McCann's (1971) retinex model proposed that the visual system contains three to four retinal-cortical (retinex) systems, each sensitive to a band of wavelengths. Each retinex is an independent computation in terms of lightness, which is compared to other retinexes to determine colour. Their model was concerned with how lightness values are computed within each retinex image. The retinex model treats edges, which are derived at the retina, as an important source of information. The edges are subsequently integrated so that the lightness of local and remote areas of the visual field can be compared. This is computed by multiplying the luminance ratio at the edge of a first and second adjacent area by the luminance ratio at the edge of the second and third adjacent area. This multiplication of sequential edge ratios continues to a remote target area of interest to be compared to the starting area. Land and McCann claimed that the lightness ratio of any two areas in an image can be obtained by this method, while taking into account shallow gradients that may be present, caused by variations in the illumination. In a later version of the theory (Land, 1986), Land adopted an anchoring rule in which the highest relative reflectance computed by retinex is perceived as white.

Filling-in

The retinex model represents a computational approach to lightness, albeit one that does not focus on physiological mechanisms that might underlie these computations. Filling-in theories have suggested that brightness (or colour) signals interact with and propagate from edges, and have proposed a neural basis for this process. This proposed brightness propagation can stop at edges (e.g. Grossberg & Todorović, 1988), or can be spread and integrated across edges (e.g. Rudd, 2013).

Grossberg and colleagues (Cohen & Grossberg, 1984; Grossberg, 1983; Grossberg & Mingolla, 1987; Grossberg & Todorović, 1988) presented a filling-in model that proposes how edges are segmented and how the regions lying between edges are filled in with achromatic or chromatic colour. The model is comprised of interactions between two systems: a boundary contour (BC) system and a feature contour (FC) system. Prior to these systems, luminance is preprocessed by on- and off-centre ganglion and geniculate cells. The output signals from this processing provide input to both the BC and the FC system. The BC system consists of contrast-sensitive and orientationally tuned units, which putatively correspond to simple and complex cortical cells. The simple cell units are sensitive to orientation and contrast polarity, and are activated by the output from on-cells in the preprocessing stage. The complex cell units are sensitive to orientation regardless of contrast polarity, and are activated by simple cell units with the same axis of orientation, but opposite contrast polarity preference. Thus, the BC system encodes the location and orientation of edges, which is used as input to the FC system. The FC system contains an array of intimately connected cells that rapidly spread signals between each other. Bottom-up input from on-cells from the preprocessing stage generate FC signals that laterally diffuse from edges (defined by the BC system) to fill in regions between edges. Neural activity in the FC system corresponds to the percept of brightness. Note that Grossberg and colleagues assume that the BC system does not generate any visible percepts; we only consciously "see" colour and brightness that is propagated between BC signals.

In Grossberg and colleagues' model, the lateral spread of (FC) brightness signals is moderated by inhibition from (BC) edge signals, meaning that brightness signals are diffused and averaged within boundaries, but do not cross boundaries. More recently, Rudd and colleagues (Rudd, 2001, 2003, 2013; Rudd & Arrington, 2001; Rudd & Popa, 2007; Rudd & Zemach 2004, 2005, 2007) have developed a model combining edge integration and filling-in mechanisms, which allows filling-in signals to spread across edges. Similar to retinex theory, edge ratios are mathematically computed and then integrated across space. However, in Rudd's model, the value of the weight that edges are given falls off with distance from the patch of interest. Additionally, the edge weights depend on whether the target patch is a decrement (lower in luminance) relative to its surround, or an increment (higher in luminance) relative to its immediate surround. For decrements, far edges are weighted about 30% as large as the weight of an adjacent edge, while for increments far edges are weighted about 79% as large as the weight of an adjacent edge. Rudd has proposed that the edge integration path begins with regions of common background that surround surfaces of interest and propagates from there to the location of each individual surface via the shortest route. He has suggested that figural organisation from higher visual areas (e.g. cortical area V4) plays a role in selecting edges for long-range spatial integration.

Similar to retinex, some filling-in models (e.g. Grossberg & Todorović's, 1988) can take into account shallow gradients that are caused by inhomogeneous illumination (i.e. illumination that varies gradually across a scene). However, they are inherently brightness models because there is no principled way to differentiate illumination or depth edges from reflectance edges. Rudd talks about the propagation of lightness, but he has explicitly stated that the neural mechanisms involved in lightness and brightness computations are the same, and that cognitive or "top down" factors are involved in the interpretation of the neural image produced by the edge integration mechanism (Rudd, 2013). Thus his model is also essentially a brightness model.

Spatial filtering

A class of brightness models descended from early contrast theories involve spatial filtering and contrast normalisation of the retinal image. Like the early contrast theories, spatial filtering models generally do not distinguish between lightness and brightness, or if they do they regard the recovery of surface lightness as a separate, subsequent step to brightness computation and is not dealt with by the models. Over the years many spatial filtering models have been developed by researchers such as Marr (1982, Marr & Hildreth, 1980), Watt and Morgan (1985; MIRAGE), Morrone and Burr (1988), Kingdom and Moulden (1992; MIDAAS), Heinemann and Chase (1995), McArthur and Moulden (1999), and Shapiro & Lu (2011). These models were motivated by receptive field properties of early visual cortex, which is comprised of multiple filter channels tuned to different spatial frequencies and orientations (DeValois et al., 1982; DeValois & DeValois, 1988; Wilson & Wilkinson, 2003). Similar to filling-in models, each model applies a set of spatial filters based on receptive fields to an image. However, the receptive fields filter the image directly, and so the output of this filtering directly corresponds to perceived brightness. This is more similar to early contrast models involving lateral inhibition. Perhaps the most influential spatial filtering model is Blakeslee and McCourt's oriented difference-of-Gaussian multiscale filtering model (ODOG; Blakeslee & McCourt, 1999, 2001, 2004; Blakeslee et al., 2005, 2008). In the ODOG model, filters are constructed from the difference of Gaussians whose centres are shifted with respect to each other. They are characterised as Gaussian blobs with inhibitory flanks that are orientation and spatial frequency selective, and have been suggested to correspond to cortical simple cells (except that they are elongated in the wrong direction relative to those in the cortex; Blakeslee & McCourt, 2004). Six multiscale filters with different orientations are convolved with the image and then contrast normalised (using RMS contrast). This leads to six normalised outputs, which are summed to produce the final model output, which is a prediction of perceived brightness. Blakeslee and McCourt have stated that an advantage of their model is that it can explain effects such as White's effect, grating induction, and Mach bands, which are not predicted by filling-in models. However, unlike filling-in models, brightness computations in spatial filtering models are limited by the size of the largest scale spatial frequency filters or receptive fields.

Framework models

The theories outlined so far have emphasised low-level contributions to the perception of lightness[1]. Framework models distinguish lightness and brightness. They have also attempted to account for the effects of scene layout on lightness perception, so represent a mid-level approach to lightness perception. These models propose that the visual system divides images into regions of common illumination, without requiring that the visual system generates an explicit representation of the intensity of the illuminant. The visual system computes lightness within these regions, which are called atmospheres (Adelson, 2000) or frameworks (Bressan, 2006; Gilchrist, 2006; Gilchrist et al., 1999).

The most widely known of these models is Gilchrist's anchoring theory (Gilchrist, 2006; Gilchrist et al., 1999). The motivation behind anchoring theory came from the problem of mapping ambiguous luminance values and ratios onto lightness values. For example, two adjacent dark surfaces under bright illumination (e.g. black and dark grey) could produce the same luminances on the retina as two light surfaces under dim illumination (e.g. light grey and white). An anchoring rule is needed to assign fixed lightness values to ambiguous luminance values. Gilchrist resolved the anchoring problem in the same way as edge-integration theories (Land, 1986; Rudd, 2013). He asserted that the visual system assigns a fixed lightness value (white) to the highest luminance in a scene, which serves as a lightness anchor. Other lightness values are computed by forming luminance ratios relative to this anchor point. To determine how these luminance ratios are scaled, the theory proposes that the perceived range of greys

---

[1] with the exception of Rudd's model, which proposed that figural organisation affects the path of brightness propagation.

tends towards 30:1, which is the canonical range between black and white (Gilchrist, 2006). Scale normalisation occurs in the form of expansion when the range of luminances is less than 30:1, and compression when the range is greater than this.

For scenes containing more than just two luminance values, Gilchrist has asserted that luminance values are anchored and scaled relative to their *local framework* (an area of approximately constant illumination) and their *global framework* (the entire display or visual field; an idea first introduced by Kardos, 1934). Gilchrist has claimed that there are strong and weak factors that generate frameworks. Strong segmentation factors include penumbra and depth boundaries. Weaker segmentation factors include Gestalt grouping factors such as proximity, surroundedness (closely related to the Gestalt principle of figure/ground), common fate (Agostini & Proffitt, 1993), similarity (Laurinen et al., 1997) and T- and X-junctions. The strength or weight given to a framework depends on a number of factors. Stronger frameworks are larger, have a greater degree of articulation (defined as the number of surfaces in the framework), and stronger segregation (determined by the factors mentioned above).

Bressan (2006) proposed a modified version of anchoring theory, called double anchoring theory (DAT). The rules defining frameworks differ from the original model, and the “double” in double anchoring theory refers to the addition of a second anchoring step. The first anchoring step is the same as in the original anchoring model, i.e. the highest luminance is assigned the value of white. In the second step, the “surround” is also assigned a value of white, thus resulting in two lightness values being assigned to every surface within a framework. Final lightness values are calculated by weighting the two values according to factors such as the relative size of the surround, the degree of articulation in the surround, and its absolute luminance. Unfortunately, Bressan has not clearly specified what constitutes the “surround”, so it remains unclear how to apply the model to arbitrary scenes.

Decomposition models

Decomposition models were developed from Helmholtz’s recognition that our perception resembles components of the distal stimuli. Decomposition models suggest that the visual system perceptually decomposes luminance values into representations of their causal sources (e.g. reflectance, illumination, and shape). This process is often referred to as *intrinsic image analysis*, but does not necessarily imply a strict "inverse" of the image formation process, i.e. the result need not be veridical. Compared to theories outlined in previous sections, one major advantage of the intrinsic image analysis approach is that the process reflects our perception of the world – we get distinct impressions of reflectance properties, illumination, and object shape. Decomposition models explore *how* the visual system might compute intrinsic images.

Intrinsic images and layers

Adelson and Pentland (1996) described the problem of computing intrinsic images using a workshop metaphor. The workshop has three specialists: a painter, a lighting designer, and a metal bender. Any image can be reproduced using each worker's speciality alone while holding the other two specialties constant. The painter's solution accounts for all the information in an image with variations in grey shades of paint (assuming the scene is flat and uniformly illuminated; Figure 1.2A). The lighting designer's solution accounts for all the image information with variations in local illumination (assuming the scene is flat with constant reflectance; Figure 1.2B). The metal worker's solution accounts for all the image information with variations in shading caused by changes in the surface normal with respect to a single distant light source (assuming the scene also has constant reflectance and viewed from a single position for the surfaces to properly line up; Figure 1.2C). Although each specialist could construct a given image on their own, they can also cooperate and combine their specialities to produce the same image. The many-to-one mapping problem of intrinsic image analysis is portrayed by the fact that each specialty (painting, lighting, and bending) can be used together in an infinite combination of ways to achieve the same image. It is impossible to "undo" or reverse the entanglement of sources of image variance, so any decomposition approach must rely on constraints or assumptions imposed by the visual system on a given image. Below we discuss the various methods proposed by decomposition models for applying these constraints.

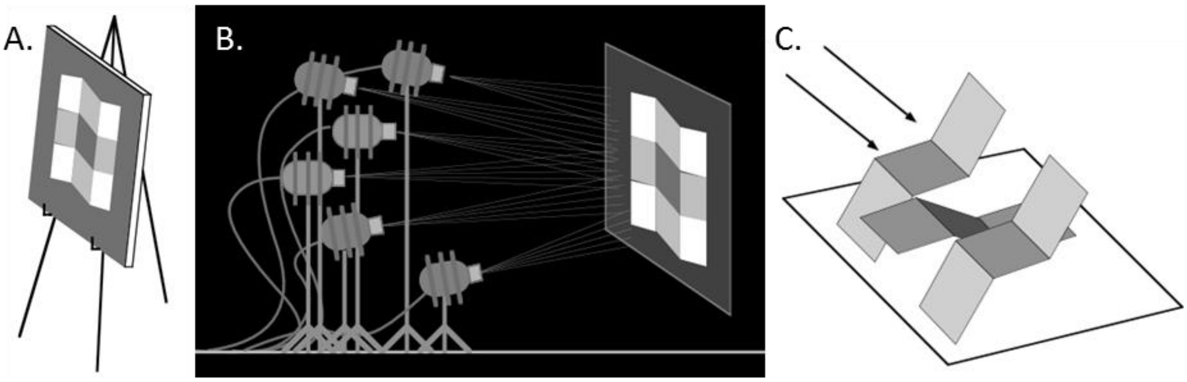


Figure 1.2. Workshop metaphor taken from Adelson and Pentland (1996). The same image can be produced solely by the painter (A), the lighting designer (B), and the metal worker (C). See main body text for details.

Layers models explicitly describe the phenomenal decomposition (or *scission*) of the retinal image into multiple images or separate layers. This concept was first applied in Metelli's (1970, 1974a, 1974b) work on transparency, which described the phenomenal decomposition of a single region of uniform luminance into two surface layers, one of which was transparent. Unlike brightness models, which have no capability of differentiating illumination or depth boundaries from reflectance boundaries, in layers models the interpretation of edges plays a key role in the recovery process.

In a seminal paper, Barrow and Tenenbaum (1978) proposed that the visual system recovers a set of layers called *intrinsic images* from the retinal image. Each intrinsic image maps a specific component of the scene such as reflectance, orientation, distance, or incident illumination at every point in the retinal image. Importantly, the different layers mutually constrain one another rather than existing independently. Barrow and Tenenbaum addressed the problem of computing intrinsic images by proposing that, "while isolated fragments of an image have inherent ambiguity, interactions among fragments resulting from assumed constraints can lead to a unique interpretation of the whole image" (p. 9). The concept of interactions among local features to reduce ambiguity is integral to the layers approach. In the intrinsic images model, constraints are imposed on different parts of an image such as regions, edges, and junctions. For example, regions have either smoothly varying intensities, which correspond to a curved surface with constant reflectance, or have constant intensity, which corresponds to a shadowed surface not directly facing the light source, or alternatively an illuminated planar surface. Edges either correspond to the boundary of a surface or to the boundary of a cast shadow, and this is classified according to the appearance of the regions on either side of the edge. In the model, T-junctions are caused by extremal boundaries (edges of objects), and provide constraints that can resolve the ambiguities in edges. Barrow and Tenenbaum suggested that when intrinsic characteristics are ambiguous, for example when regions lack information, the visual system relies on plausible estimates derived from assumptions about likely scene characteristics. While Barrow and Tenenbaum used relatively simple scenes, they argued that their model could be extended in a straightforward way towards real-world scenes. They proposed that other characteristics that are not present in their simple scenes, but are present in complex scenes, may form intrinsic images such as transparency, specularity, and luminosity.

In a series of papers, Anderson and colleagues (Anderson, 1997, 1999, 2003a, 2003b; Anderson & Winawer, 2005, 2008; Anderson & Khang, 2010; Anderson et al., 2011) have presented a model of layered image decomposition (which they have termed *scission*) based on geometric and photometric relationships that occur at contour junctions (where edges meet causing X, T, or I junctions) and terminations of gradients. Geometric relationships arise from the geometric continuity of targets and their surrounds. Photometric relationships refer to the consistency of polarity relationships of the borders separating targets from their surrounds (Anderson & Winawer, 2008). When the contrast polarity of at least two aligned edges is preserved, i.e. they both change from light to dark, or dark to light, this meets the conditions for the image to perceptually divide into causal layers, for example a transparent layer and an underlying surface, or the reflectance of a surface and the prevailing illumination (Adelson & Anandan, 1990; Beck et al., 1984; Metelli, 1970, 1974a, 1974b). Anderson and colleagues proposed a *transmittance anchoring principle* (TAP; Anderson, 1999) which states that “the visual system treats the highest contrast image regions as regions in plain view and only infers the presence of transparent surfaces if there are spatial or spatio-temporal (Anderson, Singh, & Meng, 2006) perturbations in the contrast magnitude along contours, surfaces, or textures” (Anderson & Winawer, 2008, p. 5). A transmittance value of 1 (completely opaque) is assigned to regions in plain view, and transparent layers are scaled relative to this anchor (Anderson & Khang, 2010).

Anderson and Winawer (2005, 2008) demonstrated that the way an image is parsed into layered image representations can have a large effect on lightness perception, in that the way the image is decomposed determines how luminance is partitioned between different layers. Anderson and Khang (2010) showed that destroying percepts of transparency by adding a ring around the target destroys these shifts in perceived colour caused by the transparency.

Work on transparency has demonstrated that when scission occurs, it can induce profound shifts in perceived lightness. However, Anderson and colleagues have stated that not all lightness effects are necessarily the consequence of decomposing images into layered representations (Anderson & Winawer, 2008). For example, it remains to be determined whether decomposition into layers is involved in lightness computations when all surfaces are under the same illuminant and no transparent surfaces are present. Illumination estimation models, which are described in the next section, advocate the idea that the visual system decomposes images into representations of reflectance and illumination.

Illumination estimation

Illumination estimation theories are broadly consistent with scission theories in that they advocate an explicit decomposition of luminance into illumination and surface reflectance. The decomposition theories discussed so far have been ambivalent to the order of lightness and illumination computations. However, proponents of illumination estimation have suggested that the visual system uses estimates of the illumination to subsequently (or concurrently) compute surface lightness (e.g. Logvinenko, 1999, 2003; Logvinenko et al., 2005; Logvinenko & Ross, 2005; Ikeda et al., 1998; for colour see e.g. Kraft & Brainard, 1999; Kraft et al., 2002; Lee, 1986). This is similar to Helmholtz's views.

An advantage of the illumination estimation approach is that forming a representation of the illumination conditions from information-rich areas of a scene can potentially constrain lightness estimates of surfaces in remote areas that lack information and are thus ambiguous. This would be particularly useful in 3D scenes. For example, Ikeda et al. (1998) presented the concept of the recognised visual space of illumination (RVSI), which is described as the state of an observer's recognition for illumination in the space between objects. They proposed that once the RVSI is constructed by the visual system, it determines lightness judgements of objects in the room. However, the question remains as to how the visual system forms the RVSI in the first place. Like the circularity inherent in Helmholtz's ideas, to generate a representation of the illumination from complex areas of a scene, the reflectance of those surfaces must be known.

A number of authors have proposed that the visual system uses cues such as shading, shadows and specular highlights to estimate the spatial and spectral distribution of the illumination in 3D scenes (Boyaci et al., 2003, 2004, 2006a, 2006b; Doerschner et al., 2007; Kraft, et al., 2002; Maloney, 2002; Ripamonti et al., 2004; Snyder et al., 2005; Yang & Maloney, 2001). Similar to the RVSI proposed by Ikeda et al. (1998), these authors have suggested that the visual system generates an estimate of the scene illuminant, which they term an equivalent illumination model (EIM). An EIM may contain estimates of the direction, intensity, chromaticity, and diffuseness of the light source[2]. These estimates are inferred by applying algorithms to observers' lightness judgments of flat surfaces that vary in orientation or depth with respect to a light source in a 3D scene containing various other objects. It has been suggested that various cues to the illumination, such as specular highlights, shadows, and shading provide the visual system with information about the EIM (Boyaci, et al., 2006a, 2006b; Kraft et al., 2002; Maloney, 2002; Snyder et al., 2005; Yang & Maloney, 2001). Similar to the RVSI, an EIM can be generalised to surfaces in the scene that lack explicit illumination cues, and used to constrain estimates of the colour or lightness of those surfaces. This will be explored further below.

---

[2] In traditional colour and lightness constancy research, the term "EIM" has been used to refer to an estimate of the illumination chromaticity/intensity inferred by the visual system. Here Boyaci and colleagues have posited that the internal representation of the illumination might include both spatial and spectral properties. Although most models that assume an EIM do not include all of these properties, the use of the term EIM in this thesis explicitly refers to Boyaci and colleagues' definition.

Relating Existing Lightness Theories to Real-World Scenes

Much of the field of lightness perception has been dominated by simple, matte displays, most of which are two-dimensional centre-surround or Mondrian-like images like those in Figure 1.1. This is reflected in experimental evidence that supports the various theories and models outlined above. For example, Wallach's ratio principle was developed from an experiment where observers were presented with two centre-surround displays, similar to the simultaneous contrast display in Figure 1.1A, but comprised of a circular disk surrounded by an annulus. Each annulus was different in lightness, and the observers adjusted the luminance of one of the disks until the two disks appeared equal in lightness. The results showed that the disks were perceived to be equal when the luminance ratios of the two displays were almost equal. From this Wallach concluded that lightness is determined by the luminance ratio between a surface and its adjacent surround.

Early contrast theorists like Wallach justified such an experimental approach because they believed that the effects of local stimulation could generalise to the entire visual field. However, this assumption has been undermined by studies that show that remote areas of a scene influence the lightness of a target patch, not just its immediate surroundings (Adelson, 1993, 1995, 2000; Hillis & Brainard, 2007; Hochberg & Beck, 1954; Logvinenko, 1999; Logvinenko & Ross, 2005, Williams et al., 1998). Furthermore, there are many examples of reverse contrast effects where the shift in the lightness of a patch is in the opposite direction to that predicted by contrast, such as in White's illusion (White, 1981; also see Agostini & Galmonte, 2002; Bressan, 2001; Economou et al., 1998). Spatial filtering, filling-in and edge integration models take into account the relationship between a target area and its larger context. However, these approaches do not account for the effect of perceived depth on surface lightness. Various studies have shown that the lightness of a test patch changes with perceived depth, even when retinal location remains constant (Gilchrist, 1977, 1980; Schirillo et al., 1990). Furthermore, they are not capable of differentiating different physical sources of image structure, for example discriminating reflectance edges from illumination edges. The predictions of these models are restricted to flat, matte, scenes and are thus difficult to extrapolate to structurally complex natural scenes.

Framework models have attempted to account for the effects of scene structure on perceived lightness by suggesting that lightness values are computed within frameworks divided by depth edges or shadows. However, it is unclear how to define frameworks in images generated by most natural scenes, where there is continuous variation in image structure caused by shading of 3D objects. Intrinsic image or layers models are perhaps the most effective class of models at tackling the issue of how scene layout affects lightness computations (or a transparent surface overlaying another surface). Decomposition models in general have been accused of not accounting for errors in lightness perception because they predict veridical perception (Gilchrist, 2006). This may be true of some models, however others like Anderson and colleagues (Anderson, 1997, 1999, 2003a, 2003b; Anderson & Winawer, 2005, 2008; Anderson & Khang, 2010; Anderson et al., 2011) have only suggested that the visual system decomposes images, not that it does so veridically. Furthermore, EIMs actually model the pattern of errors made by observers when making lightness judgments. Therefore, decomposition models cannot be rejected on the basis that they do not account for errors in lightness.

With a few exceptions, the scenes used in most lightness studies are impoverished, in that light emitted from a surface contains essentially no diagnostic information about the light field (Maloney, 1999), which is defined as the spectral power distribution of light arriving from every direction at every point in the scene (Gershun, 1936/1939). Matte surfaces absorb incident light and re-emit it in all directions, so information about the source is lost (Maloney et al., 2011), especially if those matte surfaces are smooth and flat. Such displays have been used because they allow rigorous control over extraneous variables. However, disentangling the contributions of reflectance and illumination in Mondrian worlds is least constrained, and hence a maximally ill-posed problem. As mentioned earlier, most natural surfaces are made of materials that are not characterised by purely diffuse reflectance but rather contain microstructure that generates both diffuse and specular reflections, and mesostructure (medium-scale surface relief), which interacts with light to create a more complex light field. One model that does exploit information about the light field is the equivalent illumination model (EIM), which is a representation of the light field (also referred to as the visual light field).

The EIM addresses many limitations of previous theories. This approach is capable of dealing with more natural, information-rich 3D scenes compared to brightness and framework theories. An advantage over layers models, which are ambivalent about the order of lightness and illumination computations, is that an EIM formed from information-rich areas of a scene is theoretically capable of extending to other areas of the scene. Thus, lightness can be equally constrained for surfaces that lack potentially diagnostic information about the illumination. Finally, equivalent illumination models alleviate the circularity of the illumination-estimation approach by specifying what information in the image, or which image cues, the visual system might use to form a representation of the light field. Compared to other theories of lightness perception reviewed, the EIM appears to be more generalizable to the information-rich scenes normally encountered by humans. The next section will evaluate evidence that the visual system has access to various components of the light field as suggested by EIMs, and will consider additional image information not present in EIM studies that may assist light field estimations. We also evaluate whether resolving components of the light field could be used to constrain estimates of surface lightness.

Potential Information about Lightness Available in Real-World Scenes

Proponents of EIMs are uncommitted to the nature of the representation of the illumination, meaning that observers do not need to have perceptual access to representations of the light field. However, there is some evidence that the visual system can actually resolve components of the light field (e.g. the direction and intensity of the light source) based on information in the image. Below we describe how various image cues in natural scenes could provide information about the light field, and how this could potentially constrain estimates of surface lightness.

Shading, cast shadows and specular highlights

Information about the direction of a light source can theoretically be obtained through shading on curved objects whose diffuse shading is brightest at points where the surface normal (perpendicular to the tangent) is directly facing the light source

(Lambert's cosine law; Lambert, 1760). Directional information may also be given by cast shadows, which are cast away from light sources, and specular highlights, where the surface normal bisects the angle between the direction of the light source and the specular highlight. The brightness of specular highlights may also provide information about the relative intensity of the illumination for objects in plain view versus in shadow. The clarity and sharpness of these reflections is determined by how much incident light is scattered. Less scattering leads to shinier (glossy) materials, with a purely specular object (zero scattering) acting as a mirror. Unless the object is purely specular (i.e. a mirror), only a portion of incident light is reflected specularly, so estimates of the intensity of the illumination would not be veridical without knowing that proportion.

If the visual system forms a representation of the light field, then this representation would be better constrained with the additional information given by shadows, shading, and specular highlights, relative to flat, matte displays. However, there would still be ambiguities about the relative strength of directional and diffuse components of the illumination, and the strength of the specular and diffuse components of a surface's reflectance. If the visual system uses the extra information available in complex scenes, it would still have to impose assumptions or priors about illumination, objects, and surfaces to reach an estimate of these various components of the light field. In EIMs, algorithms are applied to observers' psychophysical data (their lightness judgments) to retrospectively calculate what their estimations might have been. Regardless of what these assumptions might be, or how they might have been formed over time (through learning or evolution), the studies below suggest that observers do form a representation of the light field. The ambiguities are reflected in the pattern of errors made by observers (e.g. Boyaci et al., 2006a).

Koenderink et al. (2007a) conducted an experiment to see whether observers were sensitive to different aspects of the illumination in a scene that contained 3D objects with shading and cast shadows. They photographed a 3D scene containing uniform-albedo clay penguins standing in a circle under different lighting conditions (Figure 1.3A-C). The scene was viewed binocularly and observers judged the fit of a gauge object inserted at various locations (Figure 1.3D-H). They were asked to "make the test sphere appear like it fits into the scene" (p. 1597) by adjusting a number of lighting parameters including the direction (azimuth and elevation) of the light source, the diffuseness of the beam, and the intensity of the illumination. With a few exceptions, the results showed that settings of each parameter were reliable within and between observers, meaning that observers were able to identify the various components of the light field consistently. This suggests that observers formed a consistent impression of the structure of the light field, even at locations in empty space remote from other objects. Pont & Koenderink (2007) similarly showed that observers can estimate the direction and diffuseness of the illumination from the appearance of objects. These studies support the idea proposed by proponents of EIMs that the visual system can generate estimates of the light field through various image cues.

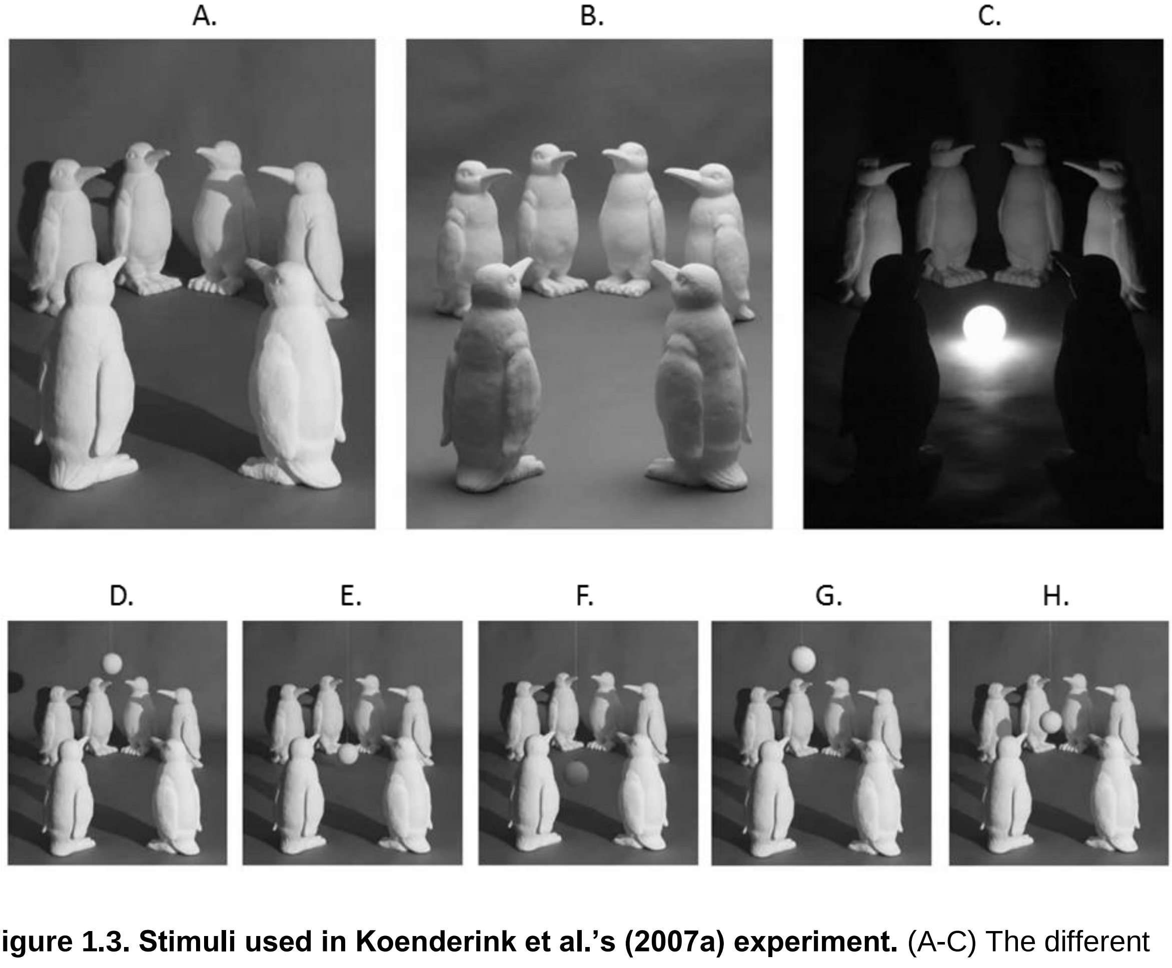
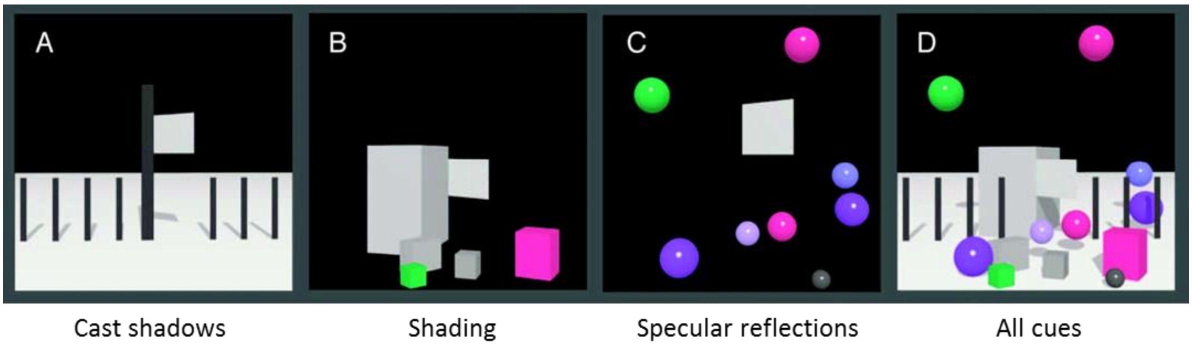


Figure 1.3. Stimuli used in Koenderink et al.’s (2007a) experiment. (A-C) The different lighting conditions that were used in the experiment. (D-H) The various locations of the gauge figure for the scene under illuminant (A).

If the visual system has a representation of the intensity and direction of the illumination, then it is potentially possible to take into account a target's orientation with respect to the light source and estimate surface lightness. For example, Boyaci et al. (2006a) had observers judge the lightness of a flat mate surface that varied in orientation in a stereoscopically viewed virtual scene, which either contained or lacked cast shadows, shading, and specular highlights (Figure 1.4). The observers' task was to match the lightness of the test surface within the scene to a nearby lightness scale. When each potential cue was presented in isolation (Figure 1.4A-C), observers' lightness judgments appeared to take 3D surface pose into account. Furthermore, observers' lightness estimates were more reliable when several cues to the illumination were present (Figure 1.4D) than when each cue was presented in isolation. Boyaci et al. (2006a) concluded that the visual system uses these cues to derive information about the illumination to constrain estimates of surface albedo. Similar arguments were made by Snyder et al. (2005), who assessed observers' ability to compensate for illumination gradients in binocularly viewed virtual scenes (Figure 1.5). In this study, observers adjusted an adjustable surface in the near room to match the lightness of a test surface that varied in depth from trial to trial. The authors found that observers judged lightness more veridically when the scene contained floating, glossy spheres (Figure 1.5C) than for similar scenes that contained no specular cues (Figure 1.5B). The experiments by Boyaci et al. (2006a) and Snyder et al. (2005) suggest that the visual system may use contextual cues such as specular highlights, shading and shadows to estimate components of the EIM, which generalises to areas of the scene where these cues are absent, thus constraining lightness judgments.



Cast shadows Shading Specular reflections All cues

Figure 1.4. The cue conditions in Boyaci et al.'s (2006a) study. (A) Cast shadows only condition; (B) Shading only condition; (C) Specular highlights only condition; (D) All cues condition.

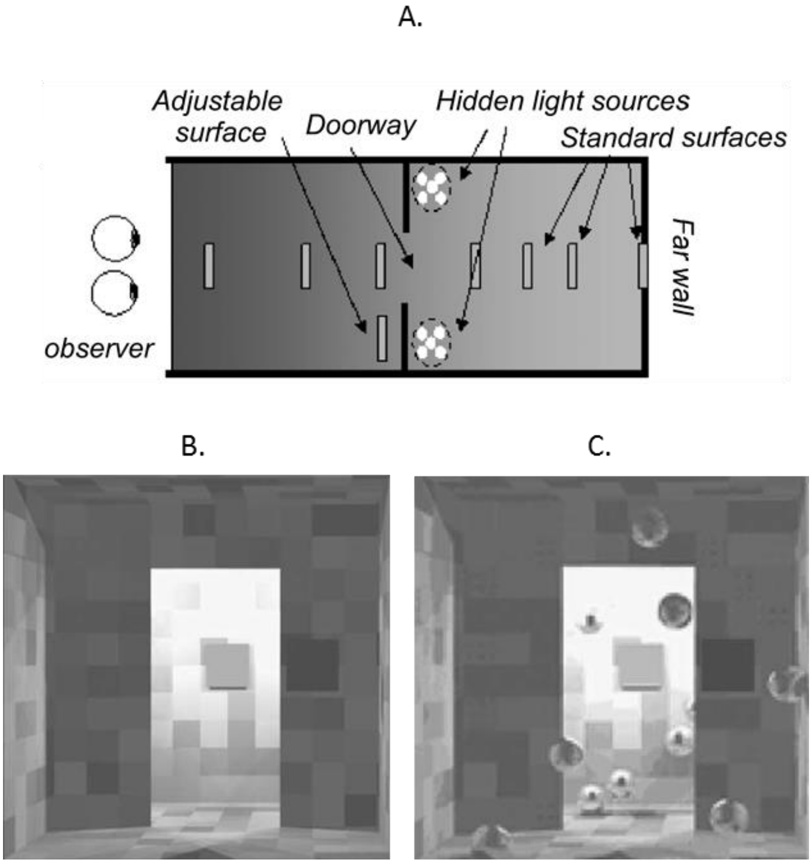


Figure 1.5. Setup of Snyder et al.'s (2005) experiment. (A) Bird's eye representation of the scene setup. (B) Observer's view of the scene without glossy spheres. (C) Observer's view of the scene with glossy spheres.

Texture caused by complex mesostructure

The 3D scenes in EIM experiments better represent real-world scenes compared to traditional Mondrian stimuli (e.g. Cataliotti & Gilchrist, 1995; Gilchrist et al., 1999; Land & McCann, 1971). However, they still resemble simple "toy-worlds" containing smooth objects and simple lighting conditions rather than the different materials, textures, and complex light fields often encountered real-world scenes. For example, the large number of purely specular spheres used in Snyder et al.'s (2005) study, while providing illuminant information in their experiment, are not likely to be present in natural scenes. Furthermore, even smooth matte surfaces in real-world scenes, such as plaster or concrete, are usually rough on a scale that is resolved by the visual system (Koenderink et al., 2007b). Texture created by illuminated roughness may provide the visual system with directional information about the light source in the form of *illuminance flow* (Koenderink et al., 2007b; Pont & Koenderink, 2003). Illuminated texture caused by surface roughness creates local luminance perturbations that look like "dipoles" (a juxtaposition of a light and dark blob; Koenderink et al., 2003). The "dipole vector" (e.g. pointing from the dark to the light side) points in the direction of the illumination. Thus averaging over these dipoles (discarding the sign) can form a strong cue to the direction (azimuth) of the illumination, albeit with 180° ambiguity (Koenderink et al., 2003). Indeed, it has been found that depending on the texture, observers can accurately estimate the direction of illuminance flow (Koenderink et al., 2007b; Koenderink et al., 2004). This information about the direction of the illumination could potentially constrain lightness estimates of a remote patch (like in EIM experiments) in the same way as shading where the luminance of the patch falls off with the cosine of the angle between the direction of the incident light and the surface normal (Lambert's cosine law; Lambert, 1760). However, like with shading from shape, there are ambiguities about strength of shading with surface relief (higher or deeper relief leads to darker shading) and angle of illumination (grazing angles also lead to darker shading). Again, the visual system would need to impose assumptions when estimating lightness from illuminance flow.

Rough surfaces are also affected by interreflections and vignetting (the surface itself partially blocking the primary light source to other parts of the surface; Pont & Koenderink, 2003). If the visual system forms a representation of the light field that constrains lightness estimates, then the extent to which interreflections counteract vignetting effects (Thompson et al., 2011, p.218) may be important in calculating surface lightness. Studies have demonstrated that 3D scenes even with only one reflectance contain cues to disentangle the relative contributions of surface reflectance and illumination. Motoyoshi et al. (2007) and Sharan et al. (2008) had observers view photographs of matte and glossy uniformly painted surfaces containing complex mesostructure that were equated for mean luminance. When these surfaces were viewed in isolation, observers' lightness judgments were positively correlated with true surface reflectance although these data exhibited a regression to the mean (white surfaces appeared darker than they were whereas black surfaces appeared lighter). Gilchrist and Jacobsen (1984) found that shading contrast generated by secondary reflections (*interreflections*) provides the visual system with information about surface lightness. They presented observers with two uniformly painted scenes with identical 3D structure except that one was painted white and the other black. In one part of the experiment they lowered the illumination of the white room so that the luminance at each point was lower than in the black room. The brightly lit black scene appeared mid-grey and the dimly lit white scene appeared light grey. The white scene appeared lighter than the black scene despite having lower average luminance. Thus observers could distinguish the two rooms independent of their brightness.

The luminance profiles of Gilchrist and Jacobsen's (1984) scenes revealed that the contrast of the shadows in the black room was stronger than in the white room. Since the scenes were structurally identical, the difference in shadow strength had to be caused by the amount of incident light reflected by surfaces in each room. Reflected light from one surface is capable of indirectly illuminating other nearby surfaces. White surfaces reflect up to 90% of incident light, whereas black surfaces reflect as little as 3%. Gilchrist and Jacobsen pointed out that even after two reflections, 81% of the light would remain unabsorbed in the white scene and could be used for further interreflections. After two reflections in the black room, only 0.09% of the light would remain unabsorbed, which is a negligible amount to contribute to further interreflections. Therefore, much more reflected light indirectly illuminated surfaces in the white scene, leading to "filled-in" shadows compared to the black scene, which had darker, more pronounced shadows. The authors suggested that the amount of shadow filling-in provided additional information about surface lightness. Note that this information is relevant in rooms or surfaces with a single albedo.

Unanswered Questions

The studies presented above suggest that there is information present in natural scenes that has not been captured by current models of lightness perception. The experiments supporting EIMs come closest to utilising the information available in information-rich scenes such as specular highlights, shadows, and shading information, but they miss out on additional information that might be provided by complex mesostructure, interreflections and illumination flow. Furthermore, there is currently no systematic attempt to assess the relative importance of different illumination cues directly. Boyaci et al. (2006) found that including all cues to the illumination in a scene was better than having one cue present, but they did not test all cue combinations to determine their relative contributions. An aim of the present thesis is to test whether information created by complex mesostructure helps the visual system constrain estimates of surface lightness, and to also systematically assess the relative importance of potential cues to the illumination.

Though EIMs provide many benefits over other theories, one potential confound is that supporting experiments do not control for low-level perceptual mechanisms that could potentially account for the results. It is unclear whether the preceding results arose from an estimation of the illumination field, as suggested by EIMs, or whether they were the result of low-level differences in image content. The addition of specular highlights and/or the presence of shadows and shading would have changed the range and distribution of luminance values in the image, which may have influenced lightness judgments. Previous studies have not attempted to tease apart low-level explanations involving luminance and contrast distributions from the mid-level explanations involving representations of the light field.

Finally, studies using rendered scenes often overlook the fact that illumination in natural environments has a high dynamic range of intensities, and the light field fills the entire space of a scene (Thompson et al., 2011, p. 205). Most studies use a combination of point source and diffuse lighting. Similar to the use of Mondrian displays, this approach is used as it allows rigorous control over variables in the experiment. Unfortunately this is at the cost of being unable to reproduce the complexity of a realistic light field. An experiment by Ruppertsberg and Bloj (2007) demonstrates that experimental control need not be compromised by the complexity of lighting in a scene. They replicated Gilchrist and Jacobsen's (1984) findings using virtual rather than physical scenes. By using virtual scenes they were able to better control aspects of their stimuli such as equating the average luminance of the black and white rooms. This demonstrates a level of control that can be exerted in virtual scenes but that is not possible for physical scenes and stimuli.

Approximating natural light fields can be achieved in rendered environments using image-based lighting with light probes (Debevec, 1998; Thompson et al., 2011, p. 202). Light probes are photographs that capture the complexity of light arriving at a single point. This information can then be mapped into the rendering environment and used to illuminate scenes, creating photo-realistic effects (Debevec, 1998). Therefore, it is possible to bring naturalistic information to controlled settings. The following experimental chapters attempt to take a first step to adding the control that is necessary to tease apart low- and mid-level contributions to lightness perception while incorporating the complexity of natural surfaces.

The experiments in the following chapters were designed to (a) assess the relative influence of different image cues and levels of image complexity on lightness perception and (b) tease apart low-level contextual influences on perceived lightness from mid-level explanations that invoke the estimates of the illuminant or layered image decomposition. Centre-surround displays like those in Figure 1.1A have an important history in lightness perception, yet they are one of the most ambiguous types of stimuli. In the following experiments we have introduced some complexity into these displays, allowing control over extraneous variables while making these artificial environment as natural as possible. The stimuli used are virtual centre-surround displays with flat, matte central test patches and various surround types rendered under a natural illumination field. In Chapter 2 we test the hypothesis that lightness constancy of the test patch will improve when the surround contains complex mesostructure (high surface relief) and specular highlights (gloss) compared to when the surround lacks this information. In a series of control experiments, we test whether any improvements in lightness constancy arise from information about the light field or whether they can be attributed to low-level attributes, such as the distribution of luminances and contrasts across different spatial scales. Chapter 3 explores whether low-level mechanisms are sufficient to explain lightness constancy of surfaces under changing illumination level, where lightness matches do not equate brightness matches. We also seek to determine whether the rocky surfaces used in Chapter 2 provide enough information to differentiate surfaces based on shading caused by interreflections. We therefore manipulate the number of interreflections rendered in the scenes and observe how lightness constancy varies as a function of number of interreflections. Experimental Chapter 4 aims to explain the qualitatively different pattern of results observed for the homogeneous centre-surround displays in Chapter 2. Again, we seek to determine whether the effects observed are the result of layered image decomposition (scission) or low-level (contrast) mechanisms.

Chapter 2. Do Surface Reflectance Properties and 3D Mesostructure influence the Perception of Lightness?

The aims of this experimental chapter are twofold. First, we aim to systematically assess the relative importance of potential cues to the illumination on lightness constancy, specifically specular highlights and shading caused by complex mesostructure. Second, we aim to determine whether lightness effects are predominantly the result of the visual system forming a representation of the illumination (mid-level explanation), or predominantly the result of differences in luminance and contrast of the surrounds between the various conditions (low-level explanation). The stimuli used in the following experiments were graphically rendered centre-surround displays with flat, matte central test patches and various surround types, rendered under a natural illumination field (Figures 2.1, 2.5, 2.9, and 2.15). The stimuli provide a number of advantages over displays used in previous lightness studies. For example, in some conditions the surfaces contained complex mesostructure, which generated additional information not present in EIM studies to potentially constrain lightness. Additionally, surfaces were embedded in a natural illumination field, which makes them more comparable to surfaces found in the real world. Importantly, these stimuli allow the results to be directly compared to simple centre-surround displays and Mondrians that have dominated the lightness literature. Finally, the control stimuli allow us to better tease apart low-level and mid-level explanations of lightness effects compared to previous studies. The experiments in this chapter are a first step to demonstrating the potential of using rendered environments to conduct controlled experiments with information-rich stimuli.

Experiment 1A and 1B: Varying Surface Relief and Gloss Level of the Surround

In Experiment 1 observers performed lightness judgments on the central test patches embedded in four surround types: low-relief (flat) matte, low-relief (flat) glossy, high-relief (rocky) matte, and high-relief (rocky) glossy (see Figure 2.1). The low-relief (flat) surfaces were similar to homogeneous centre-surround displays traditionally used in the literature. Rendering the stimuli produced almost identical images for the flat matte and glossy conditions. For completeness, we included both conditions in Experiment 1A. However, the flat glossy condition was removed in Experiment 1B after verifying that the stimuli from both low-relief conditions produced indistinguishable images and essentially identical results. We hypothesised that if image cues generated by complex mesostructure and gloss help the visual system to constrain lightness estimates, then lightness constancy should be better for high-relief rocky surfaces compared to low-relief flat surfaces, and for glossy compared to matte surfaces. We hypothesised that lightness constancy should be worst when no image cues are present (flat surfaces) and best when all image cues are present (rocky glossy surfaces).

Methods

Observers

Experiment 1 included two populations of observer. Five observers participated in Experiment 1A. Observers AS (the author of this thesis), PM, and KT were experienced in psychophysical experiments. Observers DC and RS were inexperienced and paid $20 per hour for participation. These observers performed five repeats of each condition (see procedure).

Twenty undergraduate first-year psychology students at the University of Sydney participated in Experiment 1B. These observers were awarded course credit in exchange for participation and performed one repeat of each condition. All observers except AS from Experiment 1A were naïve to the aims of the study.

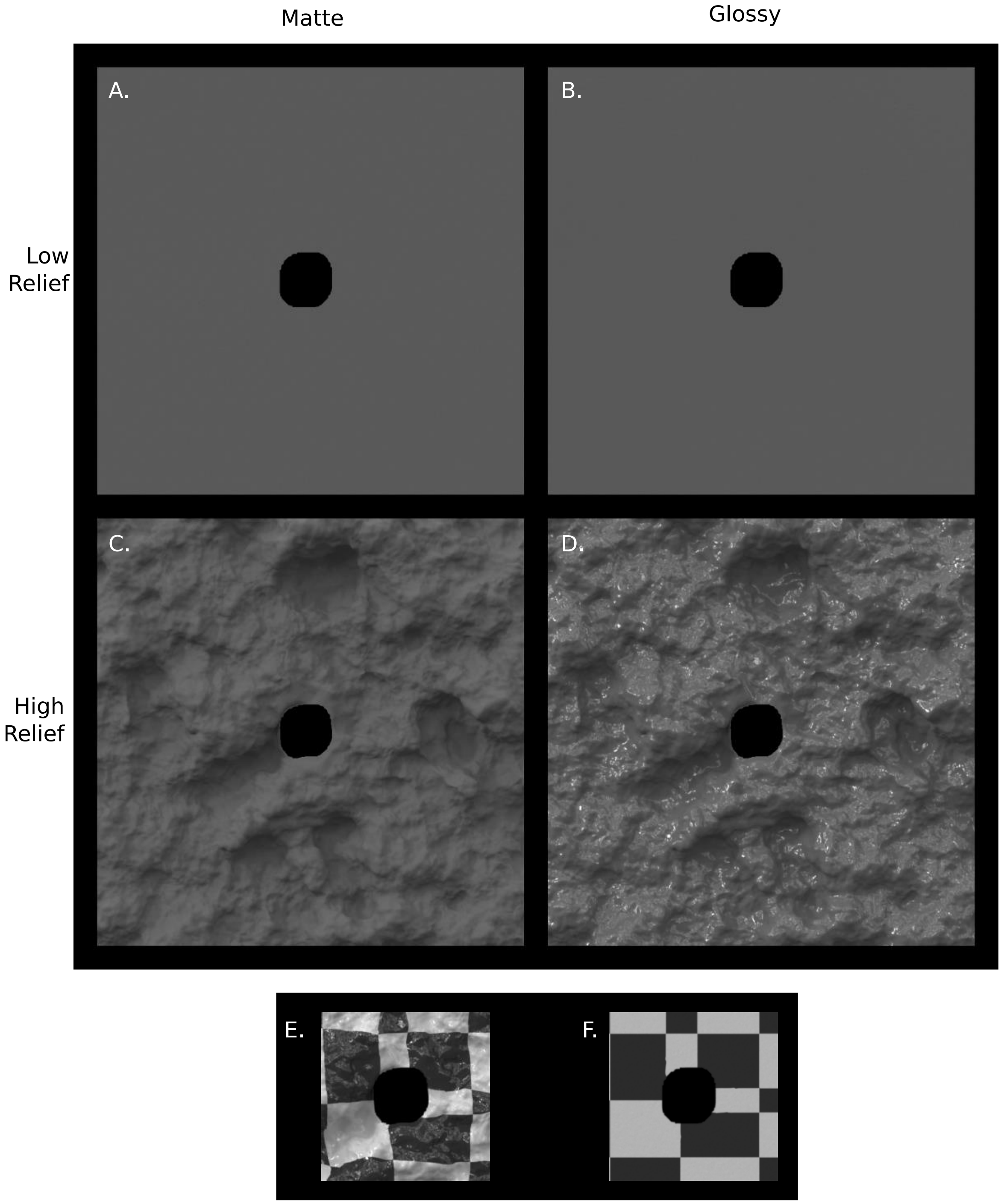


Figure 2.1. Computer rendered centre-surround stimuli used in the experiments. (A–D) Examples of target surfaces used in Experiment 1. All surrounds shown have equal reflectance (19.8%) but differ in their level of gloss and surface relief: (A) low relief (flat), matte; (B) low relief (flat), glossy; (C) high relief (rocky), matte; (D) high relief (rocky), glossy. Target central patches are shown in black but actually varied in albedo from trial to trial during experiments. (E) The adjustable surface used in Experiments 1, 2, and 4. (F) The adjustable surface used in Experiment 3. Observers moved a computer mouse left to incrementally decrease the albedo and right to incrementally increase the albedo.

Apparatus

Stimuli were presented on a LaCie Electron 22 Blue IV monitor running at a refresh rate of 75 Hz and with a resolution of 1280 × 1024 pixels, controlled by a Mac Pro computer running Mac OS X 10. Stimulus presentation and data collection were controlled by a Matlab (R2010a; Mathworks) script using the Psychophysics Toolbox (Brainard, 1997). Stimuli were viewed in a dark room at a viewing distance of approximately 70 cm. The carpet and walls of the room were black so that the only source of light came from the monitor on which the stimuli were displayed.

Stimuli creation in Blender (v. 2.6)

Stimuli were computer-rendered centre-surround displays (Figure 2.1). The displays contained flat, matte centres that varied in albedo and surrounds that varied in albedo, amount of surface relief, and gloss level. Surrounds had either low surface relief (flat; Figure 2.1A and 2.1B) or high surface relief (rocky; Figure 2.1C and 2.1D) and were either matte (Figure 2.1A and 2.1C) or glossy (Figure 2.1B and 2.1D). The surfaces were modelled in the open-source software Blender (v. 2.6). Each surface was created using an 800 × 800 mesh. The textures in each surround were generated with the displace modifier, a tool in Blender that displaces vertices in depth in a mesh based on the intensity of a texture. Various textures were used to deform the surfaces: the inbuilt cloud, marble, and stucci textures as well as textures from images of rocks and rough paper. The image of rough paper used is displayed in Figure 2.2. The rocky texture image can be found at http://junk-paris-stock.deviantart.com/art/macro-rock-texture-13-119245673. Table 2.1 shows the modifiers that were used for low- and high-relief surfaces and the order in which they were applied. Note that although the rough paper texture was used to displace vertices in the low-relief surrounds, this effect was extremely subtle, so the rendered images were essentially homogeneous. Also note that, although multiple textures were used to deform the high-relief surfaces, we refer to them as “rocky” because of their rocky appearance after rendering.

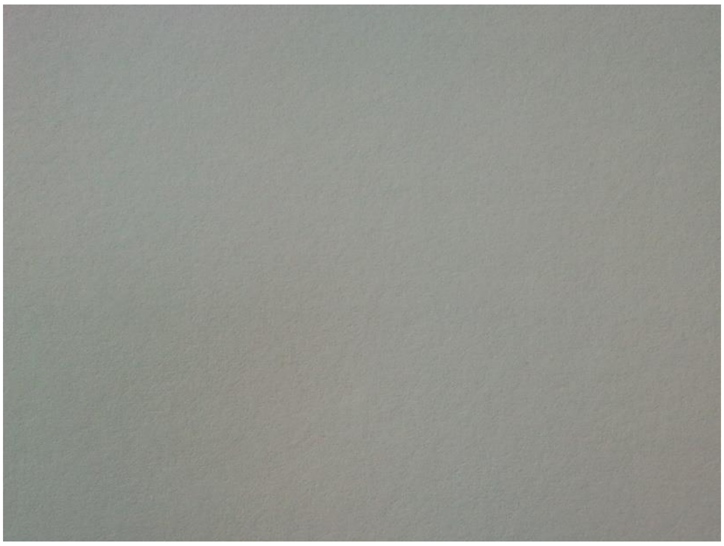


Figure 2.2. Paper texture used to deform low-relief (flat) surfaces.

	Properties	
Modifier	Low-relief surfaces	High-relief surfaces
Displace – stucci (inbuilt)	N/A	Strength 0.2
Displace – marble (inbuilt)	N/A	Strength 0.1
Displace – clouds (inbuilt)	N/A	Strength 0.07
Smooth	N/A	Factor 0.5; Repeat 50
Displace – rock	N/A	Strength 0.003
Paper	Strength 0.01	N/A
Smooth	Factor 0.5; Repeat 50	Factor 0.5; Repeat 50

Table 2.1. Modifiers used to create the effects in the low-relief (flat) and high-relief (rocky) surrounds. The order in which the modifiers are displayed is the order in which they were applied (left column, top to bottom).

Rendering software: RADIANCE

Stimuli were rendered using the RADIANCE rendering software (Ward, 1994), which simulates physical interactions between illuminants and surfaces. The surfaces were rendered using the Ward BRDF model, termed ‘‘plastic’’ in RADIANCE.[3] This model has five parameters: diffuse components R, G, and B; specularity (P_S, the proportion of light reflected by the specular component, uncoloured); and microroughness (α, which determines the amount of specular scatter). Centres were embedded in the surrounds and rendered as part of the same scene. Grey shades were assigned to the centre and surround regions by adjusting the diffuse reflectance parameters, keeping relative RGB values equal. Matte surrounds were assigned a specularity value of 0 and a roughness value of 0 whereas glossy surrounds were assigned a specularity value of 0.05 (5% of the light reflected from the surface is specular) and a roughness value of 0.01. Surfaces were rendered frontoparallel to the observer with two ambient reflections.[4]

All surfaces were illuminated by the “grove” light field from the Debevec Light Probe Image Gallery (Debevec, 1998; Figure 2.3). This light field is a high dynamic range (HDR) photograph of a real forest scene that captures the light arriving from every direction to a single point. The HDR image of the light field was mapped into the rendering environment and used to illuminate the surfaces. A grey scale version of this light field was created so that all surfaces were illuminated by achromatic light. To produce high-quality images, all surfaces were rendered 10 times larger than required and antialiased, resulting in HDR images of the surfaces with dimensions of 900 × 900 pixels. These HDR images were tone-mapped to fit the luminance range of the monitor. This was achieved by linearly compressing the diffuse component and nonlinearly compressing the specular component of the images (see next section).

---

[3] This material is not limited to the physical light-scattering properties of plastic; rather, it can be used for a wide variety of reflective materials, such as surfaces made of concrete, wood, paint, etc.

[4] Rendering with two ambient reflections allowed shadowed areas of the rocky surfaces to be indirectly illuminated by other parts of the surface as would occur in natural scenes. This was needed to test the hypothesis that observers use the amount of shadow “filling-in” as a cue to surface lightness.

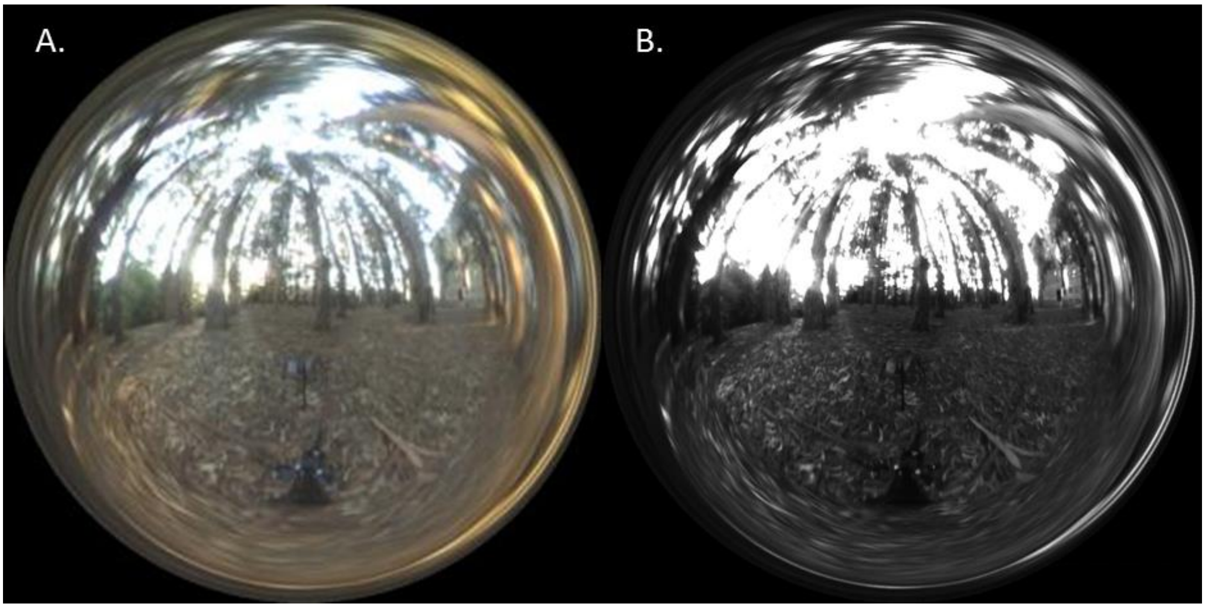


Figure 2.3. The “grove” light field used to illuminate surfaces used in the experiments. (A) Original image taken from http://www.pauldebevec.com/Probes/. (B) Black and white variant of the light field used in the experiments so that surfaces were illuminated by achromatic light. Note that the quality and dynamic range of the images displayed here is lower than what was used in the experiments to illuminate the surfaces.

Tone mapping of RADIANCE HDR images

The procedure used to tone-map each HDR image was as follows: The diffuse component was linearly compressed by transforming luminance values below 140 cd/m^2 with the equation

$$L_i^D = \frac{H_i^D}{H_{max}^D} \cdot L_{max}^D , \tag{2.1}$$

where L_i^D is the transformed luminance associated with the diffuse component for each pixel *i*, H_i^D is the original HDR luminance associated with the diffuse component for each pixel *i*, H_{max}^D is the maximum HDR luminance attributed to diffuse reflectance, and L_{max}^D is the maximum luminance assigned to diffuse reflectance in the tone-mapped (transformed) image. H_{max}^D was constant for all images and was equal to 140. L_{max}^D was constant for all images and was equal to 53.59. Thus, the brightest regions of diffuse shading in the tone-mapped image had a luminance of approximately 53.59 cd/m^2.

The specular component (HDR luminance values above 140 cd/m^2) was compressed nonlinearly to create smooth fall-off of luminance values that started at L_{max}^{D} (53.59 cd/m^2) and peaked at L_{max}^{S}, the luminance assigned to the brightest specular highlight and also the brightest luminance of the monitor (64.98 cd/m^2; see Figure 2.4). We achieved this by first subtracting H_{max}^{D} (140 cd/m^2) from each pixel and then transforming these values with the equation

$$L_i^S = \left(-e^{-\frac{1}{RSH_i^S}} + 1\right) \cdot R \,, \tag{2.2}$$

where L_i^S is the transformed luminance associated with the specular component for each pixel *i*, *R* is the luminance range of the specular highlights and is equal to $L_{max}^{S} - L_{max}^{D}$, *S* is the slope of the straight line from the linear transformation of the diffuse component and is equal to L_{max}^{D}/H_{max}^{D}, and H_i^S is the HDR luminance associated with the specular component for each pixel *i*. Finally, we added L_{max}^{D} to these specular values, and the result was a tone-mapped image with linearly transformed diffuse shading and nonlinearly transformed specular highlights (Figure 2.4).

The last step was to display the images using the eight-bit pixel values of the monitor. For this, we made a colour look-up table (CLUT) of luminance values corresponding to each eight-bit pixel value (0–255). Each luminance value in the tone-mapped image was transformed into its corresponding CLUT value.

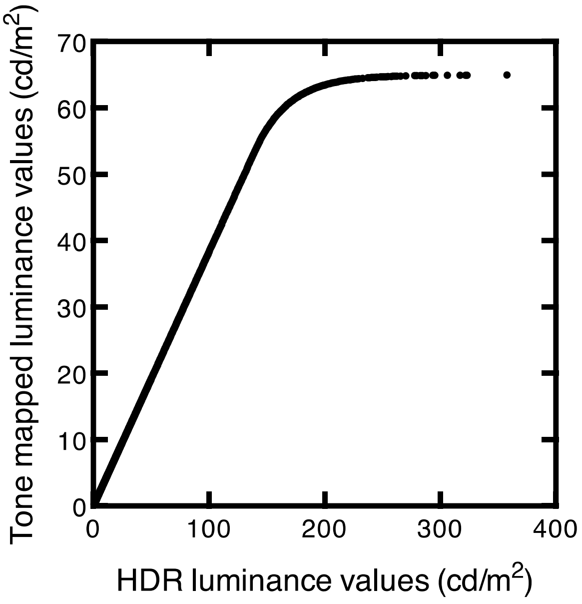


Figure 2.4. Transformation of HDR luminance values (x-axis) to tone-mapped luminance values (y-axis).

Procedure

Observers judged the lightness of flat target patches embedded in various surrounds via an asymmetric matching task (see Figure 2.1 for stimuli). In each trial, a target surface (14.88°) was presented on the computer screen. Below the target surface was a smaller surface with an adjustable test patch (5.83°). The surfaces were separated by 11.47° (centre to centre) and were presented against a black background. Observers were instructed to change the lightness of the flat central patch on the adjustable surface until it looked like it was the same lightness or painted with the same paint as the flat central patch on the target surface.

In Experiment 1A, target surfaces consisted of a flat, matte, central target patch surrounded by one of four surround types: flat and matte (Figure 2.1A), flat and glossy (Figure 2.1B), rocky and matte (Figure 2.1C), or rocky and glossy (Figure 2.1D). For each of these conditions, there were six different surround albedos ranging from black to white (see Table 2.2). This produced 24 surround conditions in total. For these 24 surround conditions, observers judged 13 to 15 test patch albedos. Table 2.2 shows specific test patch albedos included for each of the surround conditions, and Figure 2.5 displays examples of these test patch albedos for surround reflectance 19.8%. Eleven of the test patch albedos were standard for all surround conditions. We included a further two to four unique test patch albedos very close in lightness to each surround albedo. Two of these extra values were increments and two were decrements except for black surrounds, which had only two extra increments, and white surrounds, which had only two extra decrements. There were 344 conditions in total. Observers performed five repeats of each condition, resulting in 1,720 trials.

The flat, glossy condition in Experiment 1A was included for completeness. However, due to the specific viewing angle of the camera in relation to the surface and the light source, rendering matte and glossy flat surfaces generated indistinguishable images, thus producing almost identical results. For this reason, the flat glossy condition was removed in Experiment 1B. Additionally, observers did not perform repeats for any condition, resulting in 258 trials. All other aspects of Experiment 1A and 1B were the same.

Figure 2.1E shows the adjustable surface that was used for all conditions in Experiments 1, 2, and 4. The surround was a checkerboard surface with equal amounts of black (3% reflectance) and white (90% reflectance) bordering the central patch. The surface relief and gloss level of the surround were identical to that of the glossy, high-relief (rocky) test surfaces. Observers were able to adjust the albedo of the central patch by moving a computer mouse left and right. They could choose from 201 pre-rendered Munsell values ranging from zero to 10 in equal increments on the Munsell scale.

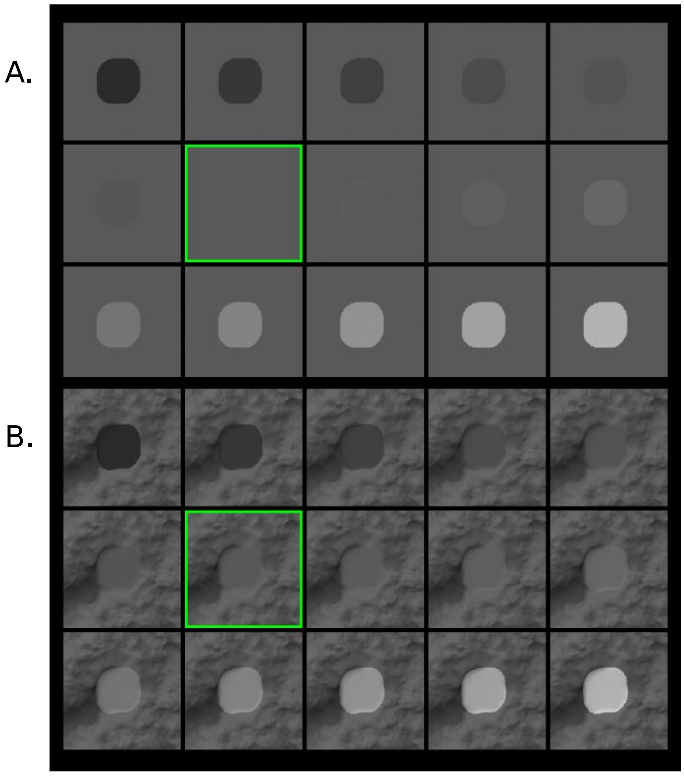


Figure 2.5. Test patch albedos for surround reflectance 19.8%. Test patches are embedded in the flat matte surround (A), and the rocky matte surround (B). Test patches increase in lightness from left to right and from top to bottom. The green square indicates the test patch that has the same albedo as the surround. The two values immediately darker and lighter than the surround were unique to this surround albedo. The other 11 test patch values were common to all surrounds. See Table 2.2 for the specific test patch values.

Surround % reflectance	Surround Munsell values	Surround luminance min (cd/m²)	Surround luminance max (cd/m²)	Extra test patch % reflectance	Extra test patch Munsell values	Extra test patch luminance values (cd/m²)
All	–	–	–	3, 5.2, 9, 13.7, 19.8, 27.2, 36.2, 46.8, 59.1, 73.4, 90	1.95, 2.75, 3.5, 4.25, 5, 5.75, 6.5, 7.25, 8, 8.75, 9.5	1.17, 2.25, 3.34, 5.13, 7.3, 9.83, 13.06, 17.01, 21.34, 26.40, 32.49
3	1.95	M: 0.031 G: 0.049	M: 2.2 G: 64.98	3.39, 3.82	2.1, 2.25	1.17, 1.49
9	3.5	M: 0.11 G: 0.12	M: 6.63 G: 64.98	7.47, 8.21, 9.84, 10.7	3.2, 3.35, 3.65, 3.8	2.56, 2.92, 3.67, 4.02
19.8	5	M: 0.32 G: 0.32	M: 14.29 G: 64.98	17.2, 18.4, 21.1, 22.6	4.7, 4.85, 5.15, 5.3	6.21, 6.56, 7.65, 8.38
36.2	6.5	M: 0.65 G: 0.65	M: 26.5 G: 64.98	32.4, 34.3, 38.2, 40.2	6.2. 6.35, 6.65, 6.8	11.64, 12.34, 13.77, 14.52
59.1	8	M: 1.28 G: 1.38	M: 43.44 G: 64.98	53.9, 56.5, 61.8, 64.6	7.7, 7.85, 8.15, 8.3	19.57, 20.32, 22.4, 23.51
90	9.5	M: 2.63 G: 2.75	M: 61.03 G: 64.98	83.1, 86.5	9.2, 9.35	30.07, 31.51

Table 2.2. Surround and centre patch reflectance and luminance values of the test surfaces used in the experiments. Notes: The first row shows the test patch values that were common to all surround types. The fifth column shows this in percentage reflectance, the sixth column shows this in Munsell values, and the seventh column shows this in luminance values. For all remaining rows, the first column contains the six reflectance values (percentage reflectance) used for the surrounds. The second column shows the values in column 1 transformed to the Munsell scale. The third and fourth columns show the luminance range of the surrounds (M = matte, G = glossy). The fifth column contains the extra two to four reflectance values (percentage reflectance) of the centre patches that were very close in lightness and unique to each surround. Two of these values were increments and two were decrements except for the black surround (3% reflectance), which contained only two extra increments, and the white surround (90% reflectance), which contained only two extra decrements. The sixth column shows the values in column 5 transformed to the Munsell scale, and the seventh column displays the luminance values of the test patches.

Results and discussion

The results from Experiment 1 are presented in Figure 2.6, which shows the average data of all five observers from Experiment 1A (left column) and the average data of the 20 observers from Experiment 1B (right column). Three trials (out of 5,160) from Experiment 1B were excluded from analyses due to observers accidentally pressing the button to set test patch lightness before they had finished making adjustments. The data revealed that test patch lightness settings were affected by surround type. The different surface relief and gloss level conditions gave rise to different patterns in the data. These patterns will be discussed in relation to lightness constancy below.

Results from the flat (low-relief) surround conditions

Figure 2.6 shows that there is a distinct difference in the shape of the data curves between the flat and rocky surround conditions (compare the top two rows to the bottom two rows, respectively). The data curves for the rocky conditions are relatively linear whereas the data curves for the flat conditions exhibit a large “step” at which test patch albedo passes through that of the surround. This sharp lightness change between low-contrast increments and decrements reflects a phenomenon termed *crispening* by Takasaki (1966), and we retain this terminology here. To emphasize the size of the step for each of the data curves in Figure 2.6, difference scores were obtained by subtracting the lowest contrast decrement settings from the lowest contrast increment settings. These difference scores are plotted in Figure 2.7. A lower score indicates a smaller step and therefore less crispening. The darkest (Munsell value 1.95) and lightest (Munsell value 9.5) surround conditions were omitted because they contained only increments or only decrements, respectively.

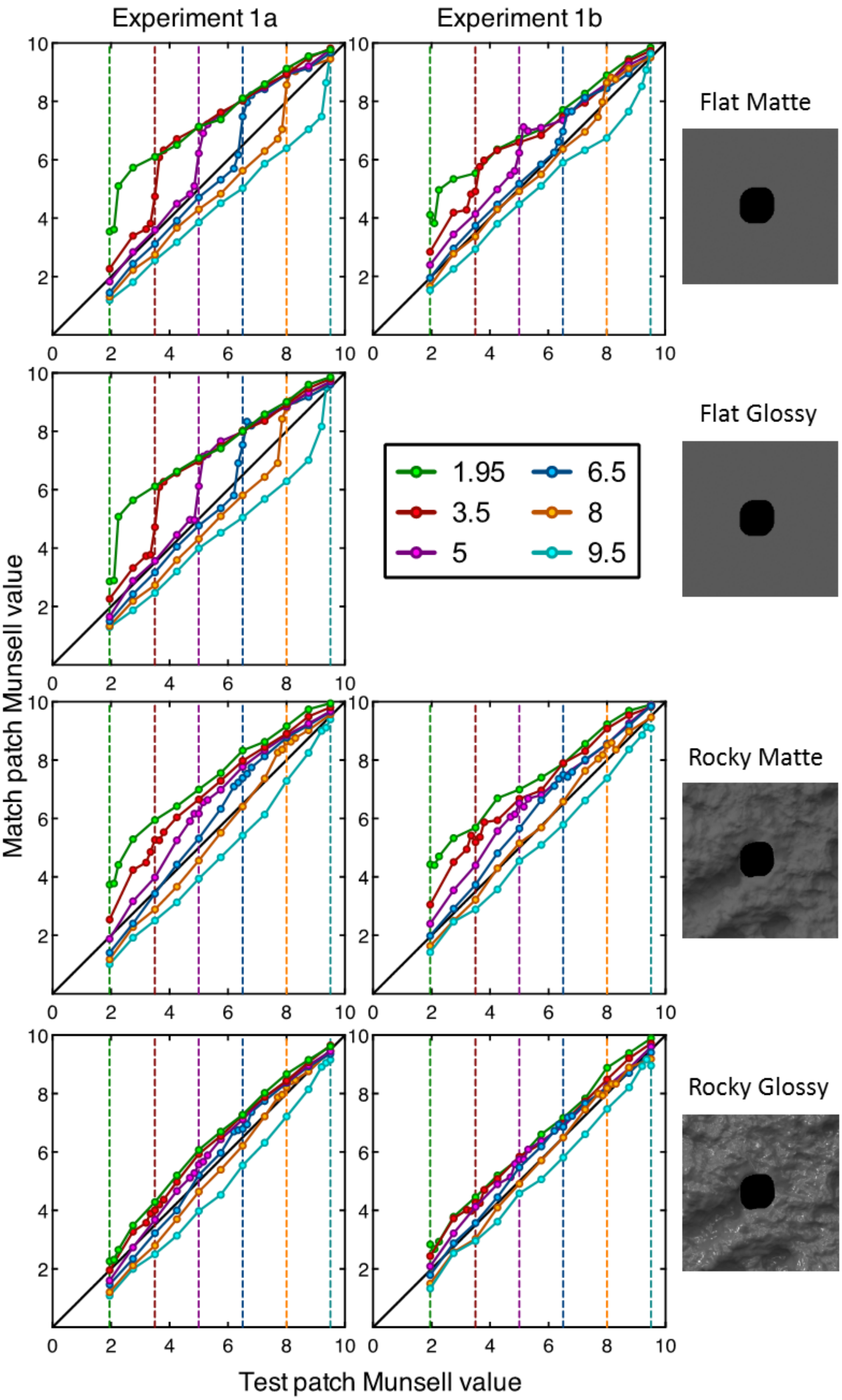

Experiment 1a
Experiment 1b
Flat Matte
Flat Glossy
Rocky Matte
Rocky Glossy
1.95
3.5
5
6.5
8
9.5
Match patch Munsell value
Test patch Munsell value
0
2
4
6
8
10

Figure 2.6. Average data for Experiment 1A (left panels) and 1B (right panels). Each coloured data curve represents test patch settings for a different surround albedo condition. The legend shows the Munsell values of each surround. For flat surround conditions (top three panels), there was an increment–decrement "step" (crispening) as the test patch albedo passes through that of the surround. This step was absent in the rocky surround conditions (bottom four panels). Comparing the rocky surround data, lightness settings were more consistent for the glossy condition (last row) compared to the matte condition (third row).

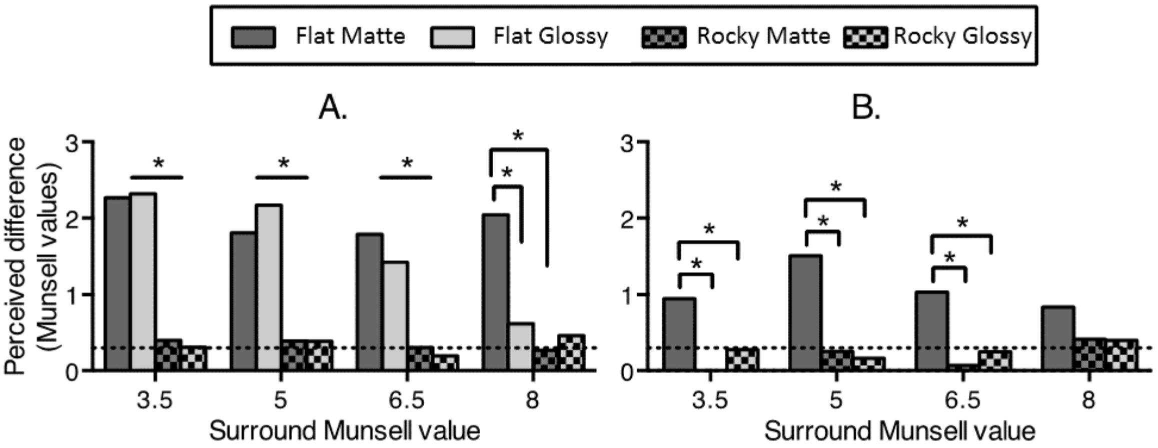


Figure 2.7. Increment minus decrement settings for Experiment 1A (A) and 1B (B). The horizontal dotted line represents the actual difference between increments and decrements. The solid bars show large increment–decrement steps for the flat surround conditions. The checked bars show that this step was eliminated in the rocky surround conditions (where surrounds contained complex mesostructure).

When the surround was essentially homogeneous (low-relief condition), increments appeared much lighter than decrements: For observers in Experiment 1A, the lowest contrast increments appeared, on average, about 2 Munsell values lighter than the lowest contrast decrements (see Figure 2.7A, solid bars); for observers in Experiment 1B, the difference was about 1 Munsell value (see Figure 2.7B, solid bars). The actual (simulated) increment-decrement Munsell difference was 0.3, indicated by the horizontal dotted lines in Figure 2.7. The checked bars in Figure 2.7 show that difference scores from the rocky surround conditions lie around this line. From these difference scores it clear that, for both sets of observers, crispening was only induced by the homogeneous flat surrounds. These observations were statistically verified for Experiment 1A by subjecting the difference scores for each surround albedo condition to a within-subjects two-way ANOVA with two levels of surface-relief (flat, rocky) and two levels of gloss (matte, glossy). *F* values and *p* values are displayed in Table 2.3. There was a main effect of surface relief for all surround albedo conditions. This confirms the observation that the increment-decrement step was larger when test patches were surrounded by homogeneous flat compared to rocky surrounds. There was one inconsistency for surround Munsell 8, at which the step seemed to be smaller for the flat glossy surround (Figure 2.7A). Indeed, for this condition, there was a main effect of gloss and an interaction between surface relief and gloss level. Follow-up tests[5] suggested that the flat glossy surround was not as effective in inducing crispening as the flat matte surround. However, Figure 2.6 (left column, second panel from the top, orange data points) clearly shows that there was strong crispening for this condition. Closer inspection of the stimuli revealed that, for surround Munsell 8, the lowest contrast decrement was practically indistinguishable from the flat glossy surround. Rendering the flat surrounds with gloss made the surfaces appear darker than their matte counterparts (because part of the light was reflected in the specular component that was not visible from the camera's point of view). This slight darkening of the surround caused the lowest contrast decrement to appear close enough to the surround albedo that it was almost undetectable. Observers were therefore likely to match this test patch to the surround albedo, reducing the lowest contrast increment-decrement step.

---

[5] Follow-up paired t-tests using Sidak-corrected alpha values of 0.0398 per test indicated that the matte surround induced a larger step than the glossy surround only for the low-relief condition, $t(4) = 3.62$, $p = 0.02$, not the high-relief condition, $t(4) = -1.74$, $p = 0.16$. Additionally, the low-relief surround induced a larger step than the high-relief surround only when the surround was matte, $t(4) = 5.86$, $p = 0.004$, not glossy, $t(4) = 0.84$, $p = 0.45$. Sidak-corrected alpha values were calculated as $\alpha = 1 - (1 - \text{FWER})^{1/n} = 1 - (1 - 0.15)^{1/4} = 0.0398$, where FWER is the family-wise error-rate for the ANOVA.

	Surround 3.5		Surround 5		Surround 6.5		Surround 8	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
Relief	19.84	0.011*	27.01	0.007*	84.25	0.001*	41.21	0.003*
Gloss	0.03	0.874	6.06	0.07	3.05	0.156	8.89	0.041*
Interaction	0.1	0.768	1.79	0.251	0.16	0.706	15.96	0.016*

Table 2.3. *F* values and *p* values for the increment–decrement difference scores of Experiment 1A. * $p < 0.05$. For all tests, $df = (1, 4)$.

For Experiment 1B, t-tests using Sidak-corrected alpha values of 0.253 per test[6] were carried out to compare difference scores in the flat and rocky surround conditions (see Table 2.4 for *t* values, *df*, and *p* values). For three out of the four surround albedo conditions, flat surrounds induced a larger increment–decrement step than both the rocky matte and glossy surrounds (Figure 2.7B). The apparent lack of step for surround Munsell 8 can be explained by the probabilistic nature of detecting very low-contrast test patches (Ekroll & Faul, 2012b). It appears that observers in Experiment 1B were, on average, less likely to detect the low-contrast decrement compared to observers in Experiment 1A.

Flat vs.	Surround 3.5			Surround 5			Surround 6.5			Surround 8		
	df	*t*	*p*	*df*	*t*	*p*	*df*	*t*	*p*	*df*	*t*	*p*
Rocky matte	19	3.67	.002*	19	4.82	<.001*	19	3.98	.001*	19	2.04	.055
Rocky glossy	19	2.43	.0249*	19	6.18	<.001*	19	2.77	.012*	19	1.55	.14

Table 2.4. *t* values, *p* values, and *df* comparing the increment-decrement difference scores between conditions in Experiment 1B. * $p < 0.05$.

The above findings raise the question of why there were differences in observers' ability to detect low-contrast test patches between Experiment 1A and 1B. Additionally, there was a discrepancy in the amount of crispening observed (the size of the step) between Experiment 1A and 1B. The results of previous research suggest that these

---

[6] Sidak-corrected alpha values were calculated as $\alpha_1 = 1 - (1 - \alpha)^{1/n} = 1 - (1 - 0.05)^{1/2} = 0.0253$.

inconsistencies can be attributed to individual differences in how the stimuli are perceived. For example, Ekroll and Faul (2009) found large individual differences in the size of the crispening effect for coloured centre-surround displays.

Figure 2.6 displays another quality about the data curves from the flat surround conditions, namely that there is an asymmetry between increment and decrement settings. The averaged data in Figure 2.6 (top two rows) reveals that increment settings were essentially independent of the surround albedo. In contradistinction, decrement settings were more affected by the surround albedo, illustrated by the greater spread in the data points for each test patch. Asymmetries between increment and decrement settings have been found previously in the brightness literature (e.g., Heinemann, 1955). Additionally, for test patches on coloured surrounds, colour induction from the surround has been found to be much stronger for decrements than increments (Bäuml, 2001; Helson, 1938; Helson & Michels, 1948).

The most notable result was that crispening was completely eliminated when surrounds contained shading and shadow information indicative of surface-relief (bottom two rows in Figure 2.6). This implies a qualitative difference in how the test patch was perceived when embedded in flat compared to rocky surrounds. In Chapter 4 we further investigate the crispening effect seen in the flat surround conditions. We address a growing view in the literature that lightness perception may contain more than one dimension (Ekroll & Faul, 2013; Logvinenko & Maloney, 2006; Vladusich, 2012, 2013; Vladusich et al., 2007) and investigate the role of mid-level perceptual phenomena, such as transparency, influencing the appearance of test patches embedded in homogeneous surrounds (Ekroll & Faul, 2013). This idea may also shed some light on the increment-decrement asymmetry mentioned above.

Results from the rocky (high-relief) surround conditions

The above results suggested a qualitative difference in how the test patch was perceived when embedded in rocky surrounds that contained shading information compared to flat surrounds that did not contain this information. For both matte and glossy rocky conditions, there was a tendency for test patches on darker surrounds to appear lighter than those same test patches on lighter surrounds. However, settings from the glossy condition seem more compressed (more consistent) than those from the matte condition (compare the bottom and second-bottom rows in Figure 2.6, respectively). For these rocky conditions, lightness constancy for a given test patch tended to be better when surrounds were glossy compared to matte.

The difference in lightness constancy (vertical spread) between rocky matte and glossy data points was statistically reliable. For each of the 11 test patch values common to all surrounds (see Table 2.2), the lightness settings in one surround condition were subtracted from the lightness settings in the adjacent darker surround albedo condition. These difference scores were averaged for all 11 test patch values and plotted in Figure 2.8. For this and subsequent experiments in this chapter, a binomial sign test was used to compute the likelihood of obtaining k or more instances in which the observers' performance in the glossy condition was more consistent than in the matte condition (11 pairs of data points per subject). The results confirmed that, for rocky conditions, lightness constancy was significantly better when test patches were surrounded by glossy compared to matte surfaces, $p < 0.001$ for Experiment 1A, $p = 0.006$ for Experiment 1B. This implies that surfaces with gloss information provided the visual system with additional cues that could be used to improve lightness constancy.

The amount of lightness constancy exhibited in the data also depended on the test-patch albedo. However, this effect of test-patch albedo on lightness constancy differed between observers in Experiment 1A and 1B. Despite these differences between observers, Figure 2.6 (third row from the top) and Figure 2.8 demonstrate a general trend, at least for the rocky matte condition: Lightness constancy was better for lighter compared to darker test patches. This is illustrated by the negatively sloped matte data curves in Figure 2.8. This is likely to be caused by lighter test patches substantially increasing the range of luminance values in most of the displays.

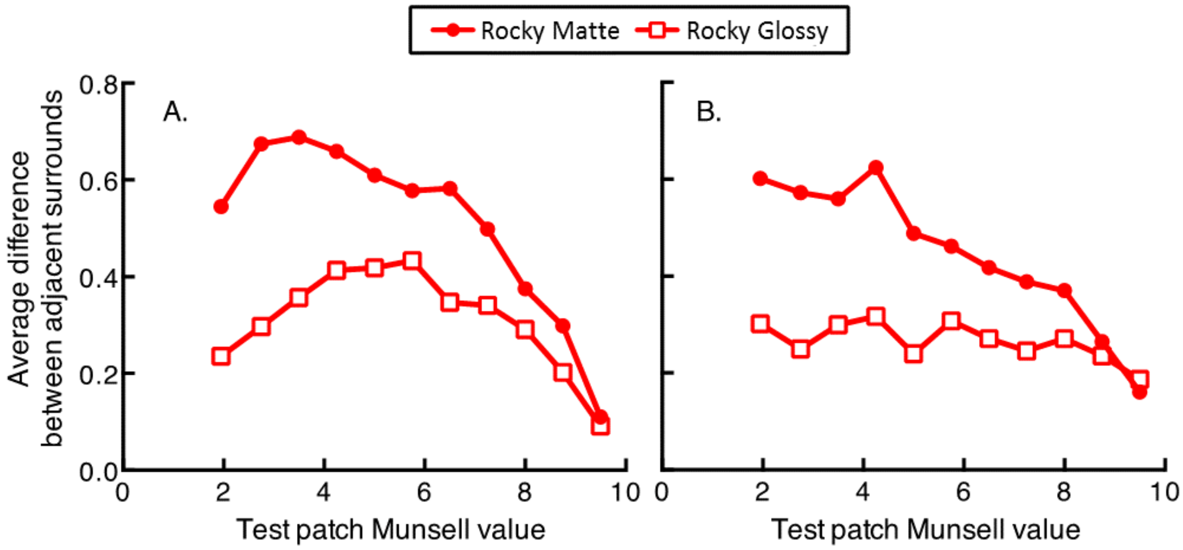


Figure 2.8. Average difference scores for the rocky conditions of Experiment 1A (A) and 1B (B). Average difference scores were calculated in the following way: For each of the 11 test patch values common to all surrounds (see Table 2.2), the lightness settings in one surround condition were subtracted from the lightness settings in the adjacent darker surround albedo condition. The plotted values are the average of these difference scores for each of the 11 test patch values. Lightness constancy was better for test patches embedded in glossy (open squares) compared to matte (closed circles) surrounds.

One last point to note is observers' settings in relation to ground truth (solid black diagonal line in Figure 2.6). For the rocky matte data points, there was a tendency for test patches on lighter surrounds to be more veridical to those on darker surrounds. One reason for this shift might be the high range of luminance values in the adjustable patch's surround. Matte surrounds with a lighter albedo also had a greater range of luminance values (due to light diffuse shading and dark shadows) compared to matte surrounds with lower albedo (which had dark diffuse shading and dark shadows). This made the matches to our adjustable patch (which was surrounded by both black and white surfaces) more symmetric when displays are lighter, leading to more veridical matches.

In Experiment 1A and 1B we found evidence suggesting that the visual system used image cues generated by rocky and glossy surfaces when estimating test patch lightness. The presence of these cues led to better lightness constancy compared to when these cues were absent: Crispening was eliminated when surrounds contained complex mesostructure; for rocky surfaces, lightness judgments were more consistent when surrounds were glossy compared to matte. Note that the rocky matte data points exhibited similar overall spread to the flat surround displays (i.e. a given test patch embedded in the lightest and darkest surround appeared similarly different for the two conditions). In this sense lightness constancy appeared to be similar for these two conditions. However, when focusing on data points near the surround value, lightness constancy was revealed to be better in the rocky displays.

One possible explanation of the above results is that the specific luminance patterns generated by naturalistic surfaces may serve as cues to help the visual system decompose luminance values into contributions of lightness and illumination. Advocates of EIMs have proposed that these cues provide the visual system with information about the light field, which is then transferred to other parts of the scene that lack this information, e.g. the flat test patch. However, there are a number of other possible explanations, which will be systematically addressed in the remaining experiments of Chapter 1. One alternative explanation is that the improved lightness constancy for glossy conditions could have been caused by the greater range and/or variation of luminance values in the image compared to matte conditions (this has been referred to as *articulation* in the literature; e.g. see Gilchrist et al., 1999). Experiment 2 was designed to directly test this possibility. Variegated centre-surround stimuli were created in a way that eliminated surface structure but retained the range and variation of luminance values in the image. Another possibility relates to the similarity between the test and matching displays, where more similar test and matching displays might lead to more symmetric matches. This will be investigated in Experiment 3. Finally, Experiment 4 tests the possibility that the specific contrast and luminance distributions in the rocky surrounds are responsible for the lightness effects in Experiment 1.

Experiment 2A and 2B: Variegated Surrounds

Experiment 2 tested whether surround "articulation" (luminance range and/or variation) was responsible for the improved lightness constancy exhibited with the rocky surrounds in Experiment 1. In Experiment 2, equivalent two-dimensional flat surfaces with variegated surrounds were created for each rocky surface (Figure 2.9). These variegated surfaces had very similar luminance histograms to the rocky surfaces but lacked information about surface structure, shading, specularity, or illuminance flow. If the specific luminance patterns generated by naturalistic surfaces are important in lightness constancy, then lightness constancy should be worse for the variegated surfaces than for comparable rocky surfaces. If surface structure is not important, and it is only the amount of articulation that is crucial, then no difference in lightness constancy between rocky and variegated conditions should be observed.

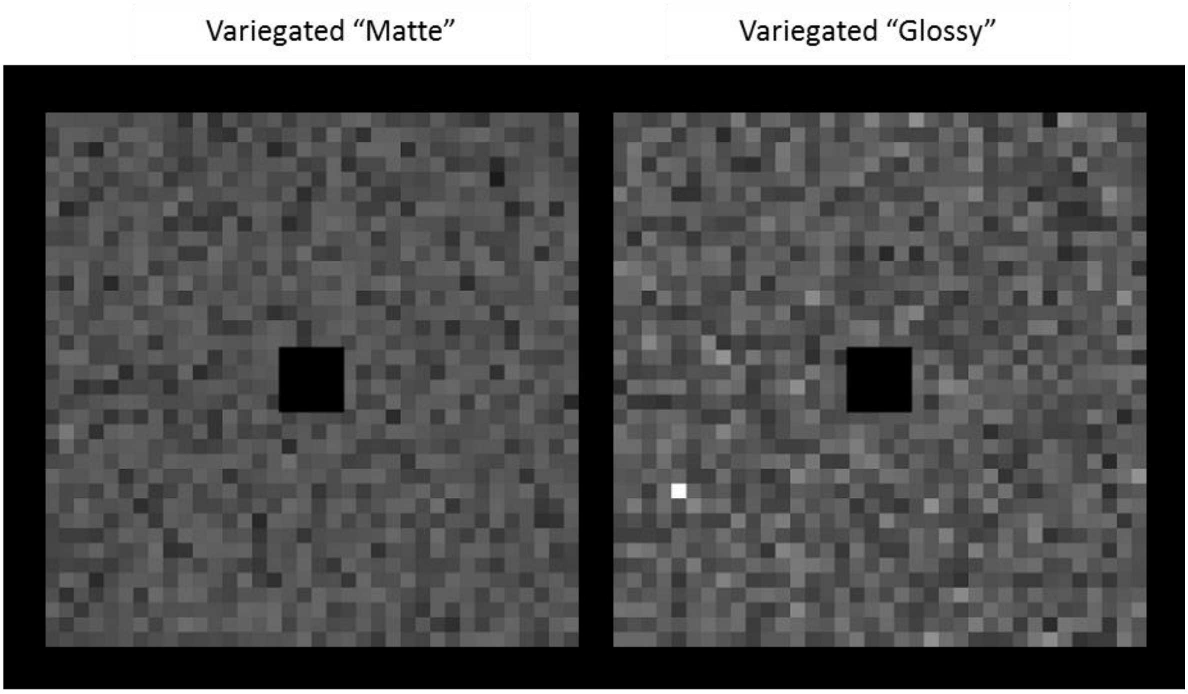


Figure 2.9. Examples of variegated centre-surround stimuli used in Experiment 2. The surrounds shown here were created from the rocky matte and glossy displays from Experiment 1 (19.8% reflectance).

Methods

Observers

Similar to Experiment 1, Experiment 2 had two parts: Observers DC and RS from Experiment 1A participated in Experiment 2A, and 20 first-year psychology students participated in Experiment 2B (none of whom had participated in the first experiment).

Apparatus and stimuli

The task was the same as in Experiment 1 as was stimulus presentation. Two-dimensional variegated surrounds were created in the following way for each gloss level and each surround albedo condition (see Figure 2.9). First, the relative frequencies of each luminance value in the rocky surround were obtained. There were 256 possible luminance values corresponding to each eight-bit pixel value of the monitor. These relative values were used to create a variegated surround consisting of 1,280 squares (36 × 36 minus 16 squares for the central patch). The number of squares of each luminance was weighted according to the relative luminance frequencies from the rocky surround. This resulted in very similar luminance frequency histograms for corresponding variegated and rocky surrounds. In this sense, the corresponding variegated and rocky surrounds are equivalent, but variations in shading and gloss in the rendered rocky stimuli appear as part of a random noise pattern in the variegated images. Because the rocky surround was made up of 799,515 pixels and the variegated surround contained only 1,280 squares, some luminance values in the rocky surround did not have enough pixels to make up one square in the variegated surround. If this occurred, we ensured that the variegated surround contained at least one square with the highest luminance value and at least one square with the lowest luminance value from the equivalent rocky surround. We also ensured the regions immediately surrounding the central patch in the variegated and rocky stimuli were matched in average luminance (the distance of one square away from the central patch).

There were very slight luminance variations within the test patches embedded in rocky surrounds, caused by interreflections. Test patch values for the variegated stimuli were created by averaging these luminance values.

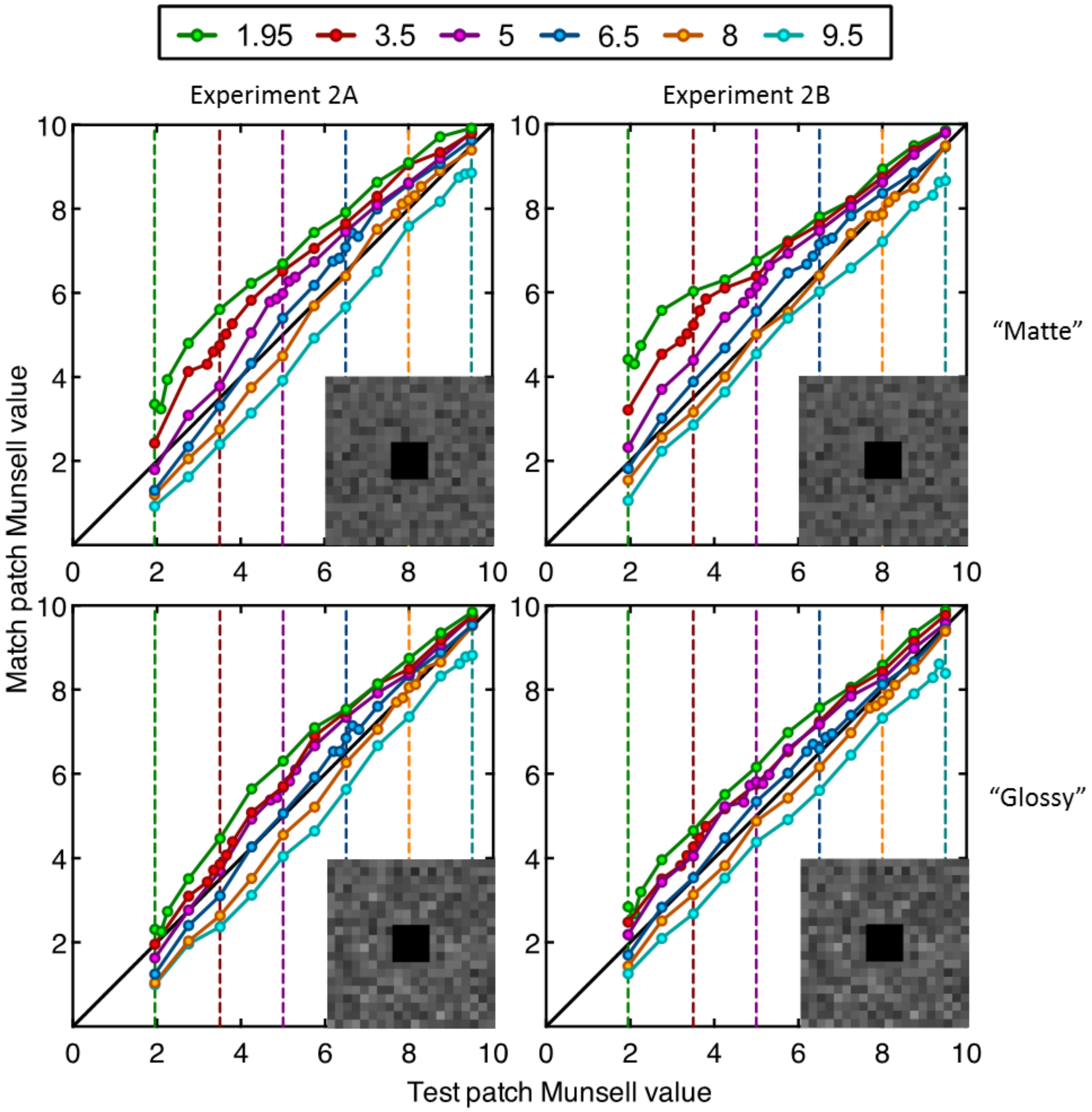


Figure 2.10. Results of Experiment 2. Left panels: average data for observers DC and RS from Experiment 2A. Right panels: average data for Experiment 2B. See Figure 2.6 caption for details about the data curves and legend. Top panels: lightness settings for the variegated matte equivalent condition. Bottom panels: settings for the variegated glossy equivalent condition.

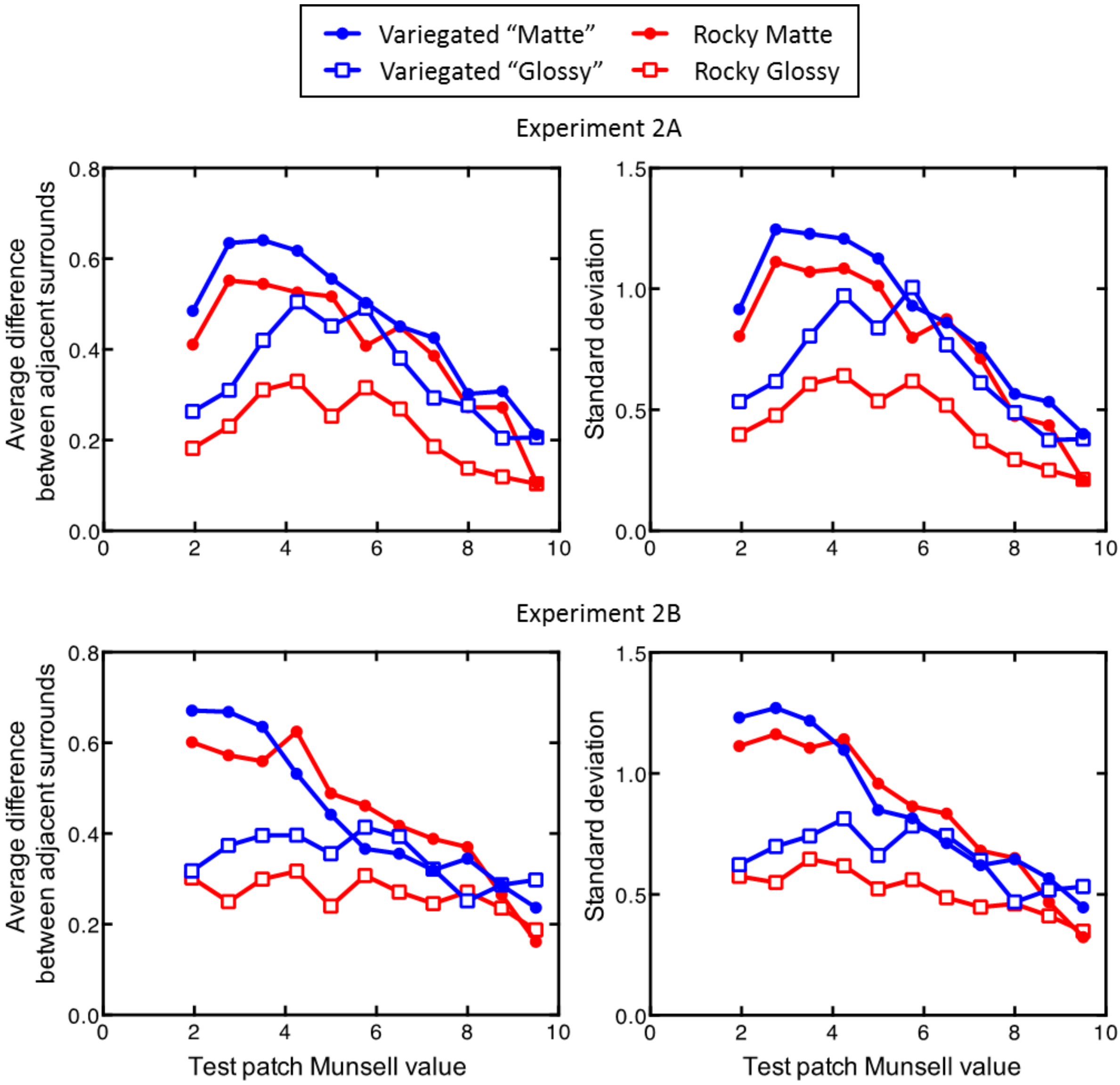


Figure 2.11. Difference scores from Experiment 2. Left panels: average difference scores for Experiment 2A (top) and 2B (bottom). See Figure 2.8 caption and main body text for an explanation of how these scores were calculated. Right panels: standard deviation scores for Experiment 2A (top) and 2B (bottom), calculated as the standard deviation of test patch settings for different surround albedo conditions. See main body text for a description of effects.

Procedure

The presentation of trials and instructions were the same as in Experiment 1. There were two surround types: variegated surrounds that were equivalent to the matte surrounds in Experiment 1 ("matte equivalent") and variegated surrounds that were equivalent to the glossy surrounds in Experiment 1 ("glossy equivalent"). The surround albedo conditions and test patch conditions were the same as in Experiment 1, leading to 172 trials in total.

Results and discussion

The average results of Experiment 2 are presented in Figure 2.10. Two trials from Experiment 2B were excluded from analyses due to observers making premature lightness judgments (see Experiment 1 results). The data reveal that test patch lightness settings were affected by surround type. To more clearly compare lightness settings, Figure 2.11 (left panels) plots the average difference scores between adjacent albedo conditions (the same method was used as in Experiment 1). A smaller difference score indicates more consistent settings and therefore better lightness constancy. We also plotted the standard deviation of test patch settings for different surround albedo conditions (Figure 2.11, right panels) to demonstrate that the effects reported are not a result of the specific method chosen to quantify lightness constancy.

Binomial sign tests were carried out on the average difference scores and standard deviation scores.[7] The results of Experiment 2A showed that when surrounds were variegated, observers exhibited better lightness constancy when there was greater articulation (luminance variation/range) in the surround (glossy equivalent condition) compared to when there was less articulation (matte equivalent condition), $p < 0.001$ (Figure 2.11, top panels; compare blue filled circles to blue open squares). However, this result was not replicated in Experiment 2B, $p = 0.5$, indicating that greater surround articulation may not influence lightness constancy for all observers (Figure 2.11, bottom

---

[7] For simplicity, only p values for tests on the average difference scores are reported. However, the results are identical for standard deviation scores.

panels). The results from both sets of observers confirmed that lightness constancy was better when test patches were embedded in rocky compared to variegated surrounds, at least for the gloss conditions, $p < 0.001$ (Experiment 2A, Figure 2.11, top panels), $p = 0.006$ (Experiment 2B, Figure 2.11, bottom panels; compare open blue squares to open red squares). This suggests that rocky surfaces containing shading and specular reflections provided additional cues over and above luminance range and variation that the visual system used to improve lightness constancy. The results were not as clear for the matte condition. Observers from Experiment 2A exhibited better lightness constancy when test patches were embedded in rocky compared to variegated surrounds, $p < 0.001$ (Figure 2.11, top panel). However, the same was not true for observers in Experiment 2B, $p = 0.50$ (Figure 2.11, bottom panel; compare filled blue circles to filled red circles).

The results of Experiment 2 suggest that the amount of surround articulation (luminance variation/range) in the rocky conditions cannot account for the results in Experiment 1. The specific luminance patterns generated by shading and shadow contrast may be important for lightness constancy by providing the visual system with some information about surface albedo (i.e., lighter surfaces reflect more light to "fill in" shadows). However, the real advantage of surfaces over variegated surrounds appears to be derived from the specular highlights. This fits with an illumination-estimation account of the data. Light reflected from the diffuse component of a surface depends on both the (direct and indirect) light sources and the surface albedo, resulting in variations in shadow and shading information between surfaces. Alternatively, the light reflected from the specular component depends primarily on the illumination intensity, causing specular highlights to remain constant in their appearance between surfaces. The visual system may be able to use this constant cue across conditions to estimate the direction and intensity of the illumination from specular highlights, better constraining lightness estimates.

Although an illumination-estimation account fits nicely with the data from Experiment 2, there are two other possible explanations for the difference in lightness constancy found between the rocky glossy and variegated glossy equivalent conditions. In Experiments 1 and 2, the adjustable patch was surrounded by a rocky glossy surface with black and white checks (Figure 2.1E). We chose the specific pattern and material of the surround to make the test patch appear as surface-like as possible. Surface properties, such as lightness, have little meaning for flat two-dimensional stimuli, and observers often cannot easily differentiate between lightness and brightness (Blakeslee et al., 2008). Although the pattern and material of the surround made the adjustable patch appear surface-like, this resulted in observers' matches being more symmetric in the rocky glossy trials compared to other trials, especially in the flat surround condition in Experiment 1 and the variegated condition in Experiment 2. This suggests that the better constancy exhibited in the glossy trials may have been due to the similarity in the match and test surrounds. Experiment 3 addresses this possibility by replicating Experiments 1 and 2 but replacing the surround of the rocky glossy adjustable patch with a two-dimensional matte surround.

Another alternative explanation of the pattern of data from Experiments 1 and 2 appeals to low-level mechanisms. Although the variegated and rocky surrounds were matched in terms of their luminance histograms, they differed in their contrast and luminance distributions across space and scale. For example, the brightest luminance values in the rocky surround (the specular highlights) were distributed across the entire surround (see Figure 2.1D). However, Figure 2.9 illustrates that the brightest luminance values from the glossy equivalent variegated surround are contained in a single square. Furthermore, the specular highlights in the rocky surround occurred near points at which the diffuse shading gradient was brightest (Marlow et al., 2011). However, for the variegated surround, the contrast of adjacent squares was random. Experiment 4 directly tests the possibility that the specific contrast and luminance distributions in the rocky surrounds played a key role in affecting lightness perception in Experiments 1 and 2.

Experiment 3: Flat, Matte Adjustable Surface

Experiment 3 addresses whether the specific surround of the adjustable patch caused the pattern of results in Experiments 1 and 2 due to more symmetrical matching between the rocky glossy test and adjustable displays.

Methods

Observers

Forty first-year psychology students participated in Experiment 3. All observers were naïve and had not participated in any of the previous experiments. Twenty observers were assigned to the homogeneous (flat matte) and the matte and glossy rocky conditions of Experiment 1B. The other 20 observers were assigned to the matte equivalent and glossy equivalent variegated conditions of Experiment 2.

Apparatus, stimuli, and procedure

The task and conditions were identical to Experiments 1B and 2 except that the adjustable patch was rendered with a two-dimensional, matte checkerboard surround (Figure 2.1F). From now on the displays with flat, matte surrounds will be referred to as homogeneous displays, as there is no longer a need to differentiate this from the flat glossy condition in Experiment 1.

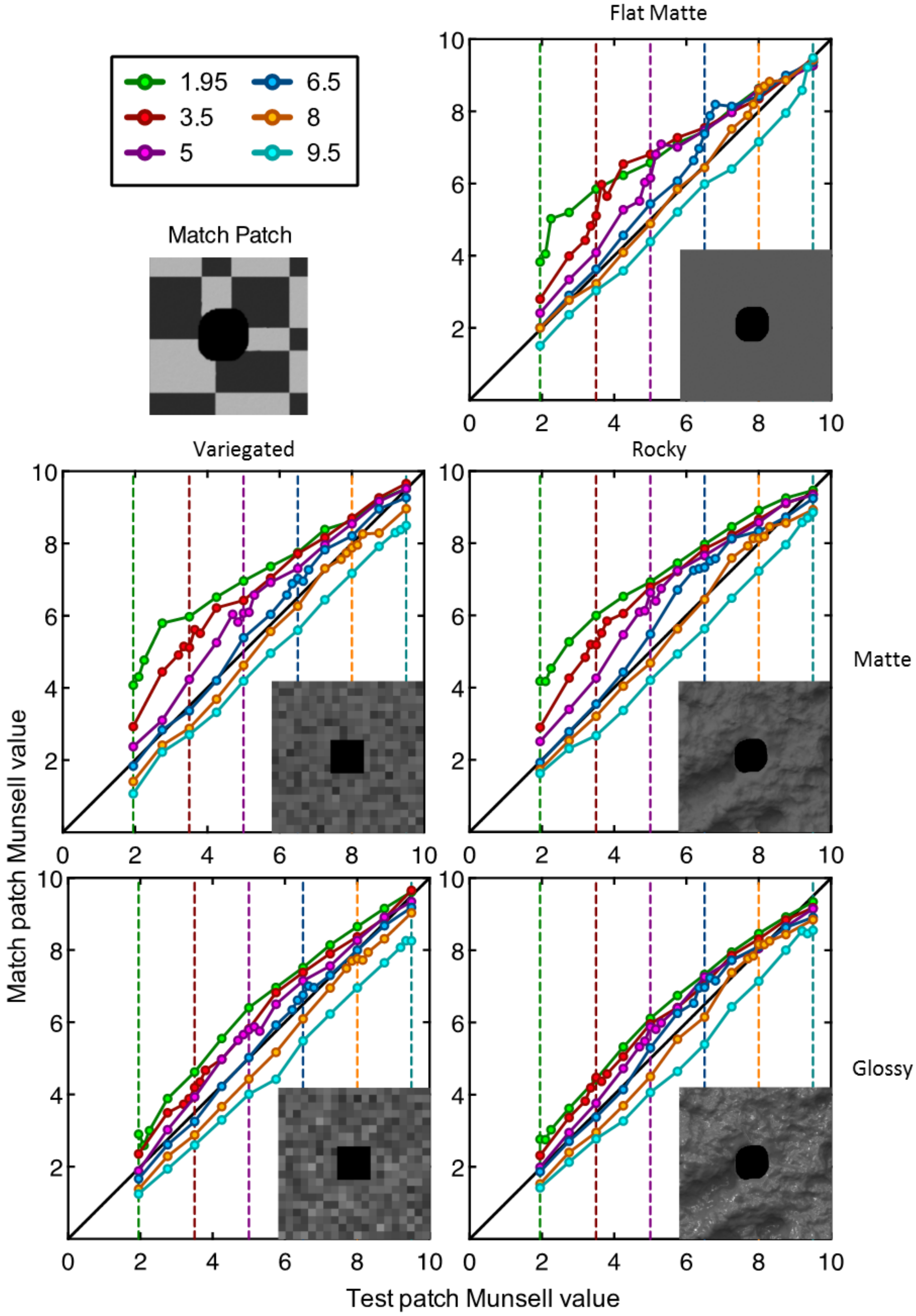

1.95
3.5
5
6.5
8
9.5
Match Patch
Flat Matte
Variegated
Rocky
Matte
Glossy
Match patch Munsell value
Test patch Munsell value
0
2
4
6
8
10

Figure 2.12. Average data for Experiment 3. See Figure 2.6 caption for details about the data curves and legend. Top left: adjustable patch used for all conditions in Experiment 3. Top right: lightness settings for the homogeneous (flat) surround condition. Middle left: lightness settings for the variegated matte equivalent condition. Bottom left: settings for the variegated glossy equivalent condition. Middle right: settings for the rocky matte condition. Bottom right: settings for the rocky glossy condition.

Results and discussion

The results of Experiment 3 are presented in Figure 2.12. Two trials were excluded from analyses due to observers making premature lightness judgments (see Experiment 1 results). The data reveal that test patch lightness settings were affected by each surround in the same way as in Experiments 1 and 2. Figure 2.12 illustrates that crispening occurred in the homogeneous (flat matte) surround conditions (top right panel) but not in the rocky conditions (middle right and bottom right panels). As in Experiment 1, increment–decrement difference scores were obtained to emphasize the size of the step for each of the homogeneous and rocky conditions (see Figure 2.13). Recall that a lower score indicates a smaller step and therefore less crispening.

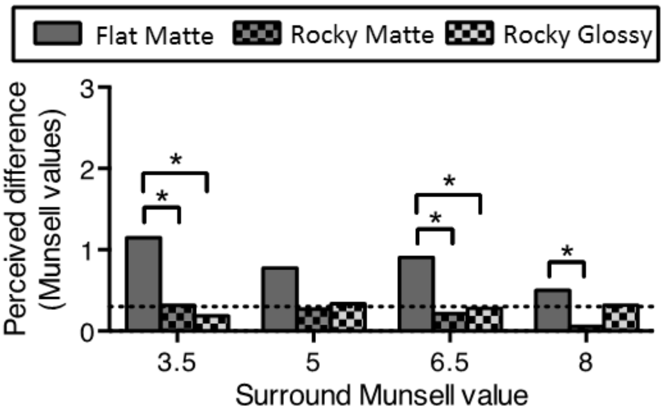


Figure 2.13. Increment minus decrement settings for Experiment 3. The horizontal dotted line represents the actual difference between increments and decrements. Replicating Experiment 1, large increment-decrement steps are present in the data from the homogeneous displays (solid bars). This step is eliminated when surrounds contain complex mesostructure (rocky conditions, checked bars).

T-tests using Sidak-corrected alpha values of .0253 per test were carried out to compare difference scores in the rocky and flat surround conditions (see Table 2.5 for *t* values, *df*, and *p* values). The results replicated those of Experiment 1B except that the comparisons for surround Munsell value 5 were not significant (see Table 2.5). Nonlinearities diagnostic of crispening were present for the homogeneous conditions (Figure 2.12, top right panel). Therefore, as in previous experiments, this slight discrepancy in results can be attributed to individual differences in observers' ability to detect very low contrast test patches.

Flat vs.	Surround 3.5			Surround 5			Surround 6.5			Surround 8		
	df	*t*	*p*	*df*	*t*	*p*	*df*	*t*	*p*	*df*	*t*	*p*
Rocky matte	19	4.52	<.001*	19	1.79	.09	19	2.47	.02*	19	2.68	.02*
Rocky glossy	19	4.57	<.001*	19	1.97	.06	19	3.39	.003*	19	0.80	.44

Table 2.5. *t* values, *p* values, and *df* for the increment-decrement difference scores of Experiment 3. * $p < 0.05$,

The results from the rocky and variegated conditions were also replicated when the flat adjustable patch was used, which can be seen by the average difference and standard deviation scores in Figure 2.14 (calculated the same way as in Experiments 1 and 2). Binomial sign tests on these scores revealed that for the rocky condition, lightness constancy is better when test patches are embedded in glossy compared to matte surrounds, $p = 0.006$. However, as in Experiment 2B, this is not the case for the variegated conditions. Importantly, for the glossy condition, lightness constancy was better when test patches were embedded in rocky compared to variegated surrounds, $p < 0.001$. This was not the case for the matte condition, $p = 0.5$. These results suggest that the effects in Experiments 1 and 2 were not caused by the symmetry between the test and adjustable surfaces.

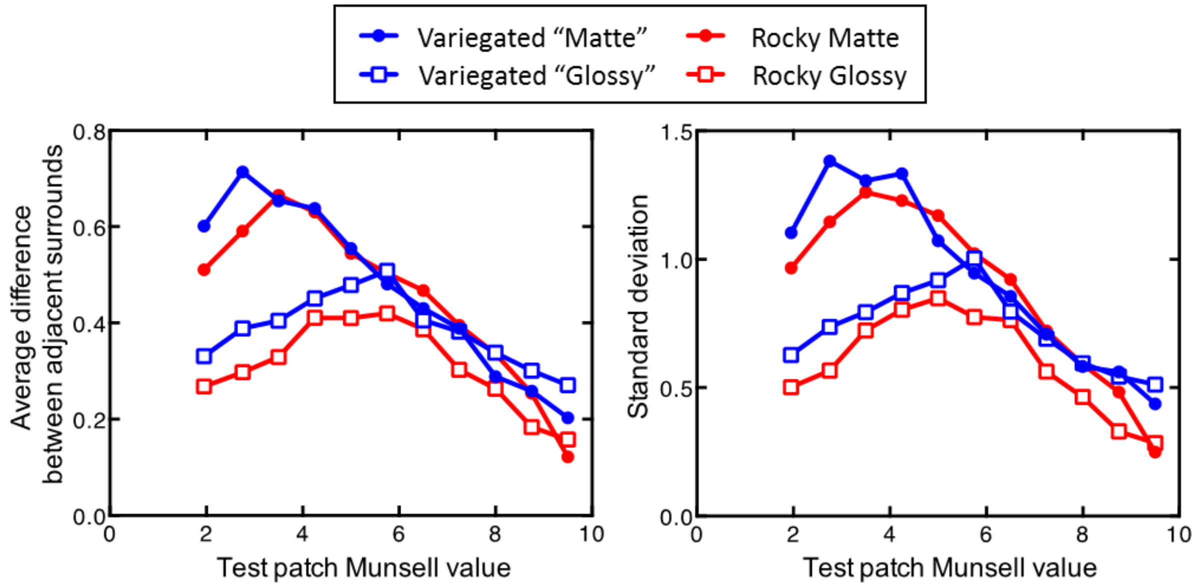


Figure 2.14. Average difference scores (left panel) and standard deviation scores (right panel) for Experiment 3. See Figure 2.8 caption and main body text for an explanation of these scores. See main body text for a description of effects.

Experiment 4: Phase-Scrambled Surrounds

Experiments 1 and 2 demonstrated that the luminance patterns generated by specular highlights improved lightness constancy for flat test patches embedded in rocky surrounds. Experiment 3 demonstrated that this improved lightness constancy was not simply the consequence of the test and match patches having similar surrounds. Experiment 4 aims to determine whether this improvement in lightness constancy was achieved through an improvement in the visual system's ability to represent the light field or whether a low-level explanation involving contrast and luminance distributions across space and scale can account for the data. In Experiment 4, the equivalent control surrounds were phase-scrambled versions of the rocky surrounds (see Figure 2.15). Scrambling the phase spectrum information ensured that the rocky and control surrounds were matched not only in terms of their luminance histograms, but also in terms of their spatial frequency content. The phase-scrambled surrounds tend to evoke impressions of grainy textures, which appear to be perceived predominantly as variations in pigment (particularly the specular highlights). If a representation of the light field via shading patterns and/or specular highlights is crucial for lightness constancy, then observers should exhibit better lightness constancy when test patches are embedded in rocky surrounds. Alternatively, if the enhanced lightness constancy of the glossy surfaces arose from the distribution of contrasts across spatial scales, no difference in lightness constancy between the two conditions should be observed.

Methods

Observers

Twenty first-year psychology students participated in Experiment 4. All observers were naïve and had not participated in any of the previous experiments.

Apparatus and stimuli

The task was the same as in Experiments 1 and 2. Phase-scrambled surrounds were created by Fourier transforming the rocky images and replacing the phase spectrum information with that of white noise. Thus, phase-scrambled and rocky surrounds had equated pixel luminance histograms and power spectrums. As in Experiment 2, the regions immediately surrounding the central patch in the rocky and phase-scrambled stimuli were matched in average luminance.

Procedure

The presentation of trials and instructions were the same as in Experiment 1 and 2. There were four surround-type conditions: rocky matte, rocky glossy, phase-scrambled matte equivalent, and phase-scrambled glossy equivalent. The surround albedo conditions were the same as in the previous experiments. However, observers only performed lightness judgments on the 11 test patches common to all the surround types in Experiment 1 (see Table 2.2), i.e., no very low-contrast test patches were used. This resulted in 264 trials for each observer.

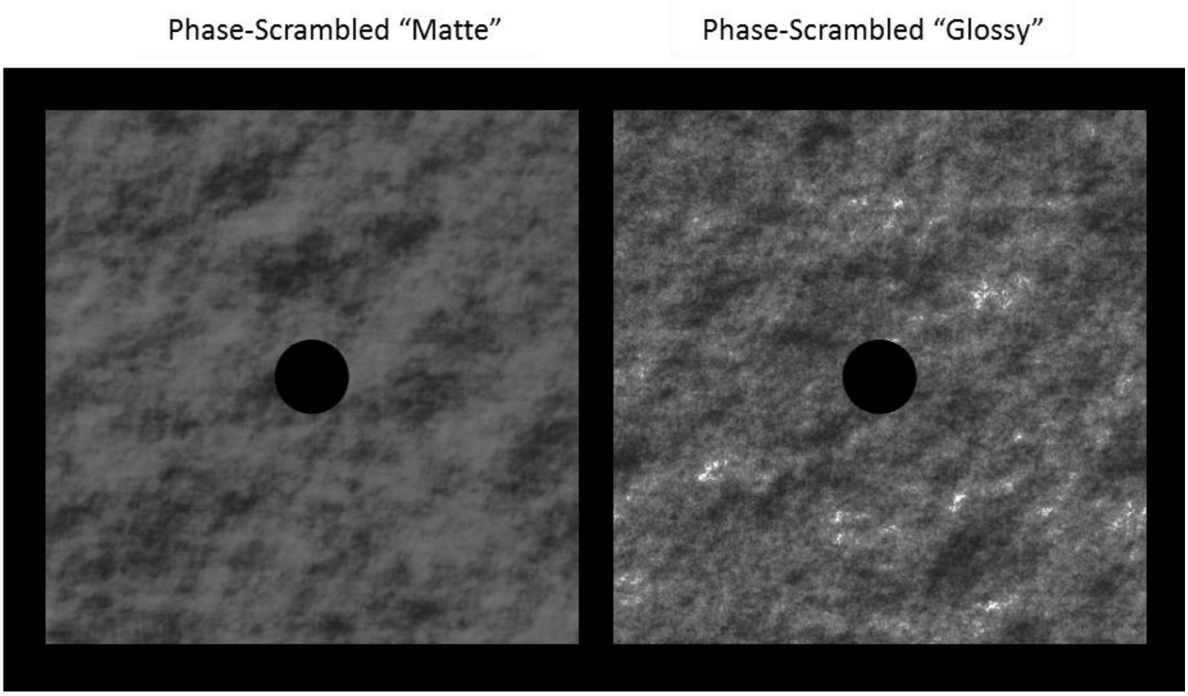


Figure 2.15. Examples of phase-scrambled centre-surround stimuli used in Experiment 4. The surrounds shown here were created from the rocky matte and glossy displays from Experiment 1 (19.8% reflectance).

Results and discussion

The results of Experiment 4 are presented in Figure 2.16. The data reveal that test patch lightness settings differed between matte and glossy conditions, replicating the results of Experiment 1. However, lightness constancy did not differ between rocky and phase-scrambled surround conditions. Figure 2.17 plots the average difference scores and the standard deviation between adjacent surround albedo conditions, calculated the same way as in Experiments 1 and 2. A lower score indicates better lightness constancy.

Binomial sign tests were carried out on the average difference scores and standard deviation scores. The results showed that regardless of whether surrounds were rocky or phase-scrambled, observers exhibited better lightness constancy in the glossy compared to the matte condition, $p = 0.006$ (rocky condition), $p < 0.001$ (phase-scrambled condition; Figure 2.17). The results also showed no difference in lightness constancy between the rocky and phase-scrambled conditions, $p = 0.11$ (matte condition), $p = 0.27$ (glossy condition). This suggests that, for these types of stimuli at least, the visual system was not using information about the light field generated by specular reflections to improve lightness constancy for glossy surfaces; observers performed just as well when control surfaces were created with the same contrast and luminance distributions as naturalistic rocky surfaces.

The results of this experiment also replicate the trends observed in the matte condition of Experiment 1: For both rocky and phase-scrambled conditions, there was a tendency for lighter test patches to be matched more consistently than darker test patches (indicated by the negative slopes of the matte data curves [closed circles] in Figure 2.17). Furthermore, as in Experiment 1, data curves in the matte condition are shifted vertically up in relation to ground truth (Figure 2.16, top row, compare data curves to the diagonal solid black line).

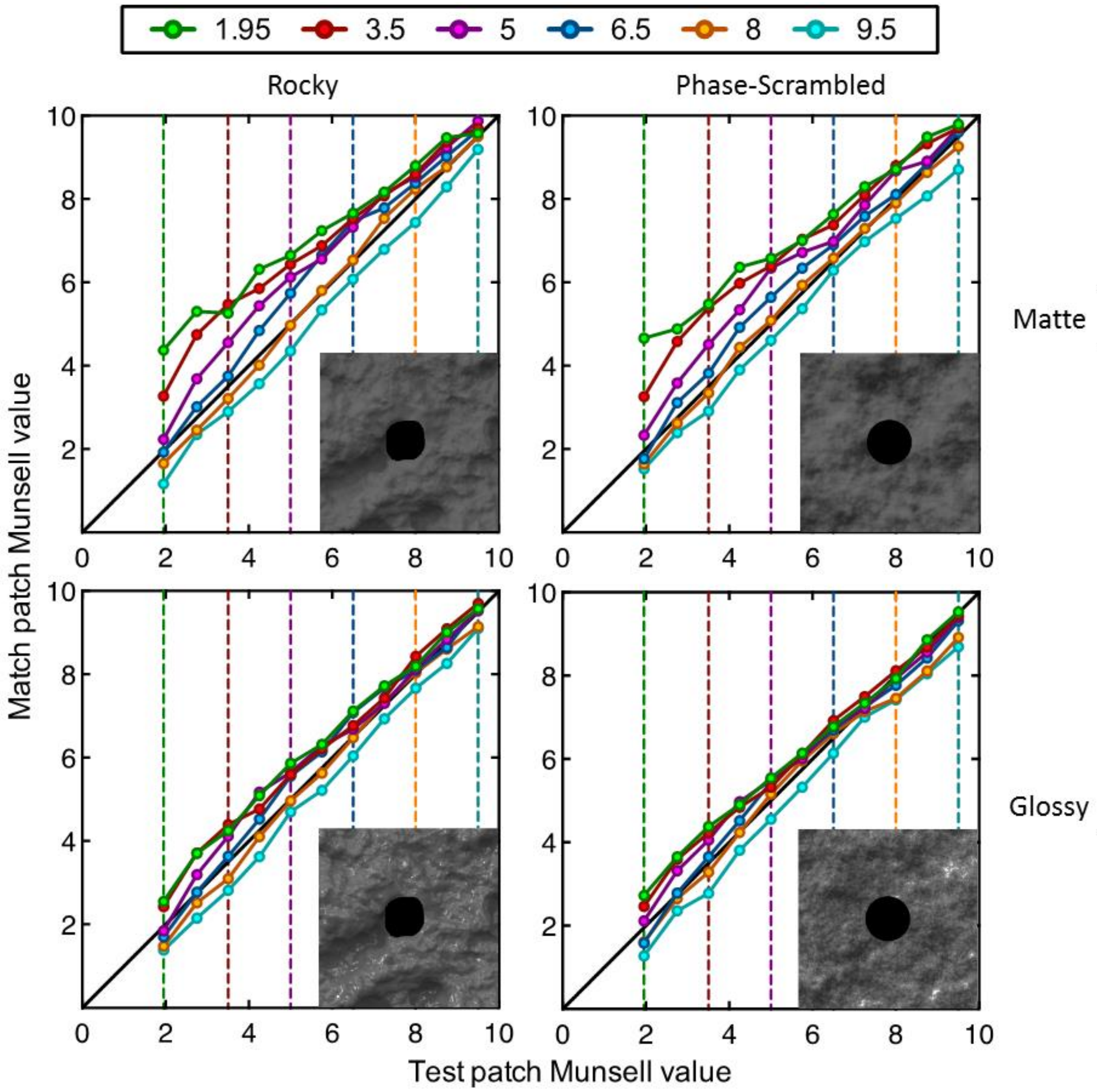


Figure 2.16. Average data for Experiment 4. See Figure 2.6 caption for details about the data curves and legend. Top left: settings for the rocky matte condition. Bottom left: settings for the rocky glossy condition. Top right: lightness settings for the phase-scrambled matte equivalent condition. Bottom right: settings for the phase-scrambled glossy equivalent condition. See Figure 2.17 caption and main body text for a description of effects.

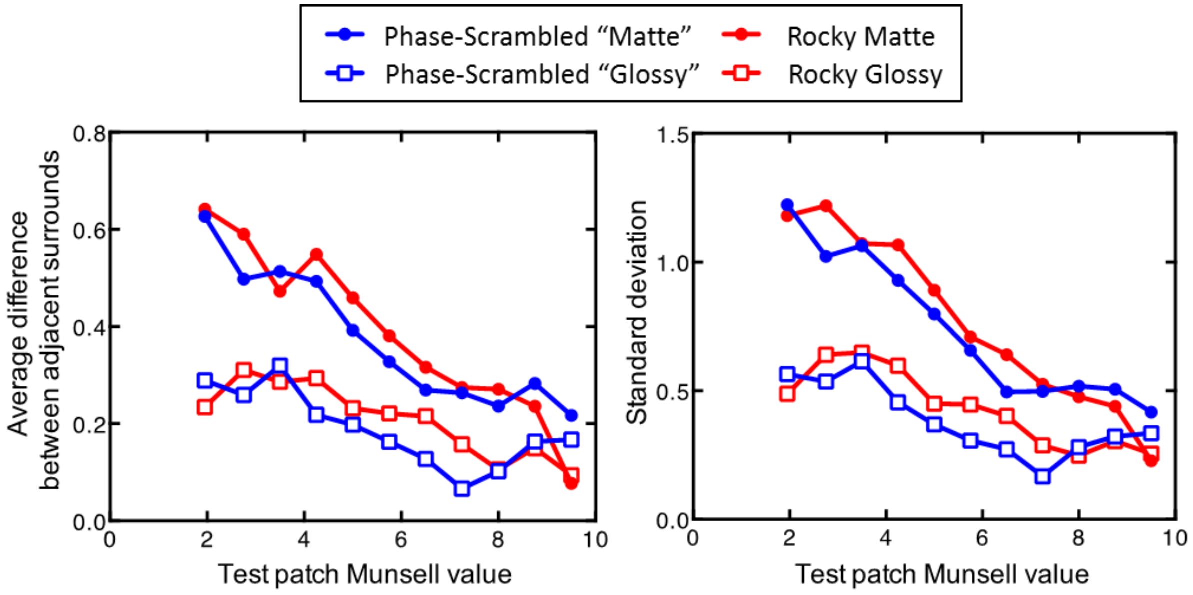


Figure 2.17. Average difference scores (left panel) and standard deviation scores (right panel) for Experiment 4. See Figure 2.8 caption and main body text for an explanation of these scores. Lightness constancy was better in the glossy (open squares) compared to matte (closed circles) conditions. However, there was no difference in lightness constancy between the rocky (red data points) and phase-scrambled (blue data points) conditions.

Concluding Thoughts

The first aim of the present chapter was to investigate whether the visual system uses cues to the illuminant created by complex mesostructure and specular highlights to improve lightness constancy of an embedded flat, matte test patch. The results support the hypothesis that more image cues lead to better lightness constancy. Low contrast test patches surrounded by rocky surfaces with complex mesostructure were judged more consistently than those with homogeneous surrounds, which exhibited a nonlinear "step" (crispening) at albedo values close to the surround. However, due to the qualitative difference in the data curves, it is difficult to directly compare lightness constancy between the rocky and homogeneous displays. The results from the matte and glossy rocky displays are more comparable. The results showed that test patches were judged more consistently when the surround contained specular highlights (glossy surfaces) compared to when these were absent (matte surfaces).

The second aim of this chapter was to tease apart mid-level and low-level contributions to the above lightness effects. From the data presented so far it is unclear what mechanisms drove the crispening effect in the homogeneous displays. This is investigated in Chapter 4. The control experiments tested whether the benefits of shading and specular reflections in the rocky displays could be attributed to the range and/or number of luminance values in the surround, the choice in matching patch surround, or differences in the energy across different spatial scales. Contrary to the hypothesis that originally motivated these experiments, the results do not provide support for explanations of lightness perception based on illumination estimation. Our results suggest that, when computing test patch lightness, there was no benefit in observers' lightness judgments for stimuli that contained surface relief, shadows, shading, and specular highlights. Rather, a low-level explanation involving contrast and luminance distributions across space and scale appears to be sufficient in explaining the pattern of results described above. However, at least two other factors need to be considered when interpreting the similar results between the rocky displays and the phase-scrambled controls, which are discussed below and investigated further in Chapter 3.

It is possible that in the present experiments observers matched the test patches on perceived luminance (brightness) rather than lightness. All stimuli were embedded in the same light field, and while there were slight differences in luminance for a given test patch due to secondary reflections from the backgrounds, these effects were negligible. Consequently, a test patch of a given albedo had the same luminance regardless of background (i.e. test patch albedo co-varied with luminance), meaning that observers could have performed the task by matching perceived brightness, not lightness. Previous research has shown that in some cases lightness and brightness matches do not differ when illumination is homogeneous (e.g. Blakeslee & McCourt, 2012), and despite variations in illumination due to vignetting and interreflections, globally our displays were illuminated uniformly (i.e., there were no illumination boundaries or gradients). Therefore, the similarity in results between rocky and phase-scrambled displays might reflect low-level mechanisms responsible for brightness computations within a display under homogeneous illumination. Such mechanisms may be inadequate explain lightness constancy under changes in illumination.

Additionally, it is possible that in the present experiments rendering the surfaces with two interreflections was not sufficient to perceptually differentiate rocky surfaces with different albedo. We reasoned that two interreflections was sufficient because it is more than is usually used in lightness experiments with computer-simulated 3D scenes, which typically have at most 1 interreflection. However, two interreflections may not have generated perceptually detectable differences in the contrast of shading and shadows between surfaces of different albedo. Therefore the extent to which shadows are “filled in” may not have been available as a cue to differentiate the rocky surfaces. This is suggested by the results of Experiments 1 and 3, where aspects of the rocky matte data were similar to the data from homogeneous displays. Specifically, for test patches that were decrements (i.e. darker than the surround) lightness settings appeared to be just as variable for rocky matte displays as homogeneous displays (e.g. compare the top to the 2nd bottom data panels of Figure 2.6). Note that this does not include test patches low in contrast and thus affected by crispening in the homogeneous surrounds. This comparison suggests that, while adding mesostructure eliminated crispening, the number of interreflections rendered may not have generated sufficient perceptual differences in shading between surrounds of different albedo. If observers were unable to use image cues generated by the mesostructure interacting with light to differentiate the surrounds, then lightness constancy would not be expected to differ between rocky surfaces and phase-scrambled controls. To address the above two issues, Chapter 3 investigates lightness constancy under changing illumination, and how this is affected by differences in the number of interreflections rendered in the displays.

Chapter 3. The Role of Interreflections on Lightness Perception under Changing Illumination Level

Chapter 3 explores whether low-level mechanisms can sufficiently explain lightness perception of flat test patches embedded in rocky surrounds when illumination level is varied. When the level of illumination is varied (i.e. from bright to dim), the luminance of a test patch also varies. Therefore it is possible to determine whether observers are making reflectance matches (their lightness judgments remain constant despite changing illumination) or luminance matches (their lightness judgments vary with illumination level). Another aim of Chapter 3 is to determine whether the number of interreflections rendered in the stimuli used in Chapter 2 provided the visual system with sufficient information to differentiate the lightness of surfaces based on the contrast of shadows in an image.

Experiment 5A and 5B: Lightness Perception of Complex Surfaces under Changing Illumination Level and Number of Interreflections

Experiment 5 had two parts, Experiment 5A and 5B. In Experiment 5A observers performed lightness judgments on central test patches embedded in various surrounds. This was similar to the experiments in Chapter 2. Surrounds were matte or glossy surfaces with complex mesostructure rendered under either “bright” (high) or “dim” (low) illumination. Phase-scrambled control displays were created for each condition (see methods) and compared to the rocky displays. We predicted that lightness judgements under changing illumination would be more consistent for test patches surrounded by rocky surfaces compared to phase-scrambled controls if mid-level mechanisms involved in the processing of surface structure affect perceived lightness in the rocky displays. Alternatively, lightness constancy should not differ between rocky and control surfaces if perceived lightness in the rocky displays can be explained by low-level distributions of luminance and contrast across different spatial scales.

The number of interreflections rendered in the displays was also manipulated to determine whether perceived lightness varies as a function of the number of interreflections. With this design it was possible determine the minimum number of interreflections (if any) required for perceived lightness to differ between the rocky and phase-scrambled conditions. We hypothesised that lightness judgments would differ between the rocky and scrambled stimuli after a certain number of interreflections if surface structure is important for lightness perception (as long as surfaces are rendered with enough interreflections). This would suggest that the filling in (or reduced contrast) of shading can be used to differentiate the lightness of surrounds with different albedos. However, lightness judgments should be similar between rocky and scrambled stimuli regardless of the number of interreflections if low-level mechanisms are sufficient to explain lightness effects in the rocky displays.

Previous studies have reported that the lightness of a surface is affected by the perceived lightness of the background rather than its physical luminance (Arend et al., 1971; Evans, 1948; Gelb, 1932; Gilchrist, 1988; Gilchrist et al., 1983; Hsia, 1943; Jaensch & Müller, 1920; Landauer & Rodger, 1964; Oyama, 1968). Experiment 5B tested how observers perceived the lightness of the surrounds in Experiment 5A to better understand the lightness effects observed in Experiment 5A.

Methods

Observers

Twenty-five first-year psychology students participated in Experiment 5A. One observer was excluded from analyses (see results and discussion section), leaving 24 observers in total.

Five observers participated in Experiment 5B. Four of them were observers AS, DC, RS and KT[8] from Experiment 1, and the fifth was observer KD who had not

---

[8] Due to a bug in the program, observer KT did not complete all three repeats for each condition (he completed 123 out of 192 trials). This bug was subsequently fixed for all other observers.

participated in any previous experiments. The participants who were not part of the lab (DC, RS, and KD) were compensated $20 for the experiment, which lasted 1 hour.

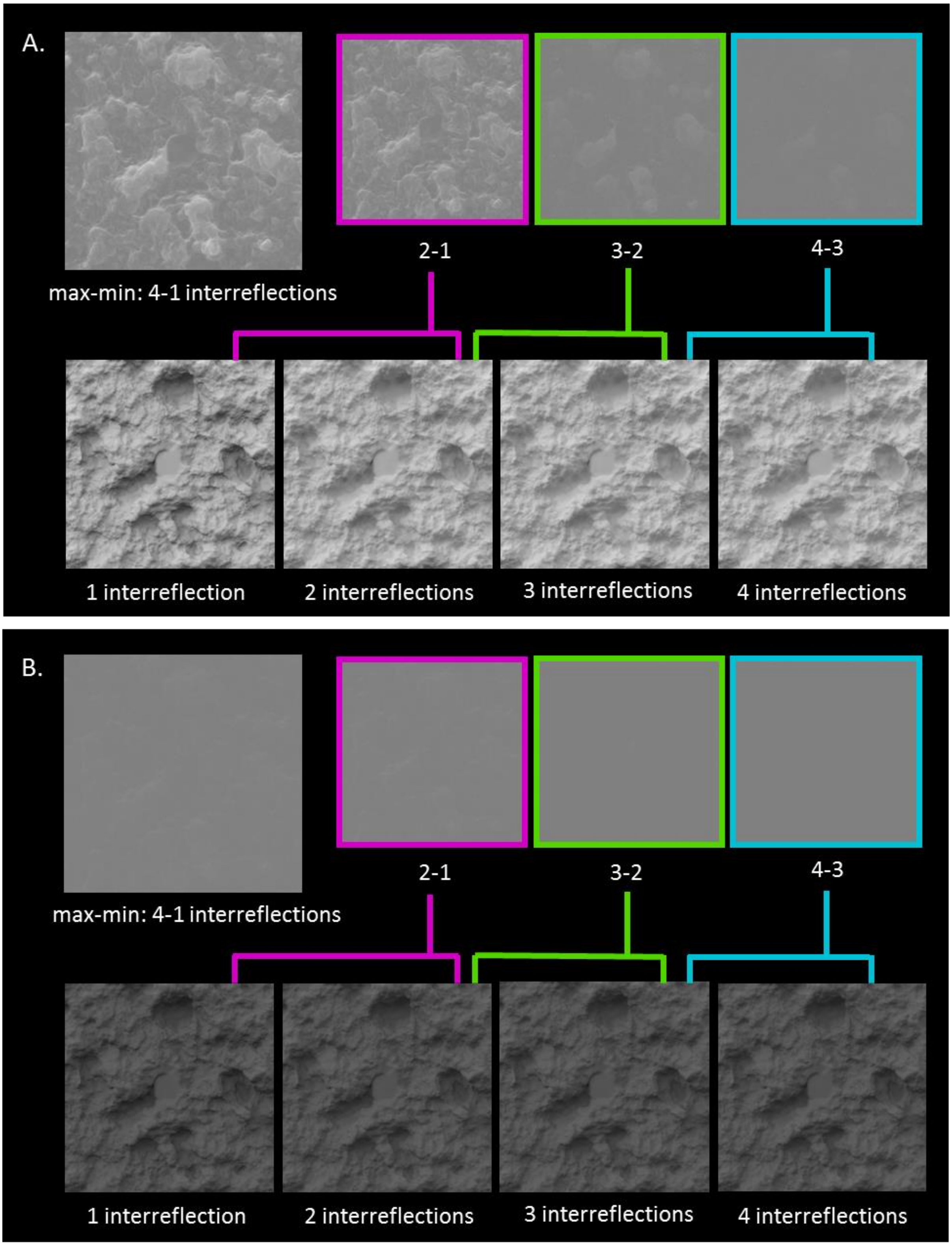


Figure 3.1. Examples of difference maps between images rendered with 1, 2, 3 and 4 interreflections, for the matte white (A) and grey (B) surrounds rendered under bright illumination. The bottom row of images in (A) and (B) shows the appearance of surfaces when rendered with different numbers of interreflections. The top row shows the difference images, with whiter areas indicating a larger difference between the images being compared, and mid-grey areas indicating no difference. Adding interreflections to the white displays produced much larger physical differences in shading contrast compared to the grey displays.

Apparatus and stimuli

The task, stimulus construction, and presentation were the same as in Chapter 2 experiments. The adjustable display was the same display used in Experiments 1, 2, and 4 (i.e. a black and white checkerboard rendered with high relief with a glossy surface reflectance, Figure 2.1E). The test displays contained rocky surfaces that varied in (simulated) albedo and gloss level, with flat matte centres that varied in albedo. Stimuli were rendered under the “grove” illumination field (Debevec, 1998; Figure 2.3), either under “bright” (high) illumination, which was the same illumination level used in Chapter 2, or under “dim” (low) illumination, which was 25% of the bright illumination. Surfaces were rendered with 1, 2, 3, or 4 interreflections. The surfaces were initially rendered with up to 8 interreflections. However, no physical differences were found between images rendered with 4 or more interreflections, so displays with more than 4 interreflections were not used in the experiment. Figure 3.1 shows examples of difference maps between images rendered with different numbers of interreflections for displays with white (90% reflectance; Figure 3.1A) and grey (19.8% reflectance; Figure 3.1B) surrounds. These were the two surround albedo values included in the experiment. The white surround was chosen to provide the optimal conditions for interreflections to modulate shading contrast (see Figure 3.1A). White surfaces reflect more incident light, which indirectly illuminates other parts of the surface and fills in shadows. Figure 3.1A demonstrates how increasing the number of interreflections in the white surround visibly reduced the contrast of the shading. The grey surround was chosen as a comparison to white for two reasons. First, the effect of rendering a greater number of interreflections for darker surfaces became negligible very quickly. The difference maps in Figure 3.1B show that even for 19.8% reflectance, adding interreflections did not affect the contrast of shadows much. Second, when darker surfaces were rendered under dim illumination, the resulting image was so dark that it was difficult to perceive surface structure.

An additional advantage of the chosen surround albedos was that the image of the grey surface under bright illumination was almost identical to the image of the white surface under dim illumination rendered with one interreflection (compare the top two images in Figure 3.2A). Consequently, we could directly compare how the number of interreflections between the two surrounds influenced areas of shading and shadows. Figure 3.2B plots luminance profiles at different cross sections of the images in Figure 3.2A. In these plots, areas of surface shading are indicated by dips in the luminance profiles. Greater dips mean higher contrast shading in that location. The first column of graphs in Figure 3.2B reveals very similar luminance profiles for the white and grey images rendered with 1 interreflection. The second column of graphs plots the luminance profiles of the same surfaces rendered with 4 interreflections. For the white surfaces (orange line), the largest areas of shading contrast were “filled in” by the additional interreflections, substantially reducing the contrast of the shading (and also increasing the mean luminance of the surround). The luminance profiles for the grey surfaces (blue line) remained relatively unchanged. These plots, as well as the difference maps in Figure 3.1, demonstrate the physical differences in shading contrast between images of white and grey surfaces rendered with 2 or more interreflections, which could theoretically provide observers with information about surface albedo.

Figure 3.3 shows examples of phase-scrambled stimuli for each interreflection condition. Phase-scrambled stimuli were created in the same way as in Experiment 4, but with one difference. Instead of a circular test patch, the test patch was the same shape as in the rocky surfaces. As in Experiment 4, the regions immediately surrounding the central patch in the 3D rocky and corresponding phase-scrambled stimuli were matched in average luminance.

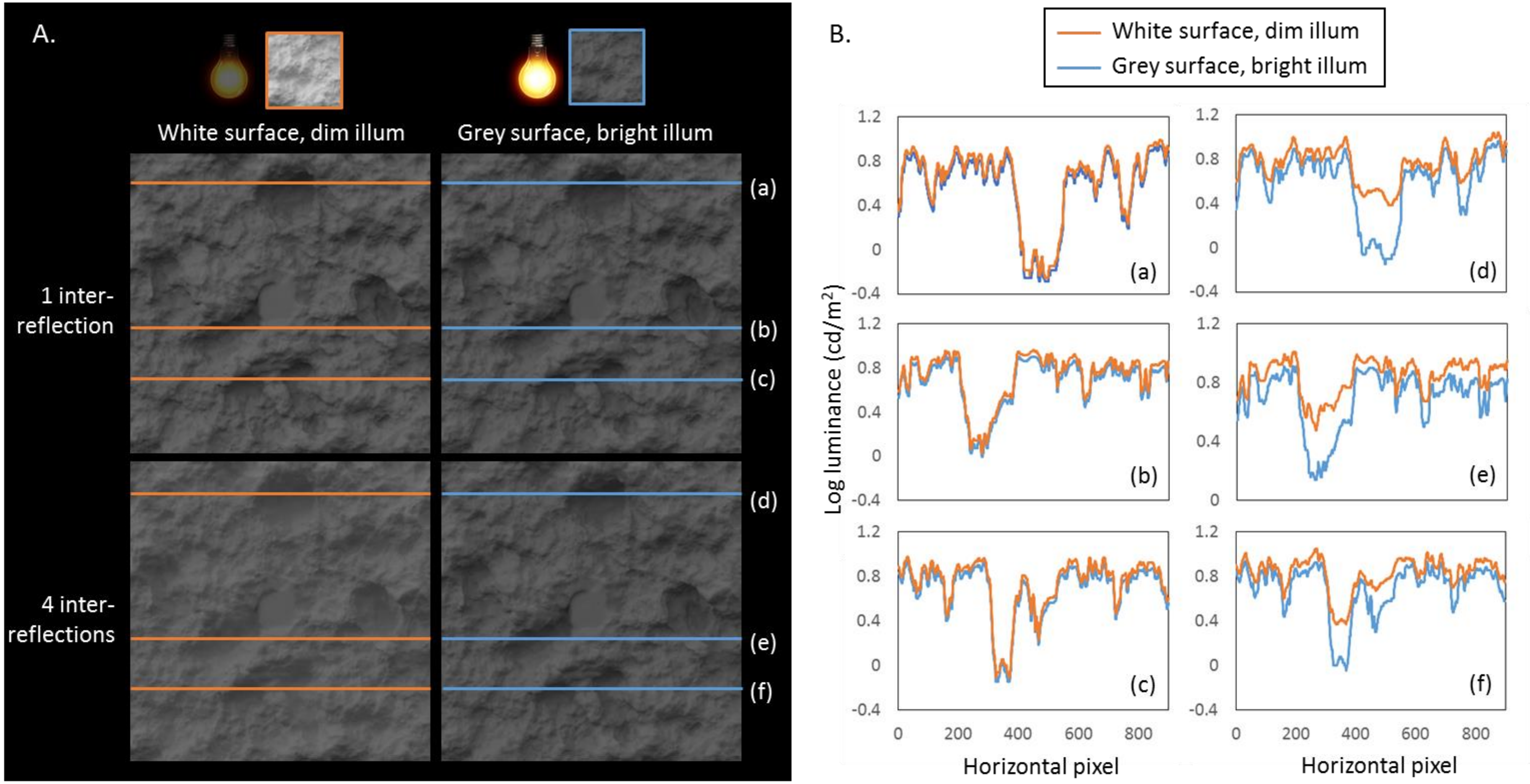


Figure 3.2. Comparison of matte white and grey displays rendered with 1 and 4 interreflections. (A) Images of white displays under dim illumination (first column) and grey displays under bright illumination (second column), for 1 interreflection (top row) and 4 interreflections (bottom row). (B) Luminance profiles at different cross sections of the images in (A). The cross sections (a-f) in (A) correspond to the plots (a-f) in (B).

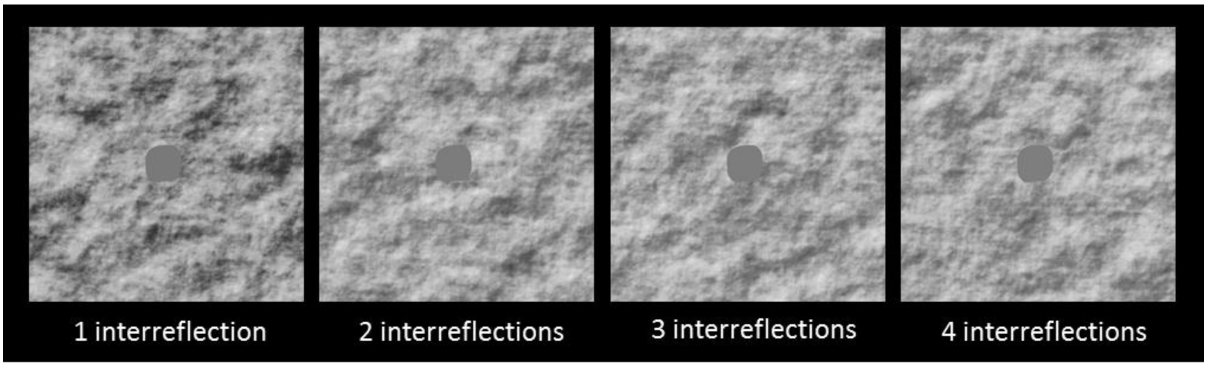


Figure 3.3. Examples of phase-scrambled stimuli for each interreflection condition. These images were created from the matte white surround condition (bright illumination).

Procedure

The presentation of trials was the same as in Chapter 2. The task instructions were also the same but additional instructions were given. In Experiment 5A, observers were instructed to change the lightness of the flat central patch on the adjustable surface until it looked like it was the same lightness or painted with the same paint as the flat central patch on the target surface. Observers were also told that different surfaces could be illuminated differently, which might affect the overall brightness of the surfaces. It was emphasised that they should match the central patches on their intrinsic lightness value, not brightness, which might be affected by the intensity of the illumination. However, it was also emphasised that they should not make cognitive decisions about the lightness of the patches, and that this was best avoided by not spending excessive time on each trial. To show that they understood the task instructions, observers completed eight practice trials in front of the experimenter before moving onto real trials. All observers demonstrated an understanding of the instructions.

There were four surround-type conditions in Experiment 5A: rocky matte, rocky glossy, phase-scrambled matte equivalent, and phase-scrambled gloss equivalent. For each condition there were two surround albedos: white (90% reflectance) and grey (19.8% reflectance). Each surface was rendered under high or low illumination, with 1, 2, 3 or 4 interreflections, resulting in 64 surround conditions in total. For these 64 surround conditions, observers judged three test patch albedos: grey (19.8%), light-grey (46.8%), and white (90% reflectance). Observers did not repeat any conditions, resulting in 192 trials for each observer.

In Experiment 5B, the task and stimulus presentation was the same as in Experiment 5A, except that observers were instructed to adjust the match patch until it appeared to be the same lightness as the surround. The instructions were otherwise the same as in Experiment 5A. The different conditions were also the same as in Experiment 5A, with some exceptions. Observers did not judge the lightness of phase-scrambled surrounds because there was no clear percept of a uniform albedo surface, and therefore no single lightness value to associate with these surrounds. Due to a ceiling effect in Experiment 5A (see results and discussion), the white test patch condition was not included in this experiment, leaving 64 conditions in total. To address the variability that was problematic in Experiment 5A, each observer repeated lightness judgments three times for each condition. This resulted in 192 trials for each observer.

Results and discussion

The results of Experiment 5A are presented in Figures 3.4 and 3.6, for each test patch albedo (grey and light grey), gloss level (matte and glossy), surround albedo (grey and white), surround type (rocky and phase-scrambled), illumination level (high and low), and interreflection condition (1, 2, 3 and 4). The white test patch condition was excluded from analyses due to ceiling effects. Observers reported that they could often not set the adjustable patch high enough to match the white test patch, especially in the high illumination condition. One observer was excluded from analyses because more than 5% of their data was outside 2.5 standard deviations from the mean. The data points in Figure 3.4 show the average lightness settings for each condition, plotted in Munsell values. The data were also plotted in log luminance (Figure 3.6) to determine whether observers’ matches were closer to luminance or Munsell matches.

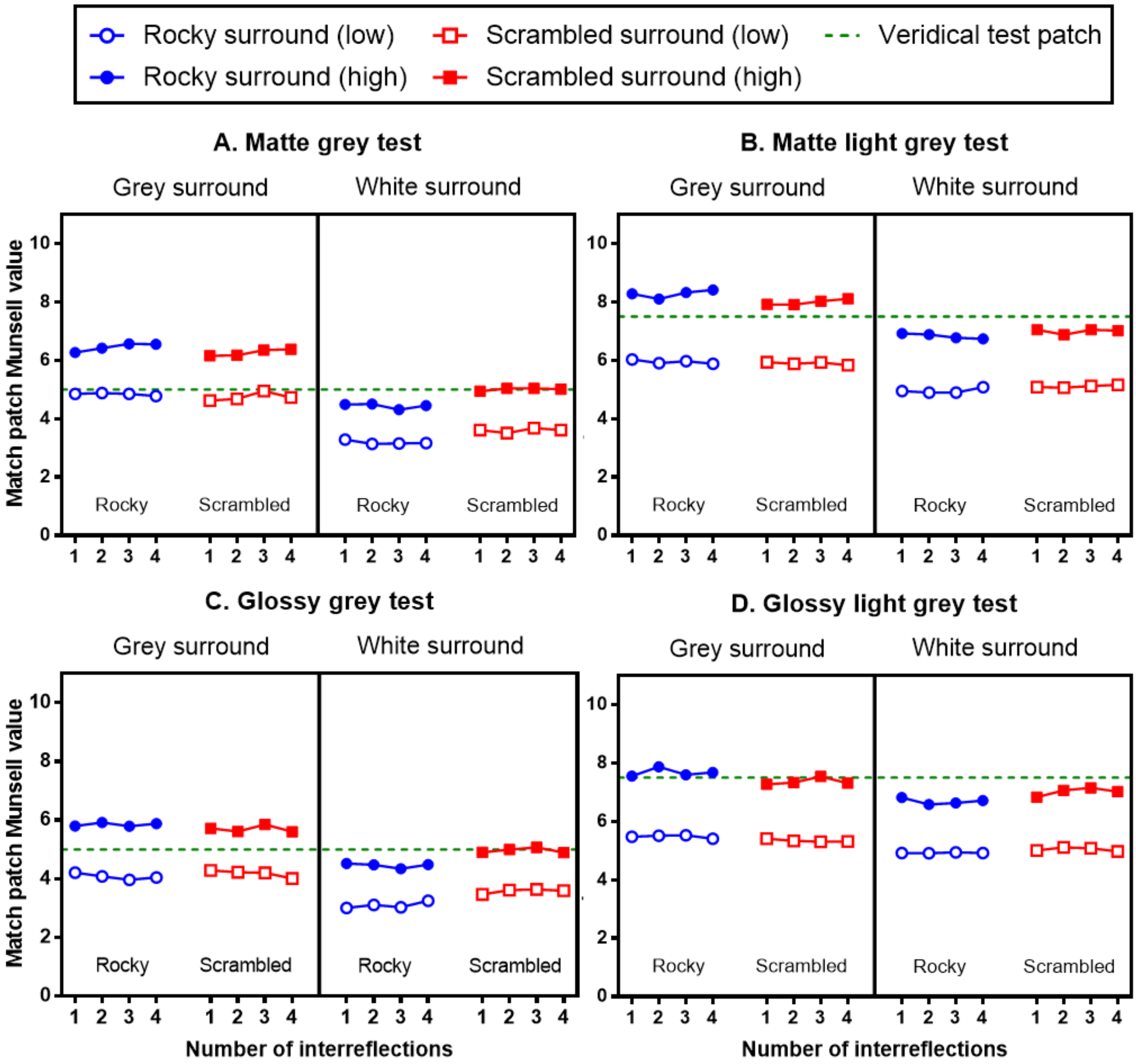


Figure 3.4. Average data for Experiment 5A, plotted in Munsell values. Closed data points are lightness settings for test patches under high illumination, and open data points are settings for test patches under low illumination. The green dotted lines indicate veridical test patch Munsell matches.

The present experiment was designed to investigate how various contextual factors influence the lightness of a test patch, such as features of the surround (albedo, gloss level, and surround type), the prevailing illumination, and the number of interreflections. To simplify the analyses of the conditions of interest, we will discuss differences between the test patch albedo conditions. Light grey test patches (Figure 3.4B and 3.4D) were consistently judged to be lighter than grey test patches (Figure 3.4A and 3.4C). The light grey test patches were both physically lighter and brighter than grey test patches, so it is unsurprising that they were judged to be lighter. Additionally, the difference between high and low illumination settings (closed and open data points, respectively) was larger for the light grey test patch (Figure 3.4B and 3.4D) compared to the grey test patch (Figure 3.4A and 3.4C). This effect is expected if a test patch's lightness is affected by its luminance, which it was in this experiment given that test patches were judged to be lighter under high compared to low illumination (compare filled and open data points in Figure 3.4). In the present experiment, increasing the illumination had a multiplicative effect on test patch luminance. Thus the change in test patch luminance from low to high illumination was greater for the light grey compared to the grey test patch. This greater difference in luminance translated to a greater difference in perceived lightness. The above effects for test patch albedo were common across all conditions. To simplify the comparison of other conditions, further analyses were averaged across test patch albedo.

The data in Figure 3.4 reveal that perceived test patch lightness was affected by the different surround and illumination conditions. However, the number of interreflections did not have a consistent effect on lightness judgments[9]. The average lightness settings in Figure 3.4 were subjected to a within-subjects four-factor ANOVA, with two levels of gloss (matte, glossy), two levels of surround albedo (grey, white), two levels of surround type (rocky, phase-scrambled), and two levels of illumination (low, high). Interreflections and test patch albedo were removed as factors to simplify the analysis. There was a main effect of gloss level, $F(1, 23) = 172.28$, $p < .001$, surround albedo, $F(1, 23) = 126.55$, $p < .001$, surround type, $F(1, 23) = 56.32$, $p < .001$, and illumination level, $F(1, 23) = 1755.29$, $p < .001$. Averaged across the other factors, test

---

[9] To verify that there was no consistent effect of interreflections, the average lightness settings for each interreflection condition were subjected to a one-way ANOVA, for each of the other 32 conditions (i.e. each connected set of data points in Figure 3.4). All ANOVAs were non-significant, all $p > .05$, except for one condition (glossy, rocky, grey surround, with the grey test patch under high illumination). A follow-up trend analysis revealed no significant linear trend, $p > .05$, and follow-up contrasts showed that the settings from 1 interreflection did not differ from the settings with 2, 3, and 4 interreflections. Thus the significant ANOVA for this one condition was not revealing any sensible differences between interreflection conditions, and will not be examined further. Furthermore, when interreflections were included in a multi-factor ANOVA with all other conditions, there was no main effect of interreflections, $F(1, 23) = 0.92$, $p = .435$.

patches were judged to be darker when they were embedded in glossy versus matte surrounds, white versus grey surrounds, and rocky versus scrambled surrounds. Test patches were also judged to be darker in the low illumination versus the high illumination condition. However, there were significant interactions involving all of these conditions, indicating that the differences between levels of the factors reported above varied depending on levels of other factors. In the following sections we report these interactions, and deviations in lightness settings from ground truth.

The effect of gloss level on perceived lightness

Figure 3.5A plots the difference between lightness settings in the grey and white surround conditions, for each gloss level and surround type. There was a significant interaction between gloss level and surround albedo, $F(1, 23) = 118.49$, $p < .001$. Glossy displays led to a smaller difference in perceived lightness between test patches embedded in white and grey surrounds (average Munsell difference of 0.74) compared to matte displays (average Munsell difference of 1.29), averaged across other conditions. This can be seen in Figure 3.4 by looking at the differences between the left panels (grey surround) and right panels (white surround) in Figure 3.4A and 3.4B (matte conditions), and comparing this to differences between the left and right panels in Figure 3.4C and 3.4D (glossy conditions). There was no three-way interaction between gloss level, surround albedo and surround type, $F(1, 23) = 0.003$, $p = .954$, indicating that the differences between matte and glossy displays were similar for the rocky and phase-scrambled conditions. This replicates the findings in Chapter 2, which showed that lightness constancy across changes in background was better for glossy displays (Experiment 1), but this improvement in lightness constancy for glossy displays was not different for the rocky and phase scrambled conditions (Experiment 4).

We also examined whether glossy displays led to more consistent lightness judgments across changes in illumination level. Figure 3.5B plots the difference between lightness settings in the high and low illumination conditions, for each gloss level and surround type. The results revealed that there was no interaction between gloss level and illumination level, $F(1, 23) = 0.001$, $p = .982$, nor was there a three-way interaction between gloss level, illumination level and surround type, $F(1, 23) = 0.02$, $p = .892$. Lightness judgments between high and low illumination conditions did not differ for glossy and matte displays (average Munsell difference of 1.75 for both). This result was similar for the rocky and phase-scrambled conditions. These results show that lightness constancy across changes in illumination did not improve for glossy surfaces compared to matte surfaces, unlike those observed with changes in surround albedo.

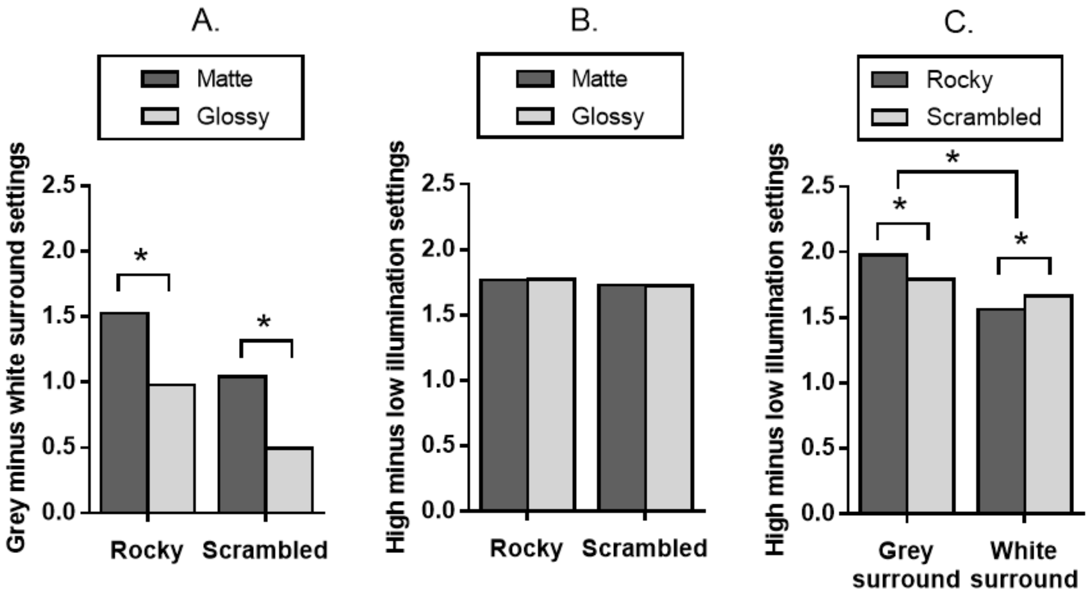


Figure 3.5. Graphs of interaction effects in Experiment 5A. (A) Interaction between surround albedo and gloss level. The difference in settings between white and grey surround conditions is plotted for each gloss level (matte, glossy), and each surround type (rocky, scrambled), averaged across all other factors. (B) No interaction between illumination level and gloss level. The difference in settings between high and low illumination conditions is plotted for each gloss level (matte, glossy), and each surround type (rocky, scrambled). (C) Interaction between illumination level, surround type and surround albedo. The difference in settings between high and low illumination conditions is plotted for each surround albedo (grey, white) and each surround type (rocky, scrambled).

The effect of surround albedo and surround type on perceived lightness

The results show that lightness constancy across changes in illumination varied depending on surround albedo and surround type. There was a significant interaction between surround albedo and illumination level, $F(1, 23) = 31.81$, $p < .001$, and a significant interaction between surround albedo and surround type, $F(1, 23) = 367.34$, $p < .001$. These interactions must be interpreted in light of a significant higher-order interaction between surround albedo, illumination level and surround type, $F(1, 23) = 32.70$, $p < .001$. This three-way interaction is represented in Figure 3.5C. The y-axis is the difference in lightness settings between high and low illumination, and this measure of lightness constancy is plotted for each surround albedo (left and right sets of bars) and each surround type (different coloured bars). A lower score indicates more consistent lightness settings (i.e. better lightness constancy) across changes in illumination.

To investigate this three-way interaction further, lightness settings were subjected to a follow-up two factor ANOVA for each surround albedo (grey and white), with two levels of surround type (rocky, scrambled), and two levels of illumination level (low, high), averaging across all other factors. For both grey and white surround conditions, there was a main effect of illumination level, $F(1, 23) = 1439.30$, $p < .001$ (grey surround), $F(1, 23) = 1204.07$, $p < .001$ (white surround), indicating that test patches under high illumination were set lighter than those under low illumination. This is represented in Figure 3.5C by the positive difference scores for each surround condition. There was also a main effect of surround type, $F(1, 23) = 114.38$, $p < .001$ (grey surround), $F(1, 23) = 232.23$, $p < .001$ (white surround). However, this main effect was opposite for the grey and white surround conditions. When the surround was grey, test patches were perceived to be lighter when embedded in rocky compared to scrambled surfaces. This can be seen in Figure 3.4 by comparing the blue data points to the red data points in the left panel of each graph. Conversely, when the surround was white, test patches were perceived to be lighter when embedded in scrambled compared to rocky surfaces. This can be seen in Figure 3.4 by comparing the blue data points to the red data points in the right panel of each graph. There was an interaction between illumination level and surround type, $F(1, 23) = 28.82$, $p < .001$ (grey surround), $F(1, 23) = 4.58$, $p = .043$. Like the main effect of surround type, this interaction was opposite for grey and white surround conditions. When the surround was grey, lightness judgments between high and low illumination conditions were more consistent for scrambled displays (average Munsell difference of 1.79) compared to rocky displays (average Munsell difference of 1.98). This can be seen in Figure 3.4 by comparing filled data points to the open data points in the left panel of each graph. When the surround was white, lightness judgments between high and low illumination conditions were more consistent for rocky displays (average Munsell difference of 1.56) compared to scrambled displays (average Munsell difference of 1.66). This can be seen in Figure 3.4 by comparing filled data points to open data points in the right panel of each graph. These results will be discussed in the section titled "*Mechanisms responsible for observed lightness effects*".

Deviations in perceived lightness from ground truth

In addition to the differences in lightness judgments between conditions, lightness settings also deviated from ground truth (Figure 3.4, green dotted lines). To investigate these deviations further, test patch settings were plotted in log luminance (Figure 3.6), which helped to determine whether observers were predominantly Munsell matching or luminance matching. Test patch settings were also plotted next to the surround lightness settings obtained in Experiment 5B (Figure 3.7).

Each graph in Figure 3.6 plots the luminance settings for a single test patch albedo in a particular gloss condition. In each graph, the data for grey and white surround conditions are presented in the first and second panel, respectively (blue and red data points for rocky and scrambled conditions, respectively). The third panel plots the theoretical settings indicating perfect test patch luminance matching (green data points), and perfect test patch Munsell matching (green dotted line across all panels). When surfaces were illuminated brightly, the theoretical settings for perfect Munsell matching were almost equivalent to perfect luminance matching (compare green filled-in data points to green dotted line). Indeed, observers' settings were similar to these theoretical values when the surround was white and illuminated brightly (filled data points, middle panels). However, test patches in the grey surround condition were consistently judged to be lighter relative to these theoretical values (filled data points, left panels). When surfaces were illuminated dimly, theoretical Munsell matches were very different to theoretical luminance matches (compare green open data points to green dotted line). Observers' settings (red and blue open data points) were not as low as the theoretical luminance matches (green open circles). Furthermore, the difference in observer settings between the high and low illumination conditions was smaller than the difference in theoretical luminance matches (compare the difference between closed and open red and blue data points to closed and open green data points). Therefore, observers' settings fell between luminance and Munsell matches.

Figure 3.7 shows the results of Experiment 5B (black data points), which are plotted together with the data from the rocky conditions in Experiment 5A (blue data points). The grey surround was perceived to be very similar in lightness to the grey test patch (Figure 3.7A and 3.7C, left panels). Thus, the consistently lighter settings for grey test patches in Experiment 5A may have arisen from the surround being perceived as lighter than it was (on the condition that the grey test patch is perceived to be the same lightness as the surround). The number of interreflections influenced the perceived lightness of the white surround. Dimly illuminated white surfaces appeared lighter for each additional interreflection. Brightly illuminated white surfaces appeared lighter after one interreflection. However, there was a ceiling effect with additional interreflections where observers were unable to set the adjustable patch lighter. Interestingly, test patch lightness did not seem to be affected by these differences in surround lightness with increasing numbers of interreflections.

Mechanisms responsible for observed lightness effects

The aim of Experiment 5 was to explore whether low-level or mid-level mechanisms are responsible for lightness perception of centre-surround surfaces under changing illumination, and to examine how perceived lightness is affected by the number of interreflections. The data from Experiment 5A showed that brightly illuminated test patches were consistently judged to be lighter than dimly illuminated test patches, indicating that observers were not completely reflectance matching. However, the log luminance plots revealed that observers were not just matching test patches on luminance either. This is consistent with previous findings in the literature (Madigan & Brainard, 2014; Ripamonti et al., 2004). Therefore, it is unlikely that lightness effects in Chapters 2 and 3 can purely be attributed solely to brightness differences, which limits the explanatory power of models of perception that do not differentiate brightness from lightness. This will be discussed further in Chapter 5.

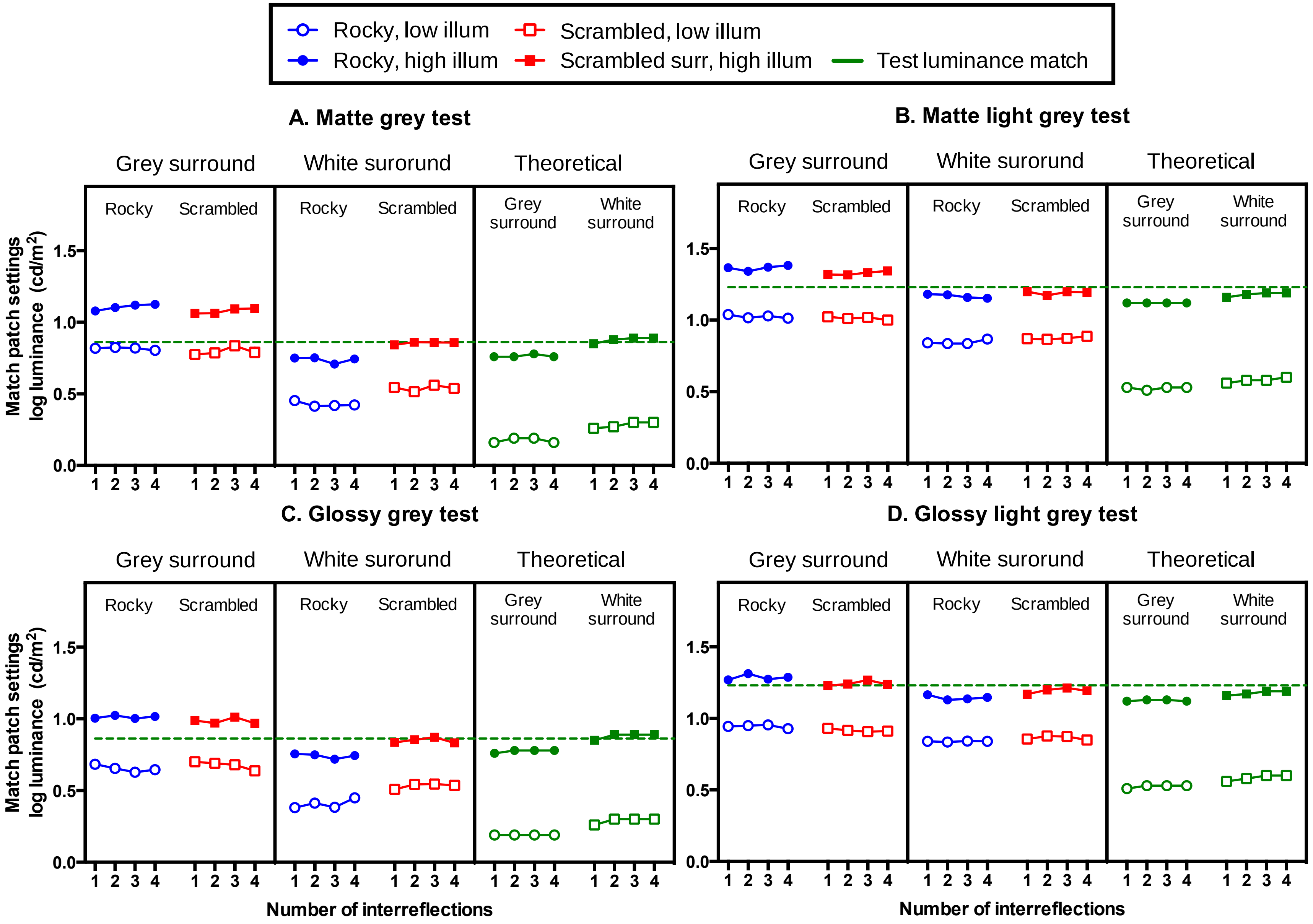

Rocky, low illum
Rocky, high illum
Scrambled, low illum
Scrambled surr, high illum
Test luminance match
A. Matte grey test
B. Matte light grey test
C. Glossy grey test
D. Glossy light grey test
Grey surround
White surorund
Theoretical
Rocky
Scrambled
Grey surround
White surround
Match patch settings log luminance (cd/m²)
0.0
0.5
1.0
1.5
1 2 3 4
Number of interreflections

Figure 3.6. Average data for Experiment 5A, plotted in log luminance. In each graph, the first and second panel plots the data for the grey and white surround conditions, respectively (blue and red data points for rocky and phase-scrambled conditions, respectively). The third panel plots the theoretical settings indicating perfect test patch luminance matching (green data points), and perfect test patch Munsell matching (green dotted line across all panels).

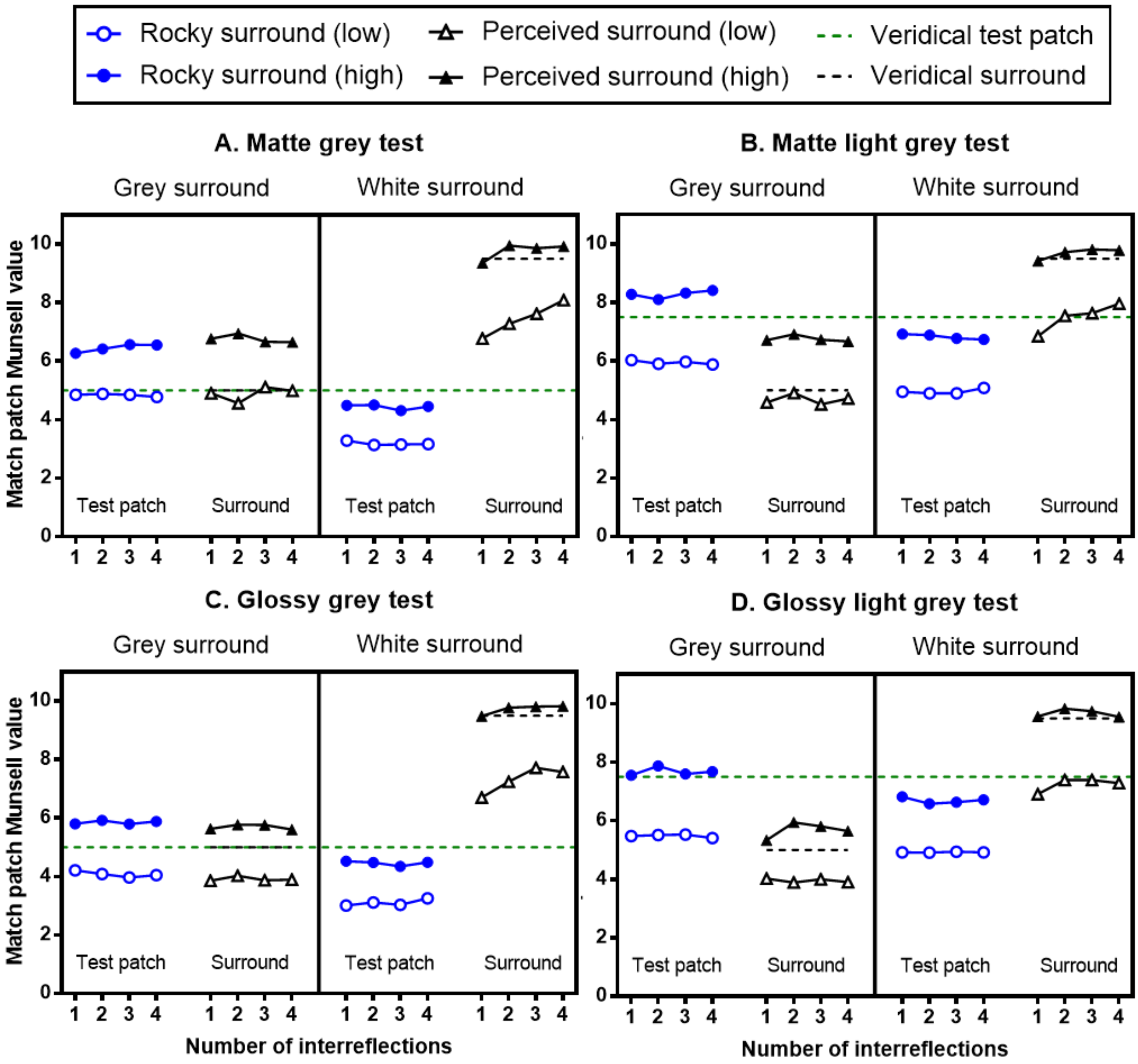


Figure 3.7. Average surround Munsell settings from Experiment 5B, plotted next to test patch settings from Experiment 5A. Both the surround (black data points) and test patch (blue data points) Munsell settings were taken from displays with rocky surrounds. The green dotted lines indicate veridical test patch Munsell matches (Experiment 5A). The black dotted lines indicate veridical surround Munsell matches (Experiment 5B).

The results also showed that test patches embedded in white surrounds were consistently judged to be darker than those embedded in grey surrounds, and test patches embedded in glossy surrounds were consistently judged to be darker than those embedded in matte surrounds. These effects were replications of effects in Chapter 2 and might be attributed to white and glossy surfaces being overall brighter than grey and matte surfaces, respectively (see Table 2.2). Finally, test patches embedded in grey surrounds were consistently set lighter than both veridical Munsell matches and veridical luminance matches. Experiment 5B suggested that the likely cause of this was the surround appearing lighter. At least two factors might have contributed to the grey rocky surround looking lighter. First, an assumption in the present experiments was that surface mesostructure was perceived to be the same depth in the various surround albedo conditions. However, greater contrast in a surface might have been interpreted as a deeper surface. Thus, grey surfaces could have appeared lighter than they were but deeper and more dimly illuminated. An alternative perception could have been that illumination was not along the line of sight, but grazing the surface, causing darker shadows in a lighter surface. Both explanations are equally plausible and cannot be differentiated by the present data. However, since grey test patch settings were similar between rocky and scrambled displays, these explanations cannot be differentiated from a simpler explanation, such as regression to the mean, or differences between the test and matching surfaces.

For white-surround displays, test patches surrounded by rocky surfaces exhibited better lightness constancy under changing illumination compared to phase-scrambled displays. This result is consistent with the hypothesis that white surfaces contained more information to constrain lightness compared to grey surfaces. However, the opposite pattern was found for test patches surrounded by grey surfaces, where phase-scrambled displays exhibited better lightness constancy compared to rocky displays. It is unclear why the opposite pattern would occur for grey surrounds based on the explanation given above. Furthermore, these effects were extremely small (less than half a Munsell value at most). It is possible that these effects were merely due to perceptual differences in the luminance range and/or contrast of rocky and scrambled surrounds. Even though they were physically matched on their range and distribution of luminance values, looking at the rocky glossy display in Figure 2.1 and phase-scrambled glossy equivalent display in Figure 2.15, the contrast of the brightest points (specular highlights in the rocky display) appears to be greater in the phase-scrambled display.

An unexpected finding was that lightness settings did not reliably differ between interreflection conditions. This was the case even though the number of interreflections affected the perceived lightness of the white surrounds (Experiment 5B). Dimly lit white surrounds were judged to be lighter with the addition of each interreflection. However, this increase in lightness was accompanied by the surfaces becoming brighter on average with the addition of each interreflection. Thus the effect of interreflections on perceived surround lightness cannot be separated from the effect of brightness on lightness. It does not make sense to have observers judge the lightness of a phase-scrambled surround (there is no clear perception of a uniform albedo surface), however future research should examine the effect of the number of interreflections on surfaces with complex mesostructure while holding constant the average brightness of the surfaces.

The above results have implications for the conclusions made in Chapter 2. The results of Chapter 2 showed that glossy surfaces led to better lightness constancy across changes in surround albedo compared to matte surfaces. This pattern was not reliably different for rocky and phase-scrambled stimuli, so it was concluded that low-level differences in contrast and luminance distributions were sufficient for explaining this effect. This finding was replicated in the present experiment where up to 4 interreflections were used. This suggests that the ability to perceptually differentiate light and dark surrounds based on the contrast of shadows does not modulate the perceived lightness of embedded test patches. Although the number of interreflections did not affect test patch lightness (Experiment 5A), it did affect lightness judgments of the surrounds (Experiment 5B). This suggests that interreflections affect the apparent lightness of surfaces that are involved in creating the interreflections, which is consistent with previous studies (Gilchrist & Jacobsen, 1984; Ruppertsberg & Bloj, 2007). It remains unclear whether these findings (from the present experiment and previous research) were caused by mid- or low-level computations. One possible low-level explanation is that the visual system attributes a higher albedo value to images with minimal contrast and/or structure. The greater number of interreflections generated by white surfaces fills in shading, leading to lower contrast images with less structure. This could be an indication of a lighter surface, regardless of surface-level percepts. Such an explanation fits with previous findings that extremely simple stimuli tend to look white (e.g. surfaces studied in Gilchrist, 2006; Gilchrist et al., 1999).

There was no effect of gloss on test patch lightness constancy across changes in illumination level. In Chapter 2 we speculated that glossy surfaces led to lightness constancy across changes in background albedo because the specular highlights had constant brightness across all background albedos. Thus the range and/or contrast of luminance values in the surround may have better constrained lightness in glossy compared to matte displays. However, changing the brightness of the illumination also changed the brightness of the specular highlights in the glossy displays. Thus, with changes in illumination, lightness may not have been as strongly constrained by luminance range and/or contrast. This idea is supported by the fact that the phase-scrambled displays showed the same pattern of results as the rocky displays.

Overall, Experiment 5 provides little evidence that mid-level mechanisms were involved in lightness perception of centre-surround displays across changes in illumination. However, observers did exhibit some degree of lightness constancy across changes in illumination, as they were not just luminance matching. Chapter 5 will discuss the relevance of low-level models of lightness perception when interpreting the data from Chapter 2 and Chapter 3.

Chapter 4. Perceptual Mechanisms underlying Lightness Perception in Homogeneous Centre-Surround Displays

The aim of experimental Chapter 4 is to understand the qualitatively different pattern of results observed for the homogeneous centre-surround displays in Experiment 1. The experiments in this chapter investigate whether the crispening effect is a mid-level phenomenon involving a layered image decomposition (scission), or whether low-level mechanisms are sufficient to explain this effect.

Experiment 6: Matching the Texture of the Central Patch and the Surround

In Experiment 1, the homogeneous displays that induced the crispening effect had centre and surround textures that were matched and continuous (i.e. uniform centre, uniform surround), whereas rocky displays that did not induce crispening had centre and surround textures that were discontinuous (i.e. uniform centres and rocky surrounds). Previous research has shown that the continuity of the centre and surround texture can influence the strength of simultaneous contrast (Hurlbert & Wolf, 2004; Laurinen et al., 1997). Laurinen et al. (1997) asked observers to judge the lightness of test patches in centre-surround displays containing modulated textures made from filtered white noise. Either the entire display was modulated with the same pattern so that the centre and surround had a consistent, continuous texture, or the centres and surrounds were modulated differently so that the centre and surround had inconsistent textures. The latter was done either by orthogonally orienting the centre and surround noise patterns, or by using different patterns that contained non-overlapping frequency ranges. The results showed that there was a reduction in simultaneous contrast when the centres and surrounds were modulated by two different patterns, compared to when centre and surround textures were continuous.

Laurinen et al'.s (1997) findings suggest that the different pattern of results obtained for the homogeneous and rocky displays in Experiment 1 might have been caused by differences in the consistency of the centres and surrounds. Like simultaneous contrast, the crispening effect is an enhanced lightness difference, and may be a consequence of homogeneous surrounds containing the same continuous, uniform texture as the central patch. The crispening effect may have been absent in the rocky surround conditions due to discontinuous centre and surround textures, i.e. the surrounds were rocky and the centres were flat. Experiment 6 tested whether crispening could be induced in rocky displays with consistent centre-surround textures. We hypothesised that if centre-surround texture consistency causes the crispening effect, then crispening should be observed in rocky displays with continuous rocky test patches. Crispening should not be observed in such displays if centre-surround texture consistency is not sufficient to induce the effect. The perceptual mechanisms that might underlie the effects of centre-surround consistency on lightness will be discussed in light of the results.

Methods

Observers

Thirteen first-year psychology students participated in Experiment 6. All were naïve to the purposes of the experiment.

Apparatus and stimuli

The task, stimulus creation, and presentation were the same as in Chapter 2 experiments. The adjustable display was the same display used in Experiment 1, 2 and 4 (Figure 2.1E). The test stimuli were centre-surround displays presented in Figure 4.1. The surrounds varied in surface relief and gloss level, and the centres varied in surface relief, gloss level and albedo. Only one surround albedo was used (mid-grey, 19.8% reflectance or Munsell 5). We reasoned that this was sufficient because in Experiment 1, the pattern of crispening was found to be the same for all surround albedo conditions. Five different types of displays were created, which are shown in Figure 4.1, rows A-E. The displays from Chapter 2 were used for the homogeneous condition (row A) and the two flat centre, rocky surround conditions (rows D and E; also see Figures 2.1 and 2.5).

New displays were rendered for the remaining two conditions (rocky centre and surround conditions; Figure 4.1, rows B and C). The rocky centres were a continuation of the surround texture, with the same level of surface relief and gloss. The rocky centres only differed from the surrounds in surface albedo. In all other respects, surfaces in these conditions were rendered in exactly the same way as the stimuli in Chapter 2 (see Experiment 1 methods section).

Procedure

The presentation of the trials and task instructions were the same as in Chapter 2. There were five display-type conditions: continuous homogeneous (Figure 4.1A), continuous matte (Figure 4.1B), continuous glossy (Figure 4.1C), discontinuous matte (Figure 4.1D), and discontinuous glossy (Figure 4.1E). In the continuous conditions, the displays contained rocky centres and surrounds, except for the homogeneous condition where displays had flat centres and surrounds. In the discontinuous conditions, the displays had flat matte centres and rocky surrounds. For each of these five display-type conditions there were 11 test patch albedo conditions, resulting in a total of 55 conditions. The test patch values were the same as in Experiment 1 (see Table 2.2), except that the first albedo and last three albedos were not included (see Table 4.2 for test patch values included). These test patch albedos were excluded because we were only interested in test patches involved in the crispening effect, i.e. those close in albedo to the surround. The two immediate increment and two immediate decrement values were the same as the "extra test patch" values used for 19.8% reflectance in Experiment 1 (Table 2.2, fourth row). Figure 4.1 shows these immediate increments and decrements outlined in orange. Each observer performed three repeats per condition, resulting in a total of 165 trials.

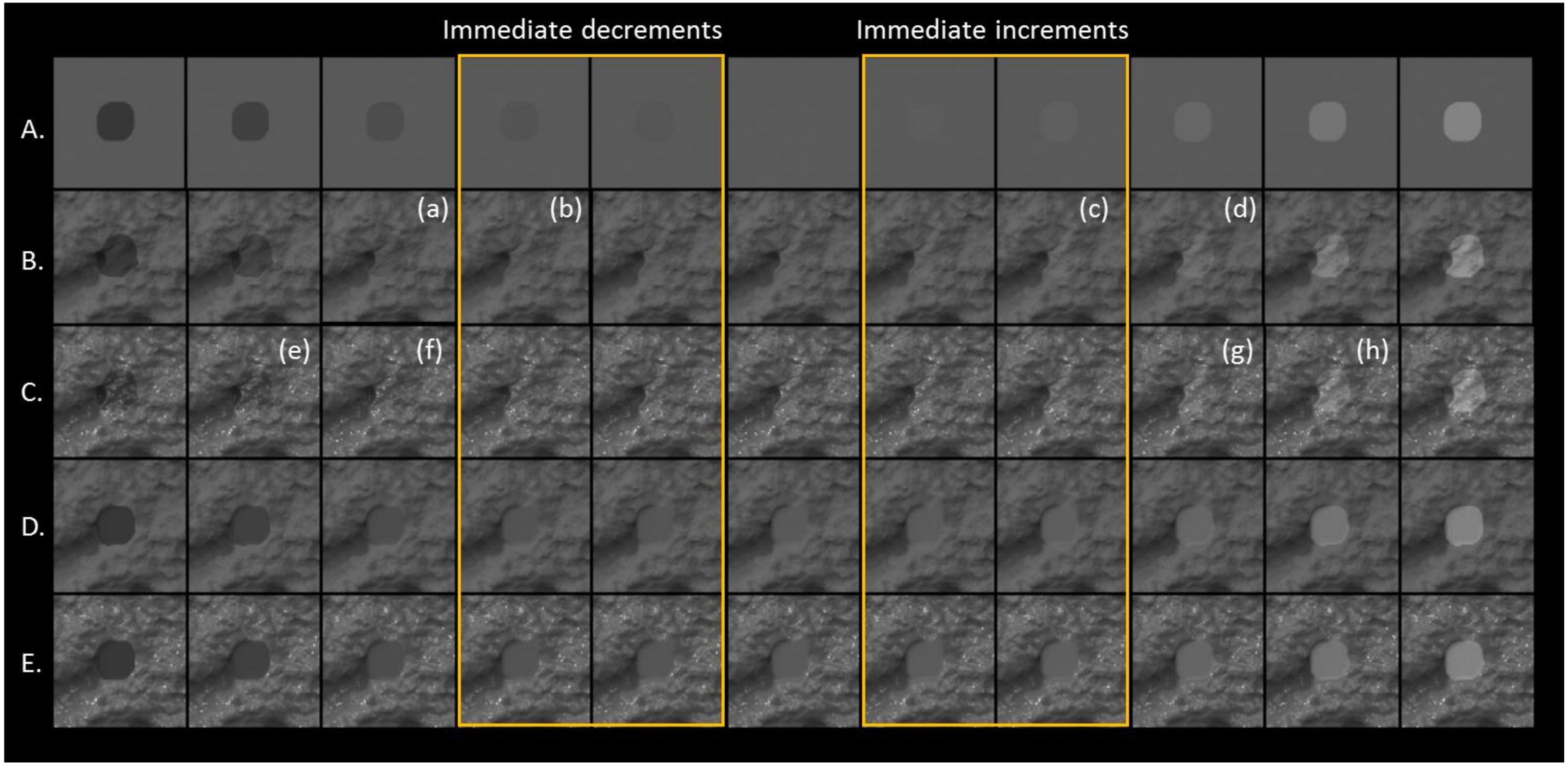

Immediate decrements
Immediate increments
A.
B.
C.
D.
E.
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Figure 4.1. Stimuli used in Experiment 6. The surrounds are cropped in this figure but were presented full size in the experiment. Each row shows the appearance of the test patches for each condition. The first three rows have continuous centre and surround textures: (A) matte flat centre and surround; (B) matte rocky centre and surround; (C) glossy rocky centre and surround. The last two rows have discontinuous centre and surround textures: (D) matte flat centre, matte rocky surround; (E) matte flat centre, glossy rocky surround. The test patches within the orange outlines are low contrast increments and decrements. Note that in the continuous matte and glossy condition (rows B and C) these low contrast test patches appear to be indistinguishable from the surround, whereas in the uniform centre-surround condition (row A) they are visible. See main body text for details about test patches labelled (a-h).

Results and discussion

The results of Experiment 6 are presented in Figure 4.2. The data reveal that the type of display affected test patch lightness settings; the homogeneous, continuous texture, and discontinuous texture conditions gave rise to different patterns in the data. The crispening effect was observed in the homogeneous condition (Figure 4.2A). There was a “step” in lightness settings as the test patch albedo passed through that of the surround. In the discontinuous texture conditions, where the surrounds were rocky and the centres were flat, there was a monotonic relationship between test patch reflectance and adjustable patch settings (Figure 4.2D and 4.2E). These findings replicated the results found in Chapter 2, Experiment 1. The results were less clear for the continuous texture conditions, where both the centres and surrounds were rocky (Figure 4.2B and 4.2C). Unlike the homogeneous condition, crispening was not observed at the increment-decrement transition. Nor was there a monotonic relationship between test patch reflectance and adjustable patch settings. Rather, there was a “flattening” of the data curve at this transition, where test patches close to the surround reflectance were set very similarly to each other. This was followed by a “step” in lightness settings between increments (c) and (d) in the matte condition and (g) and (h) in the glossy condition. These two peculiar patterns (the flattening and the step) in the continuous texture conditions will be further investigated below.

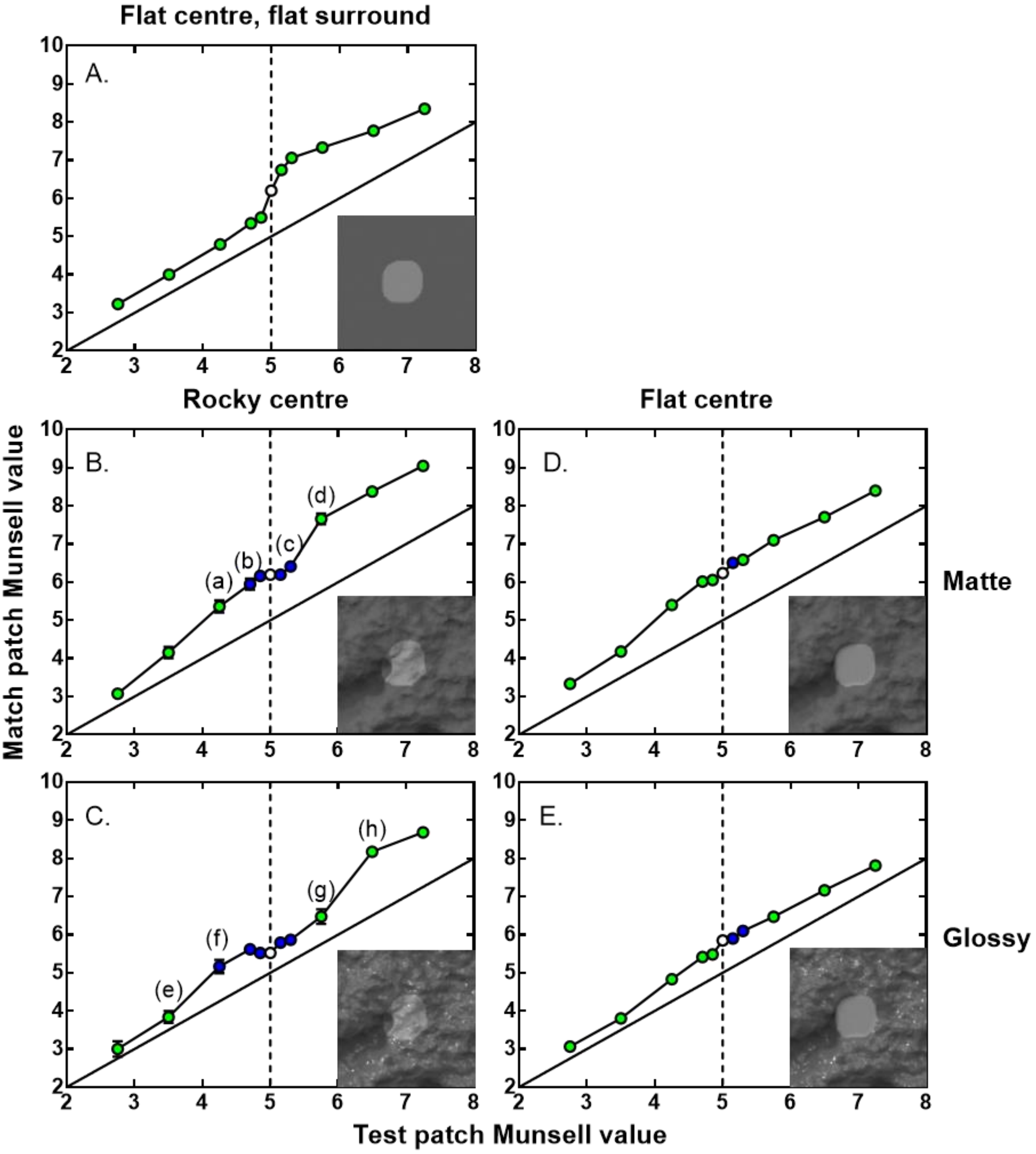


Figure 4.2. Average results for Experiment 6. The graphs (A-E) correspond to the conditions (A-E) in Figure 4.1. The test patch settings (a-h) correspond to the test patches (a-h) in Figure 4.1. The white data points are lightness settings for the surround-albedo test patch, i.e. the test patch that had the same albedo as the surround. The blue data points are test patch lightness settings that did not reliably differ from the surround-albedo test patch settings. The diagonal solid black line represents veridical Munsell matches. The vertical black dotted line indicates the surround Munsell value. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition. In a number of conditions, error bars are smaller than the data points, so are not visible.

Flattening at the increment-decrement transition

Figure 4.2B and 4.2C shows that for the continuous texture conditions there was a flattening of the data curve around the transition from decrements to increments, where test patches close in albedo to the surround (i.e. immediate increments and decrements) were set very similarly. This flattening was statistically verified via paired t-tests with Sidak-corrected alpha values of .00851 per test[10], comparing lightness settings for the surround-albedo test patch (white data point) to the six lowest contrast test patches against the surround (Figures 4.1 and 4.2, test patches a-d for matte and f-g for glossy; see Table 4.1 for *t* values, *df*, and *p* values). For the continuous matte condition, lightness settings for test patches (b) to (c) did not reliably differ from the surround-albedo test patch (Figure 4.2B, blue data points). This implies that the two immediate increments and two immediate decrements were not differentiable from the surround. These results were the same for the continuous glossy condition, but in addition the settings for test patch (f) were also not reliably different from the surround (Figure 4.2C, blue data points). This flattening in the data curve for the continuous texture conditions was compared to the homogeneous and discontinuous texture conditions. For the homogeneous condition (Figure 4.2A) all test patch settings reliably differed from the surround-albedo test patch (white data point). For the discontinuous texture conditions (Figure 4.2D and 4.2E), only the first increment for matte, and first two increments for glossy (blue data points) did not reliably differ from the surround-albedo test patch settings (white data point). Thus, in general, lightness settings for immediate increments and decrements were more similar in the continuous texture conditions compared to the other display-type conditions.

---

[10] Sidak-corrected alpha values were calculated as $\alpha_1 = 1 - (1 - \alpha)^{1/n} = 1 - (1 - 0.05)^{1/6} = .00851$.

	Homogeneous		Continuous matte		Continuous glossy		Discontinuous matte		Discontinuous glossy	
Test patch Munsell	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*
4.25	14.28	**<.001***	5.93	**<.001***	2.77	.0170	8.31	**<.001***	9.62	**<.001***
4.7	7.84	**<.001***	1.56	.145	1.40	.186	3.51	**.0043***	5.45	**<.001***
4.85	5.79	**<.001***	0.38	.711	0.02	.987	4.83	**<.001***	3.81	**.0025***
5.15	3.84	**.00237***	0.01	.991	2.75	.0176	3.08	0.0096	0.44	.671
5.3	6.80	**<.001***	1.38	.192	2.98	.0115	3.17	**.0081***	2.21	.0476
5.75	7.63	**<.001***	9.40	**<.001***	4.89	**<.001***	11.07	**<.001***	5.30	**<.001***

Table 4.1. *t* values and *p* values comparing lightness settings between the surround-albedo test patch to the six lowest contrast test patches against the surround. *$p < .00851$; *df* = 12 for all comparisons.

The stimuli in Figure 4.1 (rows B and C) suggest that the flattening observed in the continuous texture conditions may have been due to a masking effect when test patches were low in contrast against the surround. Specifically, the range and distribution of luminance values *within* the test patch and surround textures may have been masking any differences in lightness *between* the centre and surround. Figure 4.1 demonstrates that for the continuous matte condition (row B), test patches (a) and (d) visibly differ from the surround in lightness, but the test patches in between are barely or not at all visible against the surround. For the continuous glossy condition (row C), test patches (e) and (h) visibly differ from the surround in lightness, whereas the test patches in between are, again, difficult or impossible to detect. In contradistinction, for the homogeneous condition (row A), all test patches that differ in albedo from the surround are clearly visible. Furthermore, in the discontinuous texture conditions (rows D and E), the low contrast test patches appear to be very similar to each other, but they are visible against the surround. This supports the idea of a masking artefact affecting the results when both the centre and surround had a continuous rocky texture.

Step in lightness settings

Figure 4.3 plots the data from the continuous and discontinuous texture conditions together for direct comparison, for matte (left panel) and glossy (right panel) displays. There was a monotonic increase in lightness settings for the discontinuous texture conditions (purple data points), but for the continuous texture conditions (green data points) there was a "step" in lightness settings between increments (c) and (d) in the matte condition, and (g) and (h) in the glossy condition. This step was statistically reliable. Paired t-tests using Sidak-corrected alpha values of .00465 per test[11] were carried out to compare lightness settings in the continuous and discontinuous texture conditions (see Table 4.2 for *t* values, *df*, and *p* values). For the first eight test patch albedos in the matte condition, and the first nine test patch albedos in the glossy condition, there was no significant difference between lightness settings in

---

[11] Sidak-corrected alpha values were calculated as $\alpha_1 = 1 - (1 - \alpha)^{1/n} = 1 - (1 - 0.05)^{1/11} =$ .00465.

the continuous and discontinuous texture conditions. However, for the last three test patch albedos in the matte condition, and the last two test patch albedos in the glossy condition, test patches were judged to be lighter in the continuous compared to the discontinuous texture conditions (compare green and purple data points in Figure 4.3, respectively).

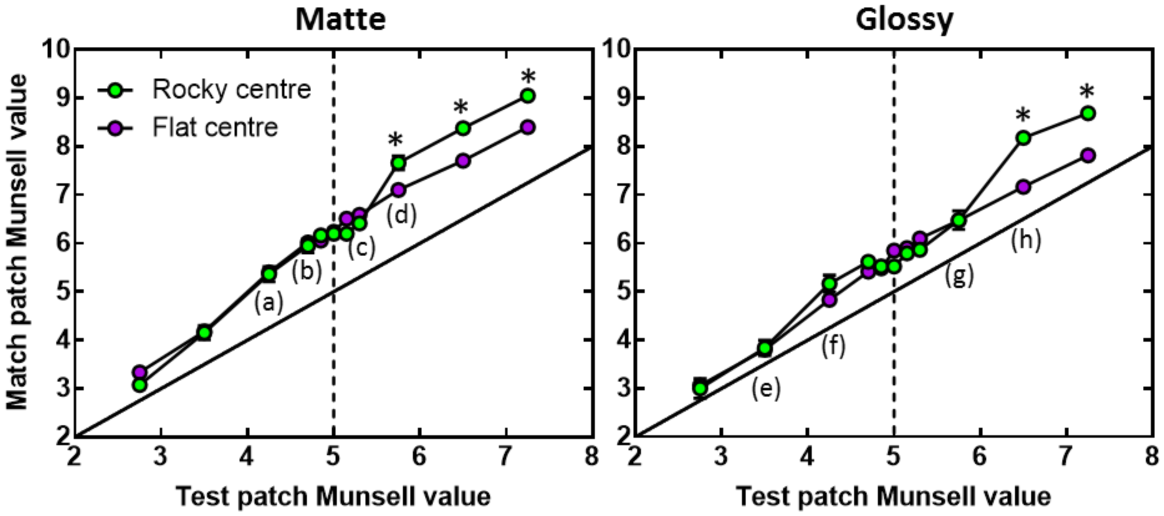


Figure 4.3. Average data from the continuous and discontinuous texture conditions plotted together, for matte (left panel) and glossy (right panel) displays. Stars indicate a significant difference between the continuous (green data points) and discontinuous (purple data points) texture conditions for a particular test patch albedo. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition. In a number of conditions, error bars are smaller than the data points, so are not visible.

The above results suggest that once increments were clearly differentiable from the surround in the continuous texture condition (Figures 4.1 and 4.3, test patches d and h), there was a step in lightness settings where visible increments were perceived to be lighter in the continuous compared to the discontinuous texture condition. It is possible that this step is similar to the crispening effect observed in the homogeneous condition, but the masking artefact was preventing crispening from occurring at the transition from increments to decrements. Another possibility is that observers used the brightest regions of diffuse shading within the test patches to judge surface lightness (e.g. see Todd et al., 2004; Toscani et al., 2013). The displays were illuminated slightly obliquely, meaning that areas of the rocky test patches faced towards the light source, resulting in brighter (and darker) regions compared to the flat test patches. However, if observers used brighter regions in the rocky test patches to judge lightness, then we would have expected decrement test patches to also appear lighter in the continuous texture conditions. This was not the case; the results revealed that there was no reliable difference between decrement test patch settings between the continuous and discontinuous texture conditions. Furthermore, the rocky test patches that were affected by masking should have appeared lighter than their flat test patch counterparts, but no reliable difference was observed between continuous and discontinuous texture conditions for these test patches either. Therefore, the results suggest that crispening may have occurred in the continuous rocky condition, but it was not possible to confirm this due to possible masking artefacts.

	Matte			Glossy		
Test patch Munsell	*df*	*t*	*p*	*df*	*t*	*p*
2.75	12	2.20	.0298	12	0.455	.650
3.5	12	0.26	.795	12	0.260	.795
4.25	12	0.36	.721	12	2.404	.0177
4.7	12	0.56	.574	12	1.476	.143
4.85	12	0.91	.365	12	0.269	.788
5	12	0.33	.746	12	2.404	.0177
5.15	12	2.66	.0090	12	0.761	.448
5.3	12	1.47	.143	12	1.671	.0974
5.75	12	4.73	**<.001***	12	0.0371	.970
6.5	12	5.70	**<.001***	12	7.351	**<.001***
7.25	12	5.48	**<.001***	12	6.274	**<.001***

Table 4.2. *t* values, *p* values, and *df* comparing the lightness settings between the continuous and discontinuous texture conditions, for each test patch Munsell value.
* $p < .00465$

If crispening occurred in the continuous texture conditions, this suggests that crispening in the homogeneous displays may have also been due to continuous centre-surround textures (homogeneous centre, homogeneous surround). Laurinen et al. (1997) suggested that low-level spatially tuned neurons that signal information about surface textures and object edges may modify early brightness signals from the initial stages of cortical processing in area V1 (Olzak & Thomas, 1991, 1992; Thomas & Olzak, 1996). They suggested that centres and surrounds with different textures might isolate different sets of early cortical mechanisms, which could lead to differently modified brightness signals compared to when centres and surrounds have continuous textures. They did not claim to know what mechanisms might underlie the modification of early brightness signals, so this remains unclear. Gilchrist (2006) suggested that differences between Laurinen et al.'s (1997) continuous and discontinuous textured displays were an effect of grouping by the Gestalt principle of similarity. An alternative explanation is based on the idea that the continuous texture displays in Figure 4.1B and 4.1C might have been perceptually decomposed into a transparent foreground layer and a background layer. A consequence of continuous centre and surround textures is that the luminance polarities along the border between the centre and surround are preserved over their entire length. Studies have shown that such conditions lead to the perceptual decomposition of luminance values into a foreground and background layer (Adelson & Anandan, 1990; Anderson, 1997, 1999; Beck et al., 1984, Metelli, 1970, 1974a, 1974b). Wollschläger and Anderson (2009) showed that image decomposition could be evoked in textured displays when the centre and surround had similar textures and satisfied the conditions for transparency (see also Anderson & Khang, 2010). Indeed, in Laurinen et al.'s (1997) continuous pattern displays, the luminance polarities at the borders were consistent with scission. Furthermore, these displays *looked* transparent. The displays in the present experiment may have also evoked percepts of transparency. The test patches in the consistent texture conditions may have been perceptually decomposed into a homogeneous transparent filter overlaying the rocky background, or alternatively the background could have been perceptually interpreted as a filter with a central hole that revealed an opaque rocky surface. From the present experiment we cannot draw conclusions about whether crispening occurred in the continuous textured rocky displays, nor whether observers perceived the test patches or surrounds to be transparent. Experiment 7 investigates whether transparency could be responsible for the crispening effect in homogeneous centre-surround displays.

Experiment 7: Matching Surfaces on Two Dimensions

The crispening effect observed for the homogeneous centre-surround displays in Experiment 1 implies that there could be a qualitative difference in the way test patches are perceived when embedded in homogenous versus inhomogeneous surrounds. Some researchers have suggested that lightness perception may involve more than one perceptual dimension (Ekroll & Faul, 2013; Logvinenko & Maloney, 2006; Vladusich, 2012, 2013; Vladusich et al., 2007). Ekroll, Faul, and colleagues (Ekroll & Faul, 2009, 2012a, 2012b, 2013; Ekroll et al., 2004, 2011) observed crispening in homogeneous coloured centre-surround displays, and suggested that impressions of transparency contributed to the appearance of the central patch. There is a significant body of work in lightness and colour that has shown that layered image decompositions can induce large lightness and colour induction effects (Anderson, 1997; Anderson & Khang, 2010; Anderson & Winawer, 2005, 2008; Wollschläger & Anderson, 2009). In Ekroll and Faul's (2013) study, observers varied the transmittance and/or the colour of an adjustable patch on a variegated surround to match a target on a homogeneous surround. The variegated and homogeneous surrounds had the same average chromaticity, which the authors argued eliminated the influence of von Kries adaptation to the crispening effect. They found that targets embedded in uniform surrounds were better matched when observers were allowed to vary both the physical colour and transmittance of the matching patch on the variegated surround. Furthermore, transmittance and saturation settings were inversely related to the chromatic contrast between the target patch and its surround; at low chromatic contrasts, homogeneous centre-surround stimuli appeared to trigger impressions of transparency with the target region being perceptually divided into a saturated, transparent filter layer and a background layer the same colour and saturation as the surround. It has yet to be explored whether this finding extends to lightness displays.

The aim of Experiment 7 is to determine whether crispening in achromatic homogeneous centre-surround displays is caused by mid-level mechanisms involving perceptual decomposition (scission) of displays into layers, as suggested by Ekroll and Faul (2013). In this experiment, observers were able to vary both the lightness and transmittance of an adjustable patch overlaying a rocky surround to match the test patch embedded in the homogeneous surround. We hypothesised that if crispening in lightness displays is caused by scission of the test patch and background, then perceived test patch transparency should increase as the centre-surround contrast is reduced, i.e. where the crispening effect occurs. Figure 4.4 illustrates how this inverse relationship between centre-surround contrast and perceptual transparency would influence test patch lightness. If perceptual transparency increases as the test patch becomes lower in contrast against the surround, more of the background would become visible through the test patch. For increments (Figure 4.4A), this means more of the "blackness" in the test patch would be attributed to the darker background layer, causing the filter layer (the test patch) to appear lighter. For decrements (Figure 4.4B), as transparency increases, more of the "whiteness" in the test patch would be attributed to the lighter background layer, causing the test patch to appear darker.

The idea that perceptual decomposition affects lightness in homogeneous centre-surround displays is appealing because it addresses reports of the difficulty in such displays of making truly satisfactory asymmetric lightness or colour matches (Ekroll et al., 2004; Faul et al., 2008; Gelb 1929; Vladusich et al., 2007). It also captures the flimsy or wispy appearance of low contrast test patches embedded in homogeneous surrounds, compared to higher contrast test patches, or those embedded in textured or variegated surrounds (see Figure 2.5). However, it is important to recognise low-level differences between the homogeneous and rocky stimuli that may have contributed to differences in the pattern of results in Experiment 1. For example, the apparent contrast of the test patches against the surround differed between homogeneous and rocky displays. Figure 2.5 shows that, in the homogeneous displays, varying test patch reflectance also changed the apparent contrast of the test patch against the surround. This did not occur as much in the rocky displays. Furthermore, adjusting the lightness of the matching patch in Experiment 1 did not allow observers to match the contrast of test patches in homogeneous displays. Adding a transmittance setting would have at least two effects on the adjustable patch. First, increasing transmittance would increase the transparency of the patch, allowing the texture of the background to be seen through the filter. Second, increasing the transmittance would decrease the contrast of the adjustable patch against the surround. It is possible that in Ekroll & Faul's (2013) study, adding the transmittance setting allowed observers to better match the chromatic contrast of the test patch against the surround. In the present experiment, we divided participants into different instructional conditions (see methods section) in an attempt to explore which mechanisms might be responsible for lightness effects in homogeneous centre-surround displays.

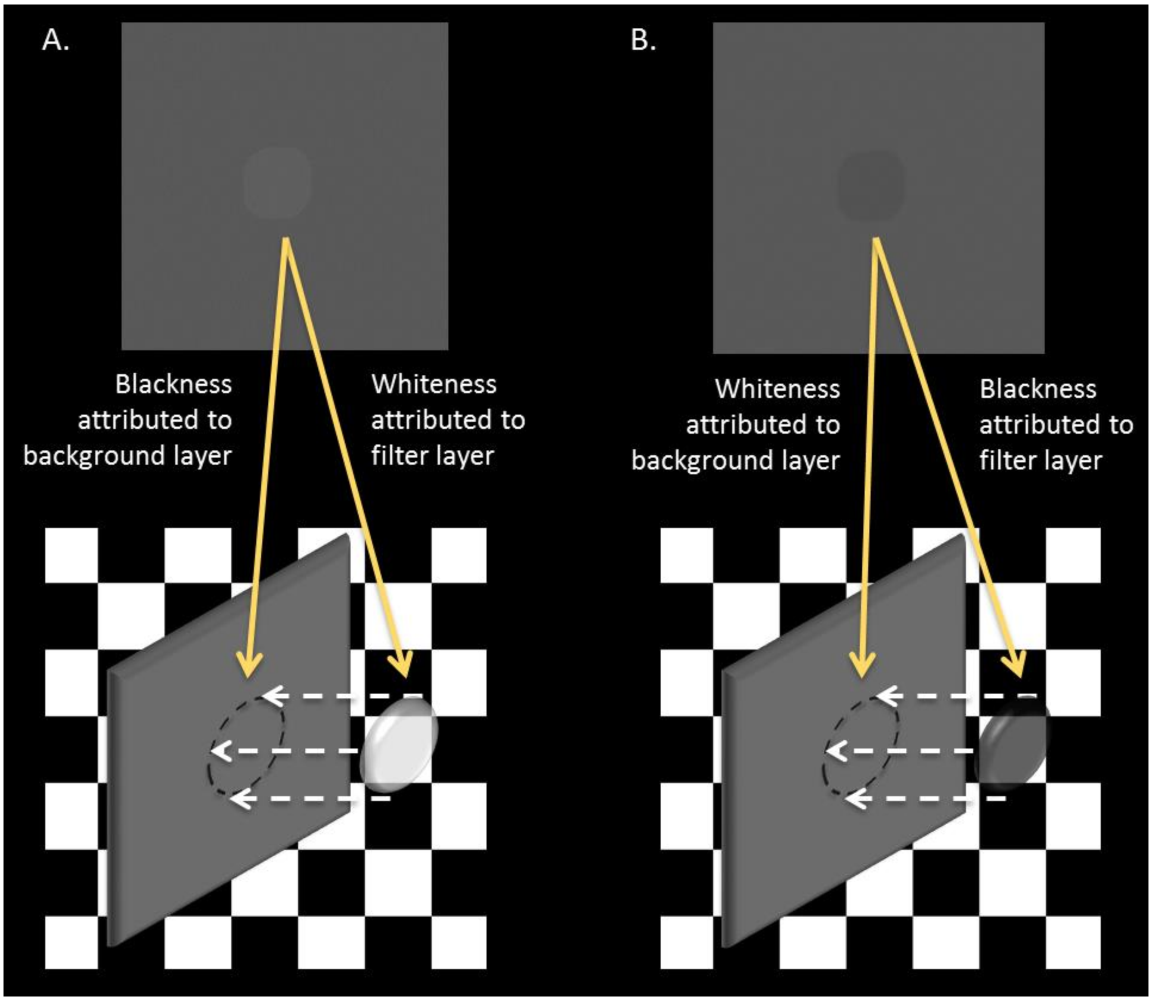


Figure 4.4. Illustration of perceptual decomposition (scission) of homogeneous centre-surround displays. (A) Displays with increment test patches are divided into a light coloured transparent test patch layer and an opaque continuous grey surround layer. (B) Displays with decrement test patches are divided into a dark coloured transparent test patch layer and an opaque continuous grey surround layer.

Methods

Observers

Sixty first-year psychology students participated in Experiment 7. Twenty observers were assigned to a transparency instructions condition, twenty observers were assigned to a contrast instructions condition, and the remaining 20 observers were assigned to a no-transmittance control condition (see procedure).

Apparatus and stimuli

The test stimuli were the matte homogeneous centre-surround displays used in Experiment 1 (see Figure 2.1A and Figure 2.5). The adjustable display was a rocky surface overlayed with a circular central disk that could vary in simulated albedo and transmittance (Figure 4.5A). The rocky backgrounds of the adjustable displays were created in Blender in the same way as the continuous texture condition in Experiment 6. The central disk in the adjustable display was created in Matlab using Metelli's episcotister model to calculate the luminance of each pixel in the location of the central disk on the monitor (see Beck et al., 1984; Metelli, 1970, 1974a, 1974b; Singh and Anderson, 2002). For a given albedo and transmittance, each pixel in the adjustable patch was calculated using Metelli's formula:

$$p = \alpha b + t(1-\alpha) \tag{4.1}$$

where p is the luminance value of the pixel in the region of overlay, α is the transmittance value of the disk (the fraction of light passing through the foreground disk), b is the luminance of the background pixel, and t is the amount of light reflected by the disk (i.e. the luminance of the disk if it were opaque).[12] The transmittance value α was bound between 0 (completely opaque) and 1 (completely transparent), inclusive. In the no-transmittance control condition the central disk had a fixed transmittance value of 0 (completely opaque) and could only vary in albedo.

---

[12] Note that t could be converted to reflectance or Munsell values, given that we knew the average luminance of an opaque test patch with a given reflectance illuminated by the grove light field.

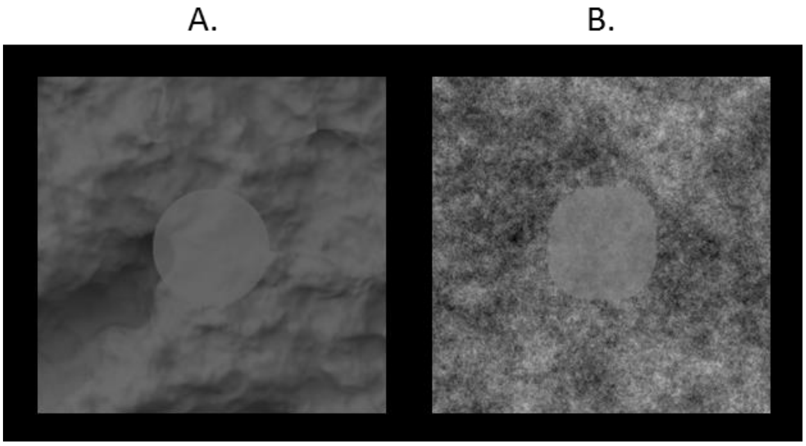


Figure 4.5. Adjustable displays used in Experiments 7 and 9. (A) Example of an adjustable surface in Experiment 7. The image has been cropped in this figure but the rocky background surface was the same size as the homogeneous test surfaces. (B) The adjustable surface used in Experiment 9. This image is not cropped.

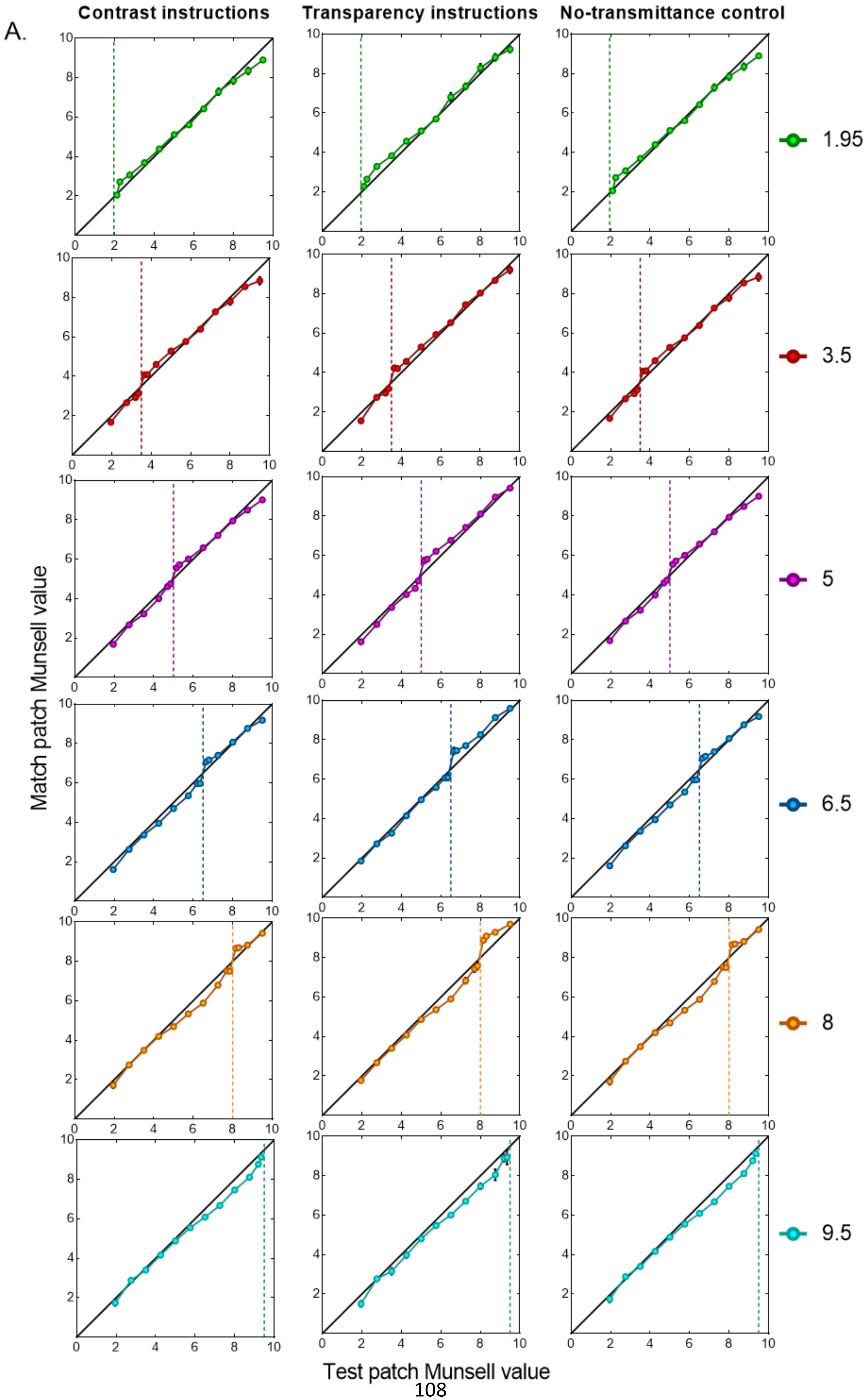
Procedure

Observers judged various aspects of flat target patches embedded in a homogeneous surround via an asymmetric matching task. In each trial, a homogeneous centre-surround test surface (14.88°) was presented on the left side of the computer screen. On the right side of the screen was a matching surface (also 14.88°) with a rocky surround and adjustable central patch. Observers could adjust the lightness of the adjustable patch by moving the mouse left and right, and could also adjust its transmittance by moving the mouse up and down. The test and adjustable surfaces were separated by 15.25° of visual angle (centre to centre) and were presented against a black background. There were three instruction conditions. In all conditions, observers were instructed to first change the lightness of the adjustable patch until it looked like it was the same lightness or painted with the same paint as the test patch. Moving the mouse left made the patch darker, and moving it right made the patch lighter. After observers matched the lightness of the test and adjustable patches, they were instructed to set the transmittance of the adjustable patch by moving the mouse down to make the patch more transparent, or up to make the patch more opaque. Observers given transparency instructions were told that if they perceived the test patch to be transparent they should adjust the transmittance of the adjustable patch until it had the same amount of transparency as the test patch. Observers given the contrast instructions were instructed to adjust the transmittance of the adjustable patch until it was the same contrast against the background, or same visibility against the background, as the test patch. In the no-transmittance control condition observers could not vary the transmittance of the adjustable patch, so were only asked to judge the lightness of the test patch. Note that the word “transmittance” was not used in any of the instructional conditions; we just use this term here to point to the dial that observers adjusted, i.e. the vertical mouse adjustment. Once observers had made the transmittance adjustment, they were instructed to fine tune the lightness and transmittance settings by making minor adjustments to each until a satisfactory match was reached. Note that for the transparency and contrast instruction conditions, observers were able to vary both the lightness and transmittance of the adjustable patch at all times (i.e. it was a two-dimensional matching task), but they were instructed to adjust lightness first.

The test surface was always a homogeneous centre-surround display. As in Experiment 1, there were six surround albedo conditions, and 13 to 15 test patch albedo conditions (see Table 2.2). The surround albedo of the matching surface varied from trial to trial, and was the same as the test surface’s surround albedo. There were also two gloss-level conditions (matte or glossy) for the background of the matching surface. This resulted in a total of 172 trials for each observer.

Results and discussion

The results of Experiment 7 are presented in Figures 4.6 and 4.8. Figure 4.6 shows the test patch lightness settings for each instruction condition, plotted in Munsell values. Figure 4.8 plots the transmittance settings for the contrast and transparency instruction conditions. The crispening effect was observed in all three instruction conditions, i.e. there was a “step” in lightness settings as the test patch albedo passed through that of the surround (Figure 4.6). To compare the size of the crispening effect in each instruction condition, difference scores were obtained by subtracting the lowest contrast (against the surround) decrement settings from the lowest contrast increment settings. These difference scores are plotted in Figure 4.7. A lower score indicates a smaller step and therefore less crispening. As in Experiment 1, the darkest (Munsell value 1.95) and lightest (Munsell value 9.5) surround conditions were omitted because they contained only increments or only decrements, respectively.


A.
Contrast instructions
Transparency instructions
No-transmittance control
1.95
3.5
5
6.5
8
9.5
Match patch Munsell value
Test patch Munsell value
0
2
4
6
8
10

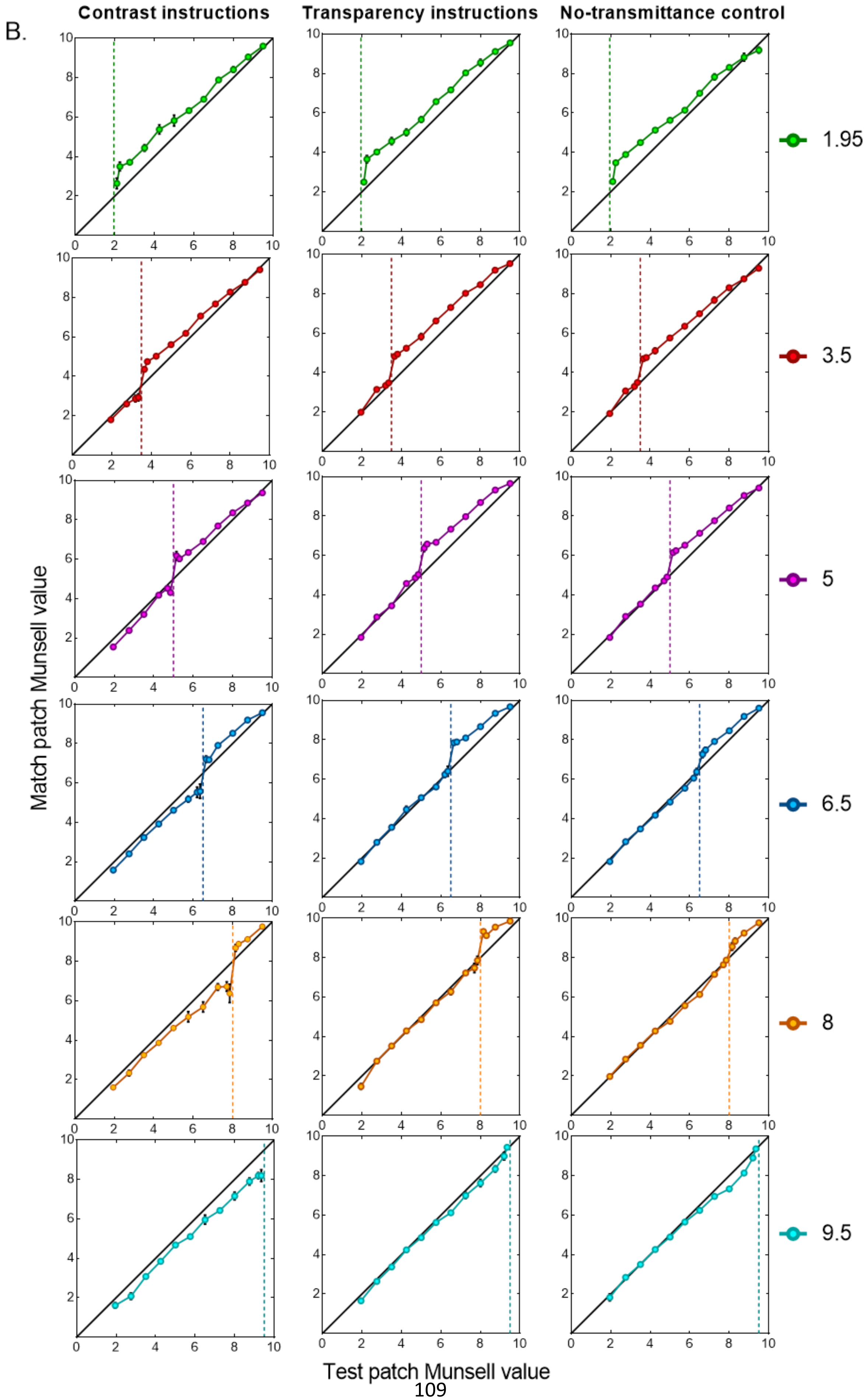

B.
Contrast instructions
Transparency instructions
No-transmittance control
1.95
3.5
5
6.5
8
9.5
Match patch Munsell value
Test patch Munsell value
0
2
4
6
8
10

Figure 4.6. Lightness settings for Experiment 7, plotted in Munsell values. (A) Lightness settings for the matte adjustable surface condition. (B) Lightness settings for the glossy adjustable surface condition. Each column shows the settings for different instruction conditions (contrast, transparency, and no-transmittance). Each row shows the settings for different surround-albedo conditions (Munsell 1.95, 3.5, 5, 6.5, 8, and 9.5). Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition. In a number of conditions, error bars are smaller than the data points, so are not visible.

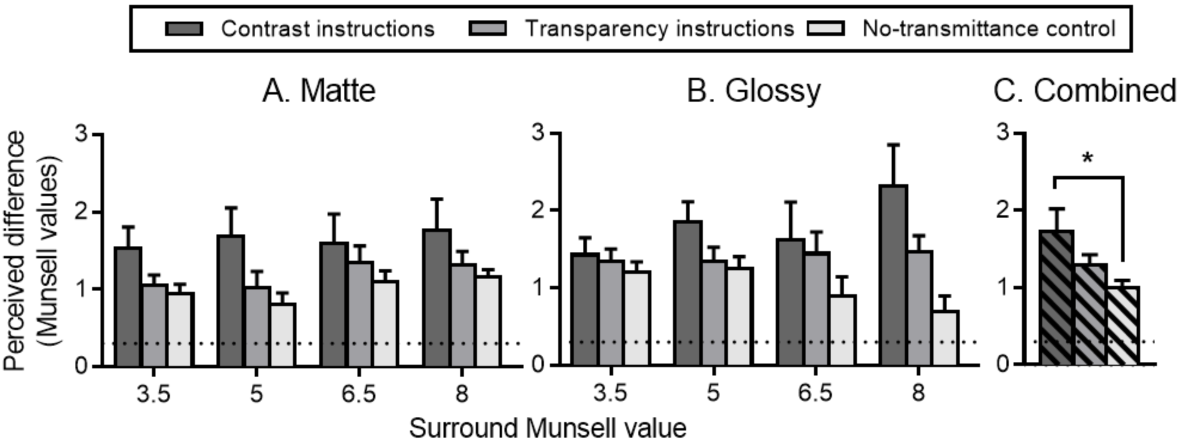


Figure 4.7. Increment minus decrement settings for the matte (A) and glossy (B) adjustable surfaces in Experiment 7. The horizontal dotted line represents the actual difference between increments and decrements. The solid bars represent the increment-decrement difference scores for each surround Munsell condition (3.5, 5, 6.5 and 8). The mean of these surround Munsell conditions was taken for each instruction condition and plotted as striped bars. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition.

In all conditions, the increment-decrement difference scores were greater than the actual (simulated) increment-decrement Munsell difference of 0.3 (horizontal dotted lines in Figure 4.7). The crispening effect appeared to be largest when observers were instructed to match the centre patches on lightness and apparent contrast against the background (Figure 4.7, dark-grey bars). In this condition, the immediate increments appeared, on average, about 1.7 Munsell values lighter than the immediate decrements. The crispening effect was smaller when observers were instructed to match the centre patches on lightness and transparency, or lightness alone (Figure 4.7, mid-grey and light-grey bars respectively). In the transparency instructions condition, the immediate increments appeared, on average, about 1.3 Munsell values lighter than the immediate decrements. In the no-transmittance control condition, the immediate increments appeared, on average, about 1 Munsell value lighter than the immediate decrements.

The increment-decrement difference scores were subjected to a between-subjects one-way ANOVA to determine whether there were any statistically reliable differences in the crispening effect between the different instruction conditions. Unlike in previous experiments, different observers were allocated to the conditions being compared (i.e., the different instruction conditions). For this reason, and since the pattern of results was identical for each gloss condition and each surround Munsell condition (see Figure 4.7), observers' difference scores were averaged across these conditions to reduce error (Figure 4.7C). The one-way ANOVA on these average difference scores revealed that there were differences between the conditions, $F(2,57) = 3.42$, $p = .0394$. Follow-up t-tests using Sidak-corrected alpha values of .0170 per test[13] indicated that the perceived difference between immediate increments and decrements was larger for contrast instructions compared to the no-transmittance control, $t(57) = 2.60$, $p = .0119$. However, the perceived difference between increments and decrements did not differ for the contrast instructions and the transparency instructions, $t(57) = 1.56$, $p = .124$, nor between the transparency instructions and no-transmittance conditions, $t(57) = 1.04$, $p = .303$. The above findings suggest that the crispening effect was largest when observers were instructed to match the central patches on lightness and apparent contrast against the background, at least compared to when observers were only allowed to match lightness. There was no clear difference in crispening between the contrast and transparency instruction conditions, even with the increased power gained by averaging each observer's score over all conditions.

---

[13] Sidak-corrected alpha values were calculated as $\alpha_1 = 1 - (1 - \alpha)^{1/n} = 1 - (1 - 0.05)^{1/3} = .0170$.

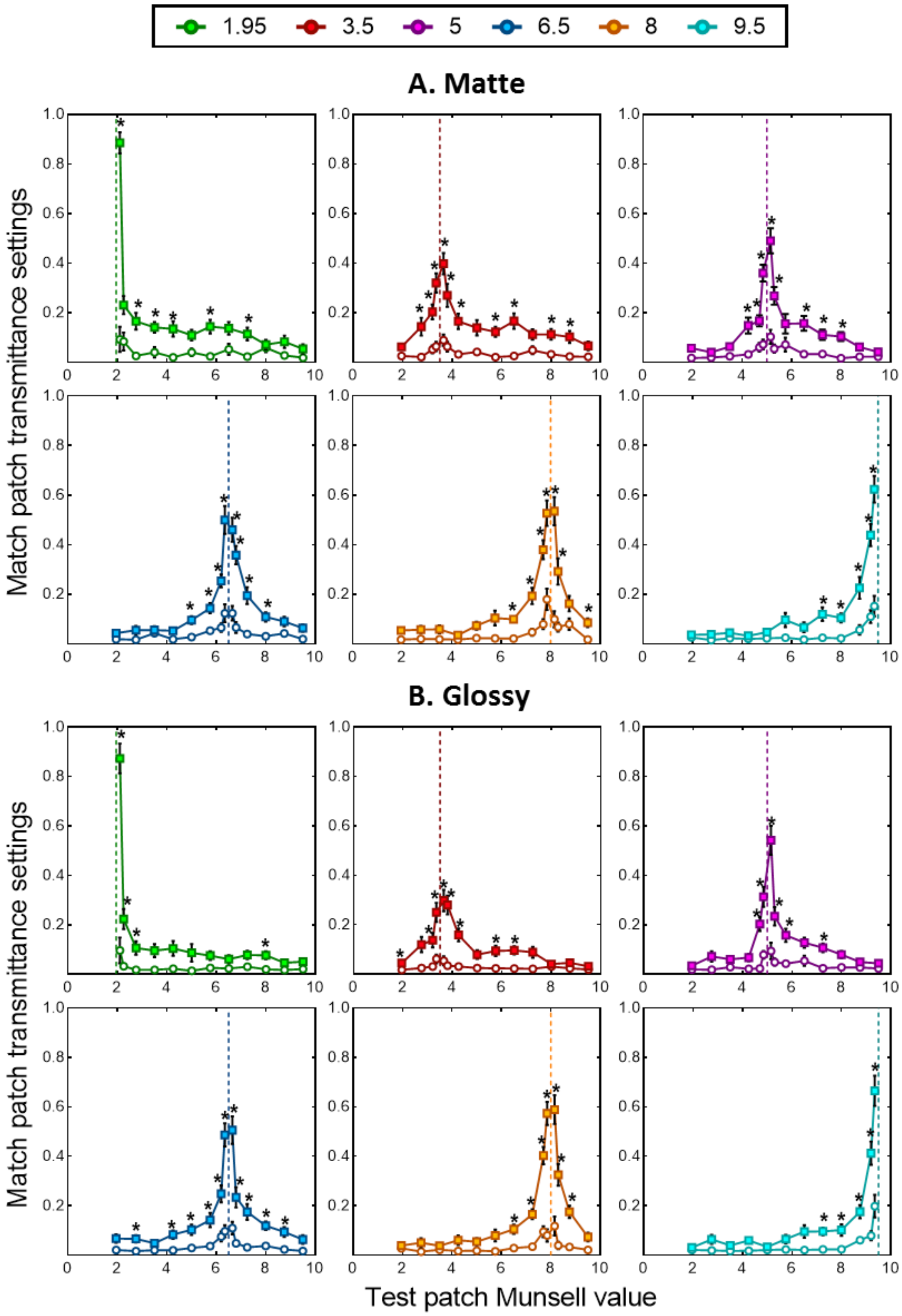


Figure 4.8. Transmittance settings from Experiment 7, for the matte (A) and glossy (B) condition. Closed square data points are settings from the contrast instructions condition, and open circle data points are settings from the transparency instructions condition. Each colour represents the settings for a different surround albedo condition. The vertical dotted lines indicate the surround Munsell value. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition.

However, there were differences between the contrast and transparency instruction conditions in the transmittance data. Figure 4.8 plots the transmittance settings for the contrast instructions and transparency instructions conditions. Each graph plots the transmittance data for each gloss condition of the matching display (Figure 4.8A and 4.8B for matte and glossy, respectively) and each surround Munsell condition. Within each graph, closed square data points represent transmittance settings for the contrast instructions condition, and open circular data points represent transmittance settings for the transparency instructions condition. While there were no reliable differences in crispening between these conditions, transmittance settings strongly depended on task instructions. Observers tended to vary the transmittance settings more when they were instructed to match the central patches on contrast against the surround compared to when they were asked to match the central patches on perceived transparency. In both conditions, transmittance settings increased as test patch reflectance became closer to the surround value (vertical dotted lines in Figure 4.8). However, this effect was much more pronounced in the contrast instructions condition, where transmittance peaked at about 0.6 for most surround Munsell conditions[14], compared to the transparency instructions condition, where transmittance peaked at about 0.1 for most surround Munsell conditions. The above observations were statistically verified by subjecting the transmittance settings to independent t-test using Sidak-corrected alpha values of .00366 per test[15] comparing transmittance settings in the contrast and transparency instructions conditions (see Tables 4.3 and 4.4 for *t* values, *df*, and *p* values). The asterisks in Figure 4.8 indicate a significant difference in transmittance settings between the two instruction conditions. As can be seen in these figures, most of the transmittance settings were significantly higher for the contrast (closed square data points) compared to the transparency (open circle data points) instructions condition (44 out of 80 for the matte match display; 41 out of 80 for

---

[14] The lowest contrast test patch against surround 1.95 was set close to 1 because this test patch was actually invisible against the surround, as mentioned in Experiment 1. Thus it would have been perceived as having zero contrast, leading observers to set the adjustable patch to be invisible.

[15] Sidak-corrected alpha values were calculated as $\alpha_1 = 1 - (1 - \alpha)^{1/n} = 1 - (1 - 0.05)^{1/14} =$ .00366.

the glossy match display). The transmittance settings that reliably differed are clustered around the test patches that were low in physical contrast against the surround. Thus, it appears that observers used transmittance as a proxy to match the central patches on perceived contrast, increasing transmittance to lower the contrast of the adjustable patch. Observers tended not to vary transmittance much when asked to match the central patches on perceived transparency, which might suggest that they did not directly perceive transparency in the homogeneous centre-surround displays. This is at odds with Ekroll and Faul's (2013) findings with coloured centre-surround displays, where observers given transparency instructions did vary the transmittance setting when it was available.

The results of Experiment 7 raise the question of why observers did not vary transmittance when asked to match transparency. There are a number of possible reasons. First, when transmittance was varied, the central adjustable patch became textured because the rocky background was visible through the filter. Perhaps observers did not like to match a textured patch to a uniform patch, so avoid using the transmittance variable altogether. Another possible explanation is that they did not explicitly perceive transparency in the homogeneous centre-surround display. The perceptual outcome may be reminiscent of transparency (low contrast test patches look insubstantial or "wispy"), but perhaps observers did not have direct access to this impression, or they could not quantify it using a matching task. The fact that transmittance was slightly varied in the transparency instructions condition suggests that observers may have perceived transparency, but were unable or hesitant to quantify this percept. Our own explorations of this revealed that, even for observers who reported perceiving the test patch as transparent, it was very difficult to judge how transparency should be matched in the adjustable patch. In contradistinction, matching the contrast of the edges of the two displays was much more intuitive and was a comparatively easy task. The ability to directly measure impressions of transparency in homogeneous centre-surround displays may be akin to judging the level of illumination in a room. Research has shown that observers can perceive differences in illumination between two rooms, but are bad at explicitly matching the illumination (Rutherford & Brainard, 2002). A possible reason for why observers varied transmittance when asked to match the contrast of the central patches is that varying the mid-level property of transparency acted as a proxy to allow observers to match a low-level property such as edge contrast, and this is something that observers could explicitly match.

The crispening effect did not reliably differ between the contrast and transparency instruction conditions. However, transmittance settings were vastly different between the conditions. Experiment 8 explored which settings observers found best matched the homogeneous centre-surround display.

	Surr 1.95		Surr 3.5		Surr 5		Surr 6.5		Surr 8		Surr 9.5	
Comparison	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value
1	11.93	**<.001***	2.75	.0091	3.04	.0043	2.29	.0280	2.94	.0056	0.94	.3529
2	2.86	.0069	3.38	**.0017***	2.34	.0246	2.21	.0334	2.57	.0143	2.55	.0151
3	4.04	**<.001***	4.49	**<.001***	2.98	.0050	0.63	.5293	2.42	.0206	1.62	.1130
4	3.36	**.0018***	5.95	**<.001***	3.52	**.0011***	2.66	.0115	2.51	.0164	1.31	.1969
5	3.68	**<.001***	6.34	**<.001***	3.73	**<.001***	3.79	**<.001***	2.97	.0052	2.55	.0149
6	2.58	.0140	4.08	**<.001***	7.31	**<.001***	3.96	**<.001***	2.70	.0102	2.44	.0193
7	4.22	**<.001***	4.16	**<.001***	6.65	**<.001***	6.01	**<.001***	5.50	**<.001***	2.42	.0204
8	2.45	.0189	2.88	.0066	5.37	**<0.001***	5.61	**<.001***	4.05	**<.001***	3.23	**.0025***
9	3.46	**.0014***	4.93	**<.001***	1.76	.0867	5.91	**<.001***	7.03	**<.001***	4.37	**<.001***
10	0.49	.6285	4.38	**<.001***	3.93	**<.001***	6.76	**<.001***	5.18	**<.001***	3.52	**.0011***
11	2.04	.0483	2.39	.0218	3.46	**.0014***	4.42	**<.001***	6.67	**<.001***	6.38	**<.001***
12	1.84	.0741	3.44	**.0014***	4.85	**<.001***	3.74	**<.001***	4.02	**<.001***	6.97	**<.001***
13			3.11	**.0035***	2.72	.0097	2.13	.0396	2.14	.0390		
14			2.31	.0263	1.67	.1033	2.90	.0062	3.76	**<.001***		

Table 4.3. Matte matching display: *t* values and *p* values comparing transmittance settings between the contrast and transparency instruction conditions, for each surround albedo and test patch albedo condition. The comparison numbers 1-14 correspond to the test patches being compared in each graph in Figure 4.8A from left to right. * $p < .00366$; $df = 38$ for all comparisons.

	Surr 1.95		Surr 3.5		Surr 5		Surr 6.5		Surr 8		Surr 9.5	
Comparison	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value	*t* value	*p* value
1	9.71	**<.001***	3.11	**.0036***	2.29	.0279	2.66	.0113	0.90	.3743	1.40	.1689
2	4.62	**<.001***	3.04	.0042	2.46	.0186	3.35	**.0019***	1.99	.0537	2.41	.0207
3	3.24	**.0025***	3.59	**<.001***	1.94	.0599	2.12	.0410	1.52	.1356	2.40	.0216
4	2.96	.0053	4.40	**<.001***	2.96	.0053	3.69	**<.001***	2.27	.0293	2.93	.0057
5	2.70	.0104	5.21	**<.001***	6.29	**<.001***	3.10	**.0036***	2.50	.0167	1.91	.0634
6	2.04	.0483	6.33	**<.001***	5.39	**<.001***	3.29	**.0021***	2.62	.0126	2.31	.0264
7	2.70	.0103	4.74	**<.001***	6.65	**<.001***	4.34	**<.001***	3.75	**<.001***	2.86	.0069
8	2.20	.0341	2.90	.0062	4.65	**<.001***	7.35	**<.001***	6.14	**<.001***	4.45	**<.001***
9	2.22	.0322	3.19	**.0028***	3.93	**<.001***	6.47	**<.001***	7.47	**<.001***	3.32	**.0020***
10	3.32	**.0020***	4.20	**<.001***	2.56	0.0144	4.39	**<.001***	9.30	**<.001***	3.83	**<.001***
11	2.19	.0347	2.93	.0058	4.42	**<.001***	4.35	**<.001***	6.75	**<.001***	6.58	**<.001***
12	2.28	.0285	0.76	.4547	3.07	0.0039	4.20	**<.001***	6.15	**<.001***	6.11	**<.001***
13			2.28	.0285	1.69	0.0989	3.42	**.0015***	5.47	**<.001***		
14			1.52	.1376	1.50	0.1415	2.58	0.0140	2.76	.0089		

Table 4.4. Glossy matching display: *t* values and *p* values comparing transmittance settings between the contrast and transparency instruction conditions, for each surround albedo and test patch albedo condition. The comparison numbers 1-14 correspond to the test patches being compared in each graph in Figure 4.8B from left to right. * $p < .00366$; $df = 38$ for all comparisons.

Experiment 8: Which Dimensions lead to Better Matches?

Experiment 8 investigates which adjustable patch settings in Experiment 7 were "best" matched to test patches embedded in homogeneous surrounds. Experiment 7 revealed that the crispening effect was similar between the instruction conditions. However, interestingly, when observers were able to vary transmittance, they tended to use this setting to match the contrast of the centre patches, but not when they matched the patches on transparency. This raised the question of which perceptual dimensions need to be varied to lead to better lightness matches. In Experiment 8, a new set of observers directly compared the settings made in Experiment 7 and judged which settings were considered to be better matched in lightness to the test patches in the homogeneous displays. We hypothesised that if satisfactory lightness matches can be made by adjusting the reflectance of the adjustable patch alone, then observers should choose the settings from the no-transmittance control or transparency instructions condition as being better matched in lightness to the homogeneous displays. If an additional dimension such as contrast or transparency is required to make satisfactory lightness matches, then observers should choose the settings from the contrast instructions condition as being better matched in lightness to the homogeneous displays. This is because observers only varied transmittance substantially in the contrast condition.

Methods

Observers

Twenty first-year psychology students participated in Experiment 8. None had participated in Experiment 7.

Apparatus and stimuli

The homogeneous centre-surround stimuli were the same as in Experiment 7. The comparison rocky displays were the same as the adjustable displays in Experiment 7, containing a rocky surface overlayed with a circular central disk that varied in simulated albedo and transmittance. On each trial, the albedo (t in Equation 4.1) and transmittance value (α in Equation 4.1) were the average albedo and transmittance settings made by observers in Experiment 7. The surrounds of the displays were cropped by 2.49° on each side so that four displays could fit on the computer monitor on each trial.

Procedure

In a two-alternate-forced-choice (2AFC) task, observers judged which instruction condition in Experiment 7 led to better lightness matches with the homogeneous display. In each trial, two pairs of displays (9.95° each) were presented on the computer screen. One pair was presented at the top of the screen, and one pair was presented on the bottom of the screen (see Figure 4.9). The surfaces within a pair were separated horizontally by 10.71° (centre to centre) and the pairs were separated vertically by 14.25° (centre to centre). Each pair consisted of a homogeneous centre-surround display and the corresponding rocky matching display overlayed by a filter with the average reflectance and transmittance settings from Experiment 7. The matching displays in each pair compared two of the three instruction conditions in Experiment 7. The homogeneous display was the same in the top and bottom pair. Observers were instructed to choose in which pair (top or bottom) the central patches were more similar in lightness. For the displays that had transparent central patches, observers were instructed to only pay attention to the pigment in the filter, and not the background that showed through. They made their decision by pressing the up or down arrow key for the top and bottom pair, respectively. After the stimuli were displayed, there was a five second delay before observers could make a response. Observers were told to use this time to carefully consider which pair was better matched, as the lightness of the central patches in each pair were extremely similar.

There were three comparison conditions: contrast condition compared to the transparency condition; contrast condition compared to the no-transmittance condition, and transparency condition compared to the no-transmittance condition. The position on the screen (top or bottom) that each instruction condition pair was presented was randomised. The position (left or right) of the homogeneous and rocky displays within each pair was also randomised. For each comparison condition there were two gloss levels (matte and glossy), and six surround Munsell conditions (1.95, 3.5, 5, 6.5, 8, and 9.5). Eight test patches that were the lowest in contrast against the homogeneous surround were chosen for each surround Munsell condition (four increments and four decrements), except for surround Munsell 1.95 and 9.5, which only had increment or decrement test patches, respectively. This led to a total of 240 trials for each participant.

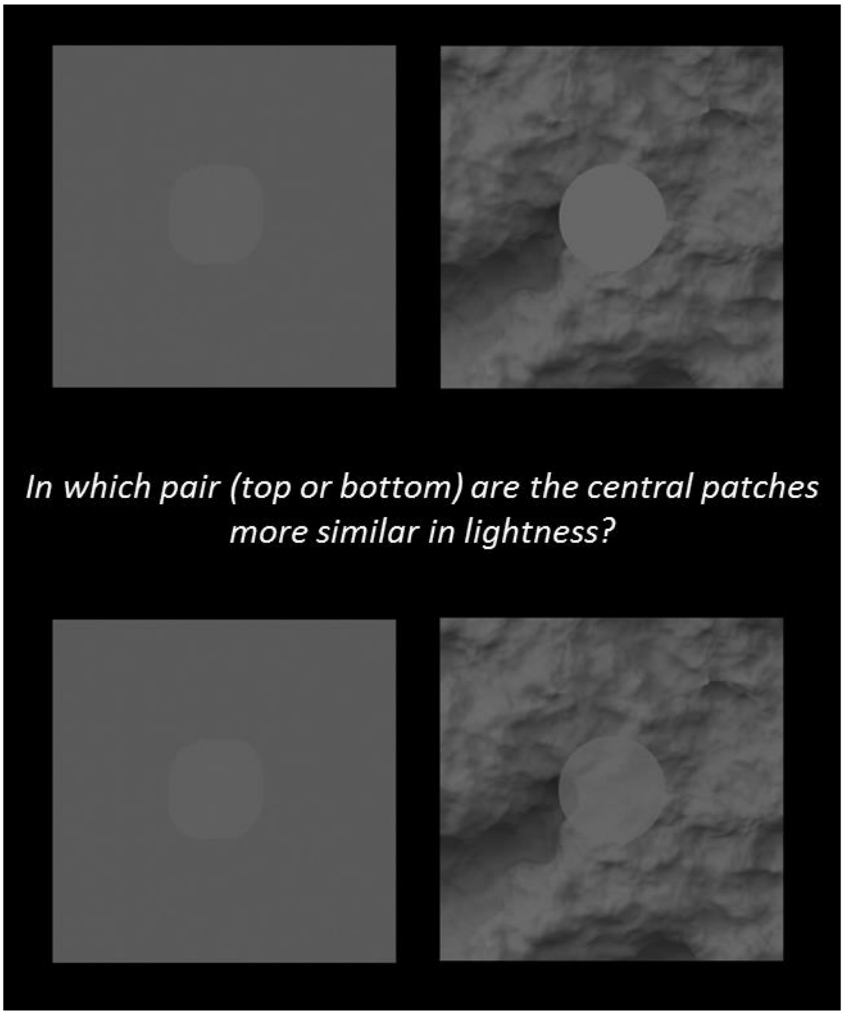


Figure 4.9. Representation of the layout of a trial in Experiment 8. The surfaces are cropped more in this display compared to in the Experiment, and the instructions text was not displayed on the screen; observers were given the instructions at the beginning of the experiment.

Results and discussion

The results of Experiment 8 are presented in Figure 4.10, which plots the number of times (out of 60) each instruction condition was chosen to have a match patch that was better matched in lightness to the homogeneous display. Each graph plots the preference for each test patch Munsell for a particular surround Munsell condition. Figure 4.10A shows the results for when the rocky displays were matte, and Figure 4.10B shows the results for when the rocky displays were glossy. The results showed that as the test patches became closer in albedo to the surround, the matching patches from the contrast instruction condition were chosen to be more similar in lightness to the homogeneous display compared to the other two conditions. Note that these were the displays with the highest transmittance settings. There was no clear preference for test patches further away from the surround albedo. Figure 4.11 shows the data averaged across test patch Munsell value, plotted as percentage of times chosen (because surround Munsell 1.95 and 9.5 had fewer data points to average across). It is clear from this figure that, on average, observers perceived stimuli that were set from the contrast instructions condition to be better matched in lightness to the homogeneous displays. To statistically verify this, the data for each surround Munsell condition were subjected to paired t-tests comparing the percentage of times the contrast instructions condition was chosen over the no transmittance control condition. We chose to compare these two conditions because we wanted to compare stimuli containing the most transparency (contrast instructions) to the stimuli containing no transparency (no transmittance control). The t-tests revealed that stimuli from the contrast instructions condition were chosen reliably more often in all matte surround albedo conditions, and the last five (out of six) glossy surround albedo conditions (see Table 4.5 for t values, df and p values).[16]

---

[16] Note that we could only validly compare two groups due to correlated measures (the percentage of times one group is chosen depends on the percentage of times the other two groups are chosen). If we compared the contrast instructions condition to the transmittance instructions conditions, the results are similar. For both matte and glossy conditions, stimuli from the contrast instructions condition were chosen reliably more often than stimuli from the transparency instructions condition, for the first five out of six surround albedo conditions, $p < .05$.

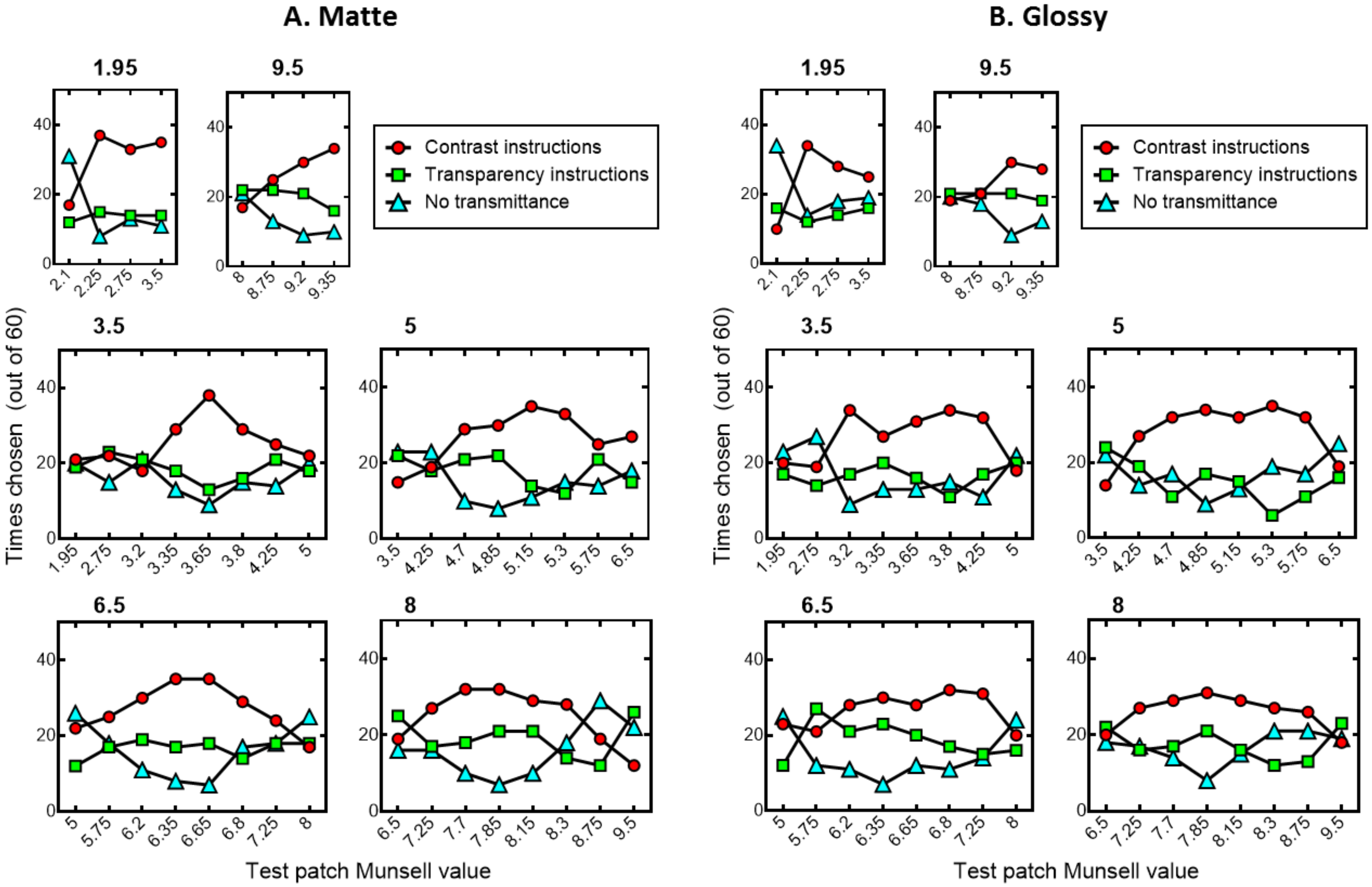

A. Matte
B. Glossy
1.95
9.5
3.5
5
6.5
8
Contrast instructions
Transparency instructions
No transmittance
Times chosen (out of 60)
Test patch Munsell value
0
20
40
2.1
2.25
2.75
3.5
8
8.75
9.2
9.35
1.95
2.75
3.2
3.35
3.65
3.8
4.25
5
3.5
4.25
4.7
4.85
5.15
5.3
5.75
6.5
5
5.75
6.2
6.35
6.65
6.8
7.25
8
6.5
7.25
7.7
7.85
8.15
8.3
8.75
9.5

Figure 4.10. Results of Experiment 8 for the matte condition (A) and the glossy condition (B). Each graph shows the results for a particular surround Munsell value, and plots the number of times (out of 60) each instruction condition was chosen to best match the lightness of the test patch embedded in the homogeneous display.

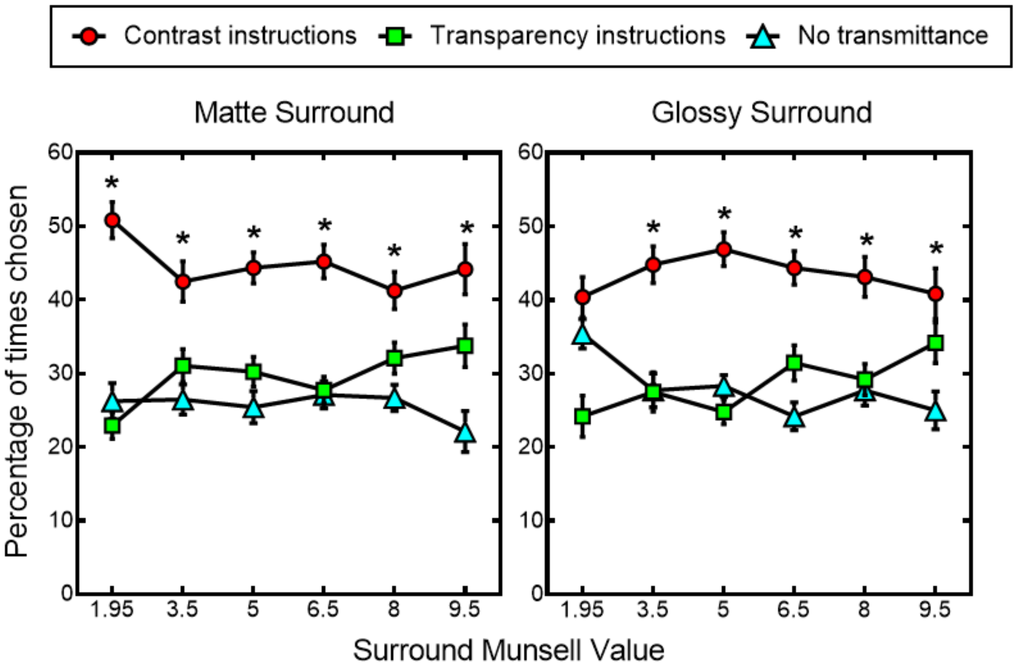


Figure 4.11. Results of Experiment 8, averaged across test patch Munsell. Each panel shows the percentage of times each instruction condition was chosen for each surround Munsell condition, for the matte condition (left) and the glossy condition (right). Error bars are standard error of the mean, and indicate inter-observer variability. Significance stars indicate when the contrast instructions condition was chosen reliably more than the no transmittance control.

The above results suggest that in order to match the lightness of a test patch embedded in a homogeneous surround to a patch embedded in a “textured” display, an extra perceptual dimension is required that allows the observer to match the patches on contrast against the surround. Giving observers the ability to vary transparency allowed them to better match edge contrast. Thus, to make more satisfactory lightness matches, observers seem to require the ability to adjust a mid-level dimension (transparency) to match a low-level construct (contrast) as well as lightness. It is also plausible that mid-level perceptual mechanisms caused the lightness shifts in low-contrast centre-surround displays without observers being explicitly aware of this. It is not possible to determine from the experiments so far whether mid-level or low-level mechanisms are dominant. However, we can conclude that observers can obtain more satisfactory matches to low-contrast homogeneous displays when they are allowed a second degree of freedom in their matches. Experiments 9 and 10 further investigate whether scission is responsible for the crispening effect.

	Matte		Glossy	
Surround Munsell condition	*t*	*p*	*t*	*p*
1.95	5.36	**<.001***	1.30	.209
3.5	3.71	**.0015***	4.27	**<.001***
5	5.00	**<.001***	5.25	**<.001***
6.5	4.86	**.00237***	5.81	**<.001***
8	3.80	**.0012***	3.55	**.0021***
9.5	4.00	**<.001***	2.91	**.0089***

Table 4.5. *t* values and *p* values comparing preferences in Experiment 8 for the contrast instructions condition over the no transmittance control condition, for each surround Munsell condition and gloss level. * *p* < .05; *df* = 19 for all comparisons.

Experiment 9A and 9B: Investigating the Increment-Decrement Asymmetry

A noteworthy aspect of the results from Experiment 1 was the asymmetry between increment and decrement lightness judgments for the homogeneous centre-surround displays. The data revealed that increment settings were essentially independent of the surround reflectance, whereas decrement settings were more affected by the surround reflectance (illustrated by a greater spread in the data points for each test patch; Figure 2.6, top two rows). In Experiment 9, we investigate whether this asymmetry persists when observers are able to vary both lightness and transmittance of an adjustable display that is common to all surround albedo conditions.[17] We were also interested to explore whether any asymmetries in lightness settings were accompanied by asymmetries in the transmittance settings. This experiment was similar to Experiment 7, except that lightness and transmittance could be adjusted in a matching surface that was common to all surround albedo conditions (this was not the case in Experiment 7). Experiment 9 had two parts. In Experiment 9A, observers varied both the albedo and transmittance of the adjustable patch and the relationship between albedo and transmittance was observed. Experiment 9B was a control experiment where participants could only vary the albedo of the adjustable patch (the transmittance setting was not available).

Methods

Observers

Four observers participated in Experiment 9. Observers RS, DC, SM and AS participated in Experiment 9A, and observers RS and DC participated in Experiment 9B.

---

[17] Note that in Experiment 7, it was not possible to directly compare surround albedo conditions because a separate matching display was used in each condition.

Apparatus and stimuli

The test stimuli were the same homogeneous centre-surround displays in Experiment 7. The matching display comprised of an adjustable filter overlaying a Brownian noise textured background (Figure 4.5B). The maximum luminance in the background was 32.49 cd/m^2, which corresponded to the luminance of the lightest test patch in the homogeneous displays (Munsell 9.5). Unlike Experiment 7, the same matching display was used for all trials. The luminance values in the central disk in the adjustable display were created in the same way as in Experiment 7, using Metelli's episcotister model (Equation 4.1).

Procedure

The procedure was the same as in Experiment 7. In each trial, a homogeneous centre-surround test surface (14.88°) was presented on the computer screen. Below the target surface was the smaller adjustable surface (5.83°). The surfaces were spatially separated by 11.47° (centre to centre) and were presented against a black background. In Experiment 9A, observers were given contrast instructions and were able to vary both the reflectance and the transmittance of the adjustable patch (see Experiment 7 methods section). In Experiment 9B, observers were only able to vary the reflectance of the test patch, equivalent to the no-transmittance control condition in Experiment 7. Similar to Experiment 7, there were six surround Munsell conditions, and 13-15 test patch Munsell conditions. This resulted in 86 conditions. Each observer performed five repeats of each condition, resulting in a total of 430 trials.

Results and discussion

The results of Experiment 9 are presented in Figures 4.12 and 4.13. Figure 4.12 shows the test patch lightness settings for Experiment 9A and Experiment 9B (Figure 4.12A and 4.12B, respectively). Figure 4.13 plots the transmittance settings for Experiment 9A where observers were given contrast instructions. The crispening effect was observed both when observers were given contrast instructions (Figure 4.12A) and when observers were not able to vary transmittance (Figure 4.12B). Figure 4.12 shows that the size of the steps is noticeably larger in the contrast instructions condition (Experiment 9A) compared to the no transmittance control (Experiment 9B). This is consistent with the results of Experiment 7. Additionally, in both Experiments 9A and 9B there was an asymmetry between increment and decrement settings (i.e. the increment settings are overall less spread out, or less dependent on surround albedo, than decrement settings, as in Experiment 1). The increment settings exhibited a higher amount of spread for lighter test patches compared to Experiment 1. However, the spread in the increments was similar for Experiment 9A and the control Experiment 9B, which was essentially a replication of Experiment 1 homogeneous condition. Furthermore, the present results replicated the findings in Experiment 1 that decrement settings were more dependent on the surround albedo compared to increments, which were relatively independent of surround albedo.

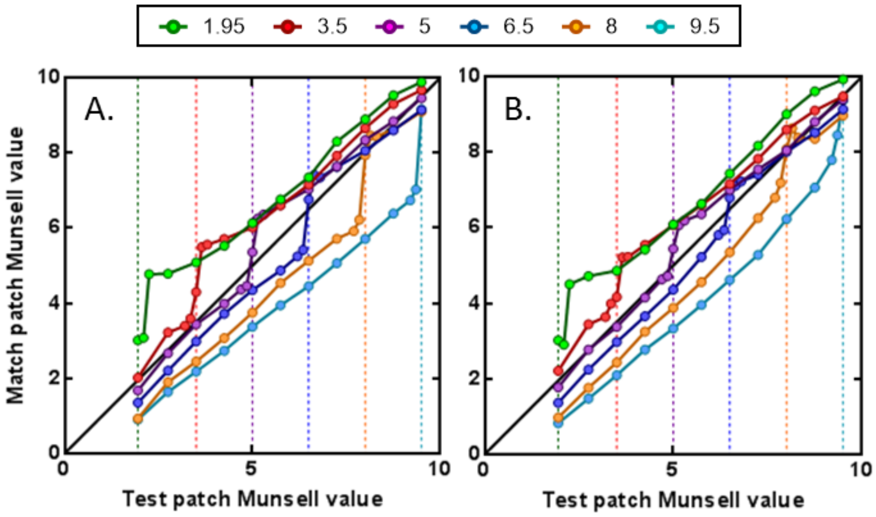


Figure 4.12. Test patch lightness settings for Experiment 9A (A) and Experiment 9B (B).

Figure 4.13 plots the transmittance settings for Experiment 9A, where observers were given contrast instructions. Each graph plots the transmittance data for each surround Munsell condition. Transmittance settings were revealed to be similar to the contrast instructions of Experiment 7; transmittance settings increased as test patch reflectance became closer to the surround value (vertical dotted lines in Figure 4.13). Transmittance peaked between 0.3 and 0.5, which was slightly less than in Experiment

7, but much higher than the transparency condition in Experiment 7. This difference between experiments might be attributed to individual differences and/or differences in the matching display used in each experiment. The matching displays in Experiment 7 contained rocky surrounds with the same albedo as the homogeneous displays (Figure 4.5A), whereas the matching displays in Experiment 9 contained a Brownian noise texture that was common to all surround-albedo conditions (Figure 4.5B). An interesting aspect of the transmittance settings in Figure 4.13 is that decrement settings appear to follow the same function for each surround, whereas the curve connecting the increment settings appears to get shallower as the surround Munsell value increases. This is better demonstrated in Figure 4.14A, where transmittance settings for each surround are plotted on top of one another. This was achieved by plotting the x-axis as test Munsell minus surround Munsell value. In this figure it is clear that the functions describing the increment settings are different whereas the functions describing the decrement settings are similar. This observation was supported statistically. The increment and decrement transmittance settings were log transformed, and linear functions were fit (via linear regression) for each surround Munsell condition (Figure 4.14B). For the transformed data (Figure 4.14B), the slope of the line corresponds to how much transmittance settings changed as the test patch became closer in albedo to the surround. The further away from zero the slope is, the greater the change in transmittance settings between test patches. A comparison of the slopes revealed that for increments, there was a significant difference between the slopes, $F(4, 29) = 3.25$, $p = .0256$ (see Table 4.6 for the slope and y-intercept for each line). However, for decrements there was no significant difference between the slopes, $F(4, 30) = 2.57$, $p = .0578$.[18]

---

[18] Since the slopes did not significantly differ for decrements, it was possible to calculate one slope for all the data. The pooled slope equalled 0.303. There was also no difference in the intercepts for decrements, $F(4,30) = 1.211$, $p = .324$. Thus there was a pooled y-intercept, which equalled 1.445.

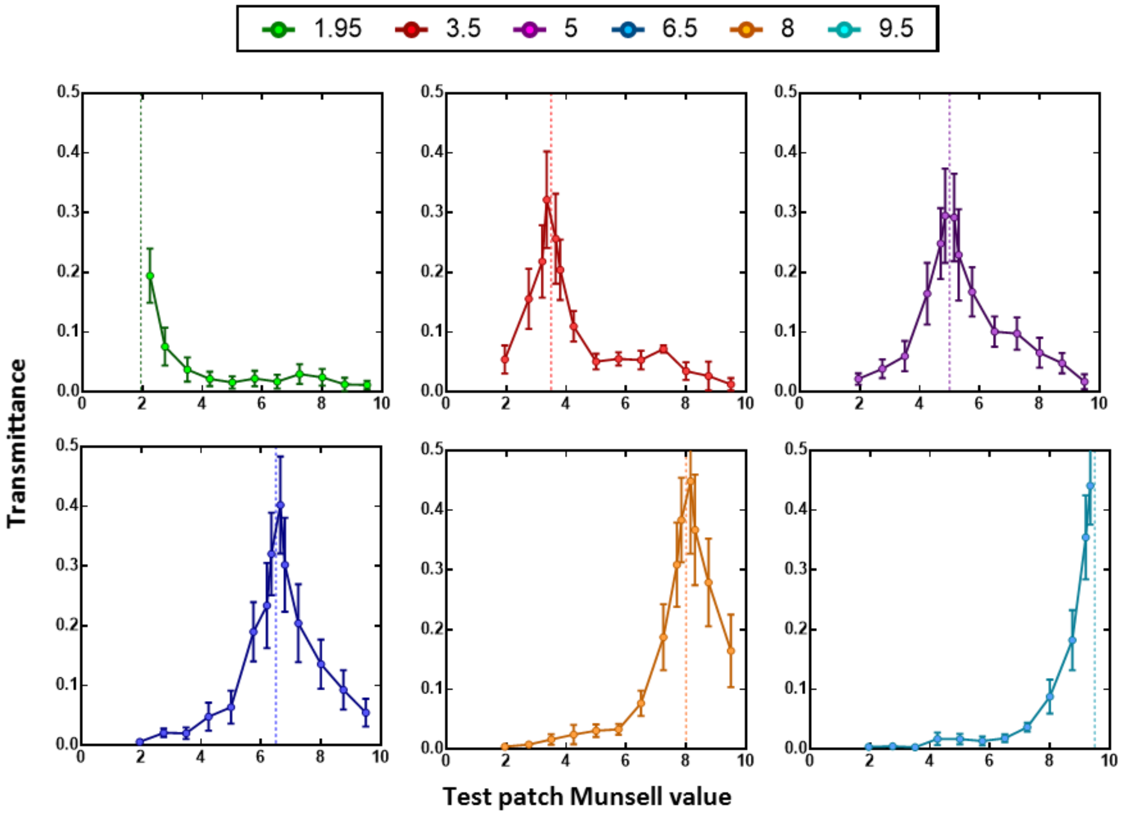


Figure 4.13. Transmittance settings for Experiment 9A. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition. Note that error bars are much larger in this experiment compared to Experiment 7 because there were only four observers in Experiment 9A, compared to 20 observers in Experiment 7.

The results displayed an asymmetry in increment and decrement lightness settings that was accompanied by the *opposite* increment-decrement asymmetry in the transmittance settings. In general, when a given test patch was a decrement, transmittance settings were similar, while lightness judgments decreased as surround albedo increased. Alternatively, when the test patch was an increment, lightness judgments were similar, while transmittance settings increased as surrounds became lighter.

The above results suggest that if transparency (whether directly perceived or not) plays a role in perceived lightness of test patches embedded in homogeneous surrounds, then there is a relationship between the perceived lightness of a test patch on various surrounds and how much a target's perceived transparency varies on different surrounds. If scission (whether directly perceived as transparency or not) affects the perceived lightness of test patches embedded in homogeneous surrounds, then the following explanation could account for the results: In a given display, the luminance within the test patch region is decomposed into a foreground filter layer, and a background layer that is a continuation of the surround (i.e. the same lightness as the surround). For decrement displays, induction from different surround albedos causes test patches to appear lighter as the surround becomes darker. This induction is attributed to changes in test patch lightness, while transparency remains relatively constant. Alternatively, increment test patches remain similar in lightness independent of the surround. Thus, the different levels of induction from different surround albedos is attributed to changes in test patch transparency, while lightness remains relatively constant.

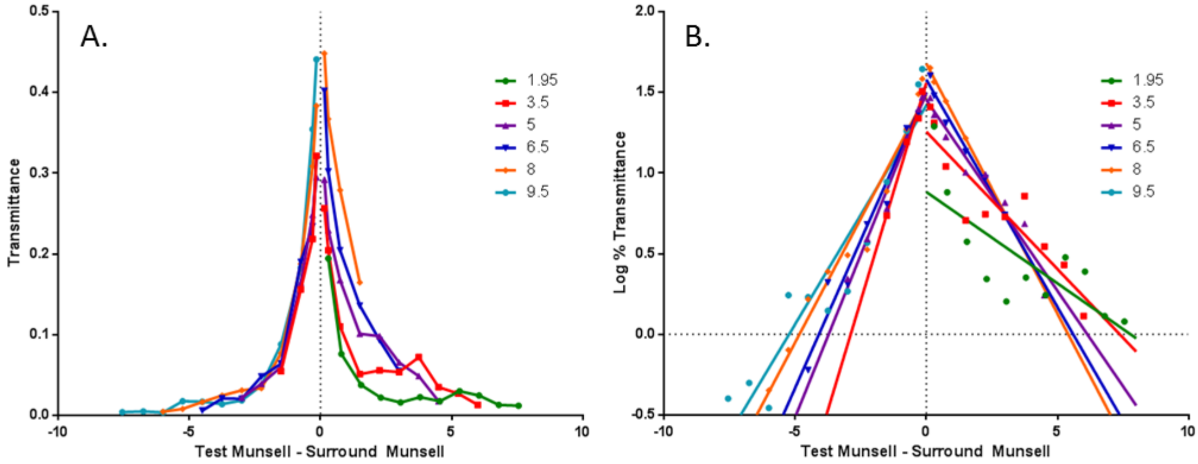


Figure 4.14. Transformed transmittance data from Experiment 9A. (A) Transmittance settings for each surround are plotted on top of one another. This was achieved by plotting the *x*-axis as test Munsell minus surround Munsell value. (B) The increment and decrement transmittance settings were log transformed, and linear functions were fit via linear regression for each surround Munsell condition. See main body text and Table 4.6 for slopes and *y*-intercepts for each equation.

	Surround Munsell value				
	1.95	3.5	5	6.5	8
Slope	-0.113	-0.17	-0.237	-0.282	-0.3104
Y-intercept	0.883	1.257	1.453	1.58	1.679

Table 4.6. The slope and y-intercept for each function fitted to increments in Figure 4.14.

It is less clear how to interpret the relationship between lightness and transmittance settings if scission is not involved. However, the results provide insight into the increment settings, regardless of whether use of the transmittance setting taps into transparency or contrast mechanisms. Although the appearance of a test patch as an increment is relatively independent of surround albedo (i.e. it looks the same lightness regardless of surround), it still does not look identical. There are perceptual differences that cannot be captured by differences in lightness alone, but may be captured by a second dimension such as transparency or contrast. At this stage we remain agnostic as to whether this dimension is transparency or contrast. In the next experiment we matched homogeneous and variegated surfaces on test patch-surround contrast. We observed how crispening was affected by these low-contrast variegated surfaces, which were designed to disrupt the conditions for transparency.

Experiment 10: Low Contrast Surrounds

The previous experiments in Chapter 4 investigated whether transparency is involved in the perception of the homogeneous centre-surround displays, or whether low-level mechanisms involving texture consistency (Experiment 6) or contrast between the centre and surround (Experiments 7, 8, and 9) modulate the appearance of these displays. The present experiment investigates the possibility that the crispening effect is caused by enhanced discrimination of test patches that are low in contrast against the homogeneous surround. In all previous experiments, the test patches in the homogeneous displays were full increments or full decrements against the surround, i.e. luminance values within the test patch were either all lighter or all darker than the surround. Alternatively, test patch luminance values in rocky displays were often between the lightest and darkest pixels in the surround. In Experiment 10, variegated displays with low contrast surrounds were created so that all test patches were either full increments or decrements against the surround, like in the homogeneous displays. We hypothesised that if transparency causes the crispening effect, then crispening should be eliminated when the conditions for transparency are disrupted (in the variegated condition). In contradistinction, if the crispening effect is caused by low-level mechanisms that enhance discrimination of test patches that are low in contrast against the surround, then crispening should be similar between the two conditions.

Methods

Observers

Eight observers participated in Experiment 10. Four were undergraduate psychology students, and the other four were observers SM, KT, AS, and RS.

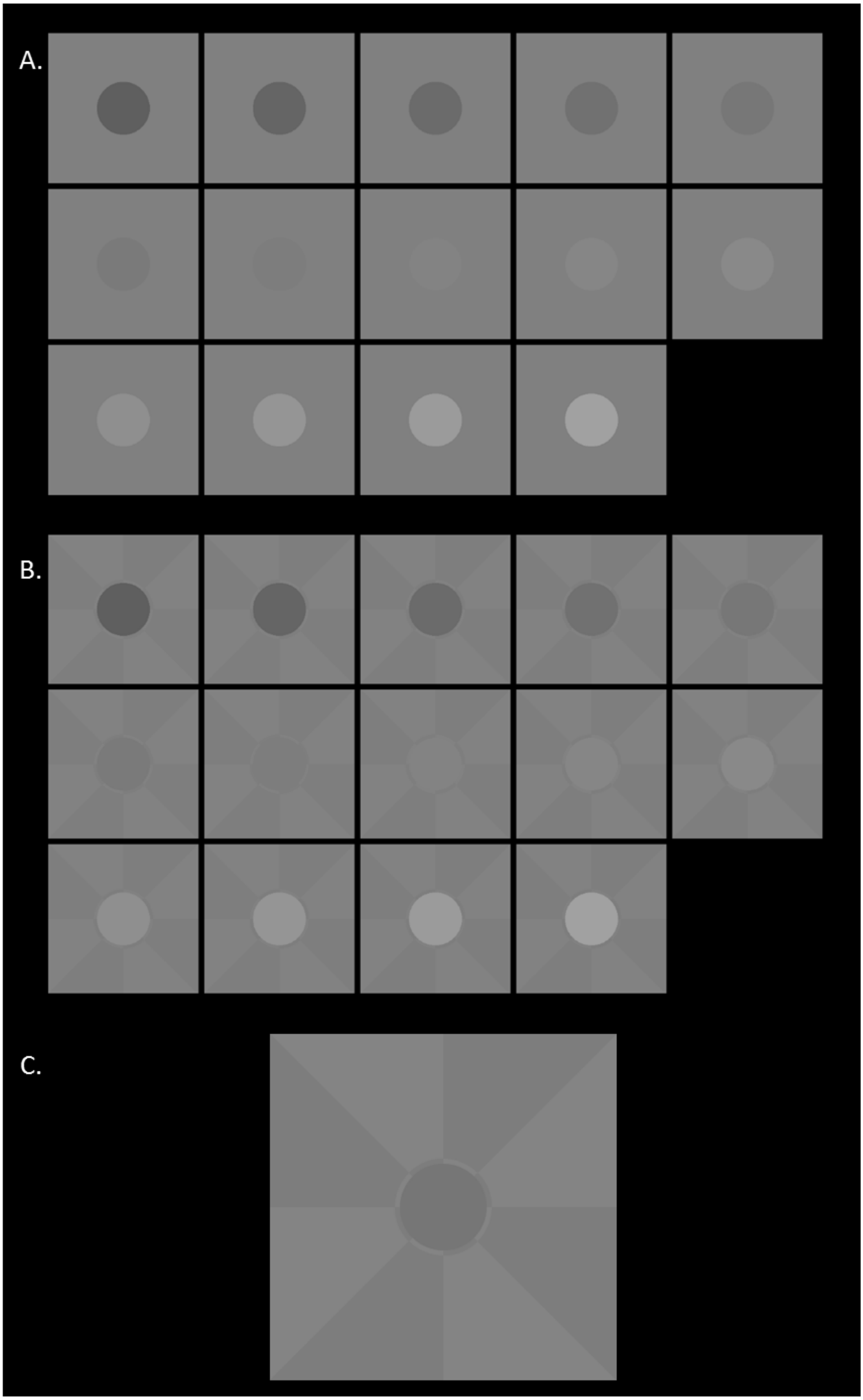

A.
B.
C.

Figure 4.15. Stimuli used in Experiment 10. (A) Homogeneous centre-surround displays. (B) Variegated displays. (C) Close up of a variegated display. For illustration purposes, the contrast in the surround in (C) is enhanced to show the pattern of the wedges and ring.

Apparatus and stimuli

The test stimuli were centre-surround displays and are presented in Figure 4.15. They were created in Matlab and did not have physical (simulated) reflectances associated with them. The homogeneous centre-surround displays were similar to those used in previous experiments. The homogeneous surrounds were mid-grey (16.16 cd/m^2). The variegated surrounds consisted of wedges that were slightly lighter than mid-grey (light wedges, 16.21 cd/m^2), and wedges that were slightly darker than mid-grey (dark wedges, 15.65 cd/m^2) surrounding a central circle. In addition to the wedges, a ring was placed around the central patch, which alternated from light to dark (the same luminance as the light and dark wedges, respectively). This ring was added to eliminate illusory filling of the surround wedges into the test patch region that occurred due to the extremely low contrast between the centre and the surround. The matching display was the same display used in Experiment 9 (Figure 4.5B). However, only lightness could be varied in the central patch, not transmittance.

Procedure

The procedure was the same as in previous lightness studies where observers could only vary the lightness of the adjustable patch. In each trial, a centre-surround test surface (8.68°) was presented on the computer screen. Below the target surface was the adjustable surface (5.83°). The surfaces were separated by 11.47° of visual angle (centre to centre) and were presented against a black background.

There were two surround-type conditions (homogeneous and variegated), and 14 test patch luminance conditions (8.36, 9.56, 10.90, 12.30, 13.71, 14.61, 15.38, 16.98, 17.81, 18.65, 20.45, 22.31, 24.2, 26.12 cd/m^2). This led to a total of 28 conditions. Each observer performed five repeats on each condition, leading to 140 trials for each observer.

Results and discussion

The results of Experiment 10 are presented in Figure 4.16, which shows the test patch luminance settings for the homogeneous (left) and variegated (right) conditions. The crispening effect was clearly present in the homogeneous condition (Figure 4.16, left). There was a “step” in lightness as the test patch luminance passed through the surround. There was also evidence of crispening in the variegated condition (Figure 4.16, right), but the size of the step was smaller than in the homogeneous condition. To further investigate crispening in the two conditions, the settings of the immediate increment and decrements were plotted in Figure 4.17. The red and blue dotted lines indicate the actual luminance of the immediate increment and decrement, respectively. The luminance settings showed that for both conditions, the perceived difference between increment and decrements was greater than the actual difference. This difference was larger, however, in the homogeneous centre-surround display. The above observations were statistically verified by subjecting the immediate increment and decrement luminance settings to a two-factor ANOVA, with display type (homogeneous, variegated) as one factor, and test patch (increment, decrement) as the other factor. The ANOVA revealed that there was a main effect of test patch, $F(1,7) = 28.94$, $p = .001$. Averaging across display type, increments were perceived to be lighter than decrements. This result was entirely expected, as the increments were physically lighter than decrements. There was no main effect of display type, $F(1,7) = 3.97$, $p = .0864$. Averaging across test patch, there was no difference in luminance settings between the homogeneous and variegated display conditions. There was a significant interaction between display type and test patch, $F(1,7) = 11.27$, $p = .0121$, indicating that the difference between increment and decrement settings was larger for the homogeneous compared to the variegated surround conditions, i.e. crispening was larger in the homogeneous condition. Follow-up t-test indicated that the greater crispening was caused by both the decrement appearing darker in the homogeneous compared to the variegated condition, $t(7) = 2.64$, $p = .0333$, and the increment appearing lighter in the homogeneous compared to the variegated condition, $t(7) = 3.59$, $p = .00890$.

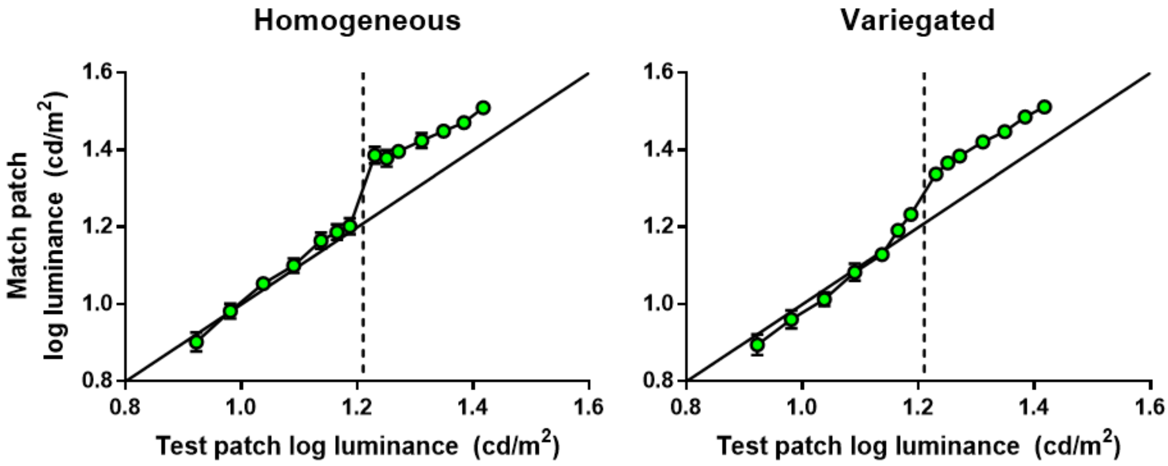


Figure 4.16. Results of Experiment 10. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition. In a number of conditions, error bars are smaller than the data points, so are not visible.

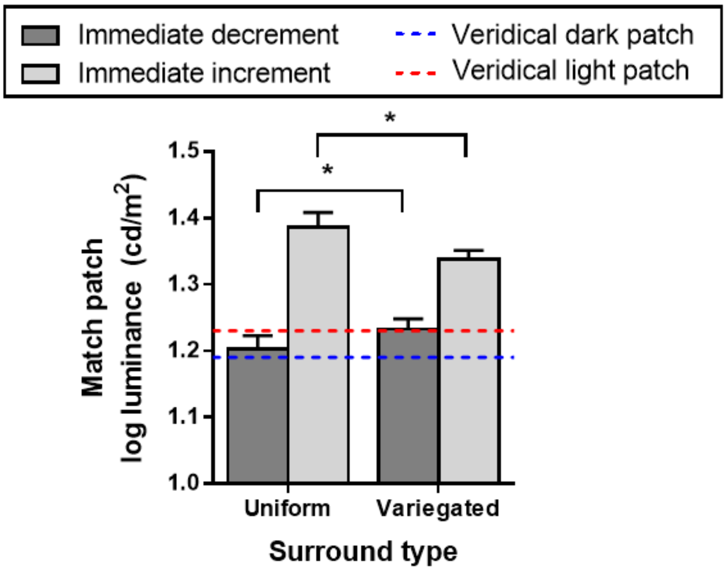


Figure 4.17. Immediate increment and decrement settings for each display type condition. The red and blue dotted lines indicate the actual luminance of the immediate increment and decrement, respectively. Error bars are standard error of the mean, and represent the inter-observer variability for a particular condition.

The above results suggest that scission may play a causal role in lightness perception in simple, homogeneous centre-surround displays. However, there was still a small amount of crispening present in the variegated displays, suggesting another reason for the different results between the two conditions. Although test patches were full increments and decrements relative to both the homogeneous and variegated surrounds, there was still higher contrast in the surround of the variegated displays compared to the homogeneous displays. This contrast in the surround may have been masking contrast effects between the surround and the centre patch, reducing the crispening effect. This may be similar to the masking artefact found in Experiment 6.

Taken together, the results from the experiments in Chapter 4 suggest that a second perceptual dimension in addition to lightness is required to capture the appearance of low contrast test patches embedded in homogeneous surrounds. When observers matched test patch lightness in a homogeneous display to that in a "textured" or rocky display, they required the ability to vary both the lightness and transmittance of the patch on the textured display to obtain the most satisfactory matches (Experiments 7 and 8). However, observers only chose to vary transmittance to match the contrast of the patches against the surround, and not to explicitly match the amount of transparency in the displays. This, along with the results of Experiment 10, suggests that homogeneous displays differ from textured displays in perceived edge contrast between the central patch and the surround, and that the only way to match the unique appearance of the edge contrast in the homogeneous displays is to vary the mid-level property of transparency in the textured displays. Chapter 5 will discuss this further in relation to existing literature.

Chapter 5: General Discussion

Summary of Experimental Results

Chapters 2 and 3 investigated whether the visual system uses cues to the illuminant created by complex mesostructure and specular highlights to improve lightness constancy (i.e. the consistency of lightness judgments) of an embedded flat, matte test patch. Lightness constancy was examined across changes in background albedo (Experiments 1 and 5) and illumination level (Experiment 5). In Experiment 1, observers judged the lightness of test patches surrounded by four different surfaces: flat matte, flat glossy, rocky matte, and rocky glossy. When the surrounds were rocky, there was a smooth, monotonic relationship between perceived lightness and test patch reflectance (replicated in Experiment 6), with glossy surfaces yielding better lightness constancy across changes in background albedo than matte surfaces (replicated in Experiment 5).

Control experiments and conditions tested whether any benefits of shading and specular reflections could be attributed to the range and/or number of luminance values in the surround (Experiment 2), the choice in matching patch surround (Experiment 3), differences in the energy across different spatial scales (Experiments 4 and 5), or the number of interreflections rendered in the scene (Experiment 5). When the original rocky displays were compared to variegated control displays with matched pixel histograms (Experiment 2), lightness constancy was better for the rocky glossy compared to the variegated glossy equivalent displays. However, this difference in lightness constancy was eliminated when the control displays had equated power spectra to the rocky displays via phase scrambling (Experiment 4, replicated in Experiment 5). In Experiment 5, observers exhibited some degree of lightness constancy when illumination varied, as lightness judgments were between reflectance and luminance matching. However, lightness constancy did not differ between glossy and matte conditions when illumination varied. Test patch lightness was also unaffected by the number of interreflections. These results were similar for the rocky and control (phase-scrambled) displays.

When surrounds were flat, there was a nonlinear "step" or crispening (Takasaki, 1966) in the pattern of matches as test patch reflectance passed through that of the surround (Experiment 1, replicated in Experiments 6, 7, 9, and 10). Furthermore, there was an asymmetry between increment and decrement settings (Experiment 1, replicated in Experiment 9). Lightness settings for decrements tended to decrease as surround albedo increased, whereas the induction for increments was independent of surround albedo.

Chapter 4 investigated whether the pattern of results from the homogeneous centre-surround displays was caused by perceptual transparency, or whether other mechanisms were responsible, such as differences in centre-surround contrast between the test and matching displays (Experiments 7, 8, and 9), or an enhanced discrimination of full increment and decrement test patches that are low in contrast against the surround (Experiment 10). In Experiment 6, observers judged the lightness of test patches that had consistent or inconsistent textures with their surround. When the rocky surround texture continued into the central test patch, test patches at the increment-decrement transition were judged to be similar in lightness (flattening of the data curve), and a "step" in lightness occurred when the increments became visible against the surround. In Experiment 7, observers matched test patches embedded in homogeneous surrounds by varying the lightness and transmittance of an adjustable patch overlaying a rocky matching display. When instructed to match the centres on lightness and transparency, observers tended not to use the transmittance settings. However, when instructed to match the centres on lightness and perceived contrast against the surround, observers utilised the transmittance settings, making the adjustable patch more transparent with test patches that were lower in contrast against the surround. Furthermore, observers judged low contrast homogeneous displays to be better matched in lightness to the matching displays whose centres were more transparent, i.e. matches were better when settings were made by observers who were given contrast instructions (Experiment 8). Experiment 9 demonstrated that the asymmetry in lightness settings between increments and decrements was accompanied by the opposite asymmetry in transmittance settings. Lightness judgments for increments were relatively independent of surround albedo, but transmittance settings varied depending on the surround. Alternatively, lightness judgments for decrements depended on surround albedo, but transmittance settings were relatively independent of the surround. Experiment 10 showed that crispening could be induced in variegated displays with extremely low contrast surrounds, where test patches were complete increments or decrements against the surround. However, the crispening effect in these displays was smaller compared to traditional homogeneous displays.

Relation to Previous Work

Lightness computations in complex displays

Contrary to the hypothesis that originally motivated our experiments, we were unable to find support that the visual system computes surface lightness by forming a representation of the light field, such as in EIMs (Bloj et al., 2004; Boyaci et al., 2003, 2004; for a review, see Brainard & Maloney, 2011). Our results suggest that, when computing test patch lightness, there is no benefit in observers' lightness judgments for stimuli that contain surface relief, shadows, shading, and specular highlights. Rather, a low-level explanation involving contrast and luminance distributions across space and scale appear to be sufficient in explaining the pattern of results described above.

One theory of lightness perception that does not require an explicit representation of the illuminant is anchoring theory (Gilchrist, 2006; Gilchrist et al., 1999). Anchoring theory proposes that the visual system assigns a fixed lightness value (white) to the highest luminance in a scene (or framework), which serves as a lightness anchor. Other values in lightness are computed relative to this anchor point by forming ratios relative to this anchor point. It is unclear how the principles of anchoring theory, such as the highest luminance rule, can be extended to naturalistic displays like those used in the present studies. Evidence supporting the anchoring rule comes from experiments that use simple displays such as dome experiments where the observers' whole field of view consists of two luminance values divided by a simple edge (Li & Gilchrist, 1999), and Mondrian experiments where observers are presented with a flat 2D array of different reflectances illuminated by a hidden spotlight (Cataliotti & Gilchrist, 1995). However, the highest luminance in most natural scenes will not be generated by the surface with the highest diffuse reflectance, but rather by specular reflections of the primary light source. Moreover, it is unclear how anchoring theory could account for the difference in lightness constancy between the variegated displays in Experiment 2 and the phase-scrambled displays in Experiment 4, which had equated luminance histograms and therefore the same highest luminance. Anchoring theory asserts that the distance from the anchor does not affect the lightness of a surface (Cataliotti & Gilchrist, 1995; Gilchrist, 2006; Radonjić & Gilchrist, 2013).

The anchoring to white rule is also inconsistent with findings in the literature from experiments that use more natural displays. Anderson et al. (2014) suggested that lightness and brightness were maximally conflated in the displays used to support anchoring theory. They demonstrated that observers did not always perceive the highest luminance to be white under more natural viewing conditions, where they were immersed in the same illumination field as the Mondrian display, or when a broad range of illumination conditions was simulated for Mondrians presented on a monitor. Additionally, a number of studies have shown that displays containing a single uniform albedo value with luminance variations caused by 3D structure do not always appear white, as would be predicted by anchoring theory (Gilchrist & Jacobsen, 1984; Motoyoshi et al., 2007; Ruppertsberg & Bloj, 2007; Sharan et al., 2008). Thus, there are many conditions where the highest luminance does not appear to serve as an anchor point when mapping luminance to lightness values.

The principle of frameworks is also difficult to apply to natural scenes. Gilchrist defines frameworks as “a group of patches in the retinal images that are segregated, or a group of patches that belong together or are grouped together” (Gilchrist, 2006, p.297). The boundaries of frameworks are depth or shadow edges, so it is unclear how frameworks should be applied to scenes with continuous luminance variation produced by shading of 3D objects and surfaces, like the rocky surfaces used in the present experiments. Furthermore, the above definition does not clearly specify how to apply the rules of frameworks to an arbitrary scene. Fundamentally, for a given image it is often not clear *a priori* what would constitute a framework. Overall it is unclear how the principles of anchoring theory can be extended to complex displays such as those in the present thesis.

One important consideration when interpreting the data from the rocky displays in Chapter 1 is that all surfaces were rendered under the same illuminant.

Consequently, test patch reflectance co-varied with luminance, meaning that observers could have performed the task by matching perceived luminance (brightness), not lightness. Thus, models that predict brightness could potentially account for the results. One such class of model are spatial filtering models, such as Blakeslee and McCourt's ODOG model (Blakeslee & McCourt, 1999, 2001, 2004; Blakeslee et al., 2005; see Shapiro & Lu, 2011, for an alternative model). Such models have the potential for dealing with more natural stimuli than anchoring theory. However, as they stand they are unable to differentiate various causes of image structure, such as the shading and specular reflections in the displays used in the present experiments.

Another potential problem with applying spatial filtering models to the present results is that they do not predict effects of illumination and transparency on brightness/lightness perception. Lightness and brightness can differ substantially when illumination is inhomogeneous (Blakeslee & McCourt, 2012). Indeed, for surfaces in Chapter 3, luminance differed substantially between test patches with the same reflectance under "bright" and "dim" illumination. Observers' lightness judgments were between reflectance and luminance matches, suggesting that observers did not merely match test patches on brightness, i.e. they were at least somewhat matching lightness. This is supported by similar findings in the literature (e.g. Madigan & Brainard, 2014; Ripamonti et al., 2004). A number of lightness illusions in the literature suggest that perceptual decomposition into layers (or scission; see Anderson, 1997; Anderson & Winawer, 2005, 2008) may affect lightness perception under some circumstances. Examples include the shadow simultaneous contrast (SC) illusion (Williams, McCoy & Purves, 1998), the snake illusion (Adelson, 2000), the checker shadow illusion (Adelson, 1995), the paint/shadow illusion (Hillis & Brainard, 2007), the argyle illusion (Adelson, 1993), and the wall of blocks illusion (Adelson, 1993; Logvinenko, 1999; Logvinenko & Ross, 2005). In a recent study, Blakeslee and McCourt (2012) applied their ODOG model to predict brightness perception in these illusions. Their model predicted the approximate magnitude of brightness effects in some of these displays (shadow SC, snake, checker shadow, and the paint/shadow illusions), but not others (argyle and wall of blocks illusions). Nor did the model predict the large shifts in perceived lightness for displays with illumination/transparency boundaries. Future studies should investigate whether low-level accounts involving distributions of contrasts can extend to lightness effects in naturalistic displays with illumination changes (spatial or temporal) and/or gradients, or whether mid-level (layered image decomposition) mechanisms may be involved.

Lightness computations in simple displays

The results from the homogeneous displays are in agreement with previous studies that have reported heightened discrimination for test patches that are close to the surround colour (Krauskopf & Gegenfurtner, 1992) or luminance (Whittle, 1992). Previous research has also reported asymmetries in luminance and colour judgments between increments and decrements (Bressan, 2006; Helson, 1938; Helson & Michels, 1948; Bäuml, 2001, Heinemann, 1955). The experiments in Chapter 4 showed mixed results for whether the lightness effects in the homogeneous centre-surround displays were caused by mid-level mechanisms such as layered image decomposition, as suggested by Ekroll and Faul (2013), or whether low-level mechanisms are sufficient to explain the effects. For example, observers judged low contrast homogeneous displays to be more similar in lightness to rocky displays that contained transparent versus opaque test patches (Experiment 8), suggesting that the homogeneous displays were perceived to be transparent. Furthermore, the increment-decrement asymmetry in lightness was accompanied by a complementary asymmetry in transmittance matches, which is consistent with the idea that luminance variations in the increment displays were attributed to changes in test patch transparency (lightness remained relatively constant), and luminance variations in the decrement displays were attributed to changes in test patch lightness (transparency remained relatively constant). Finally, the crispening effect was larger in homogeneous displays versus low contrast variegated displays, where the conditions for transparency were disrupted but test patches remained full increments and decrements relative to the surround (Experiment 10). These findings support the idea that transparency plays a role in the appearance of homogeneous centre-surround displays.

Though some of the results support a transparency explanation, it is not possible to rule out low-level contributions to the crispening effect. As mentioned above, crispening was larger for homogeneous versus low contrast variegated displays in Experiment 10. However, the variegated stimuli still induced some degree of crispening, even though transparency was disrupted. Whittle (1992) suggested that neural signals are enhanced for low contrast centre-surround displays, leading to enhanced discrimination of test patches close in reflectance to their surround. The reduced crispening in the low contrast variegated displays compared to the homogeneous displays could have been caused by contrast in the surround masking contrast effects between the central patch and the surround. Another result that supports a contrast explanation of crispening is that observers did not substantially make use of the transmittance setting in Experiment 7 when asked to match the transparency of the test patch in the homogeneous displays. This finding is at odds with Ekroll and Faul's (2013) findings with coloured stimuli, where observers adjusted transmittance when matching homogeneous and variegated displays on colour and transparency. It is possible that coloured centre-surround displays such as those used in Ekroll and Faul's study evoke a more compelling percept of transparency, compared to achromatic displays. However, observers in Experiment 7 did vary transmittance when instructed to match the patches on contrast against the surround, and preferred these settings over settings that did not make the matching patch transparent (Experiment 8). This suggests that the efficacy of using transparent match displays may have been a result of transparency providing a natural means for equating both the contrast of the edges and the lightness of the target. This idea will be discussed later.

Low-level explanations of lightness effects in simple displays

The results discussed above demonstrate that it is difficult to tease apart low- and mid-level contributions to lightness phenomena in simple displays. Local contrast explanations of lightness perception in centre-surround displays have been challenged by findings in the literature (Gilchrist, 2006). A number of experiments have shown that simple centre-surround stimuli containing identical luminance values (i.e. the same centre-surround edge ratio) can produce very different effects on perceived lightness depending on the perceptual interpretation of the scene (Arend et al., 1971; Evans, 1948; Gelb, 1932; Gilchrist, 1988; Gilchrist et al., 1983; Hsia, 1943; Jaensch & Müller, 1920; Landauer & Rodger, 1964; Oyama, 1968). In these experiments, stimuli are similar to the simultaneous contrast display in Figure 1.1 with a target at the centre of each of two adjacent backgrounds. A shadow is cast over one half of the display (one background and its target) and the reflectance of the other background is adjusted so that the two backgrounds are equiluminant. The result is a dark coloured background under bright illumination and a light coloured background under dim illumination. The observer adjusts the target on one of the backgrounds to match the target on the other background. Lightness theories based on local luminance ratios predict that observers would make the targets equiluminant because the backgrounds have the same luminance and this would make the target-surround contrast the same for each display. However, observers actually set the target patch on the light albedo background (under shadow) to be lower in luminance than the other patch, indicating that they take into account the shadow being cast on this surface.

Other brightness models, such as filling-in models (e.g. Rudd, 2013), predict effects for simple centre-surround stimuli similar to those used in the present thesis. Rudd's model predicts different induction effects for increments and decrements, although it does not predict the crispening effect. A limitation to the explanatory power of brightness models for simple displays is that there are often instances where surfaces appear to be the same lightness but vary in appearance in other ways. For example, these theories cannot explain how a white surface under dim illumination looks perceptually different from a white surface under bright illumination; they predict that the two white patches would look identical and, more importantly, indistinguishable, which they would not. Additionally, these theories cannot explain lightness perception in the case of transparency, where two colours are perceived in the same location, or a surface is seen through a transparent medium like fog or smoke (Metelli, 1970, 1974a, 1974b). In the present experiments, the crispening effect is accompanied by a qualitatively different appearance of low contrast test patches, which cannot be explained by existing brightness models.

Matching the unique perceptual quality of simple displays

Other researchers have commented on the different perceptual quality of simple centre-surround displays (Logvinenko & Maloney, 2006; Vladusich, 2012, 2013; Vladusich et al., 2007), and the inability to satisfactorily match such displays by varying lightness/luminance alone (Brainard et al., 1997a; Katz, 1935; Maloney et al., 1995;

Wuerger et al., 1995). These researchers have suggested that multiple perceptual dimensions are needed to capture the full perceptual experience of achromatic surfaces. Vladusich et al. (2007) showed that as the contrast difference between test and matching displays increased, observers were progressively less able to produce perfect achromatic matches, as measured by a ten-point scale that rated the possibility of making a perfect match. Vladusich and colleagues have suggested that blackness and whiteness form the perceptual dimensions involved in the appearance of achromatic surfaces. However, they pointed out that their model incorrectly predicted the quality of match ratings in low-contrast centre-surround displays, and suggested that this may have to do with perceptual transparency. Similar to Vladusich, Logvinenko and colleagues (Logvinenko & Maloney, 2006; Logvinenko, et al., 2008; Logvinenko & Tokunaga, 2011; Tokunaga & Logvinenko, 2010a, 2010b) have argued that asymmetric matching can never lead to an exact match; observers can only set the minimum subjective difference between stimuli. In one experiment, Logvinenko and Maloney (2006) discarded the asymmetric matching method altogether, and had observers rate the dissimilarity of pairs of stimuli under different levels of illumination. They used a multidimensional scaling technique to plot and compare the similarity of each stimulus in a two dimensional space. They found that within an illuminant, achromatic colours fell along a single locus, forming a one-dimensional space. However, two distinct perceptual dimensions were needed to represent all of achromatic surface colour, and the second dimension (which they termed “brightness”) was associated with changes in illumination. Logvinenko and colleagues accounted for perceptual differences when surfaces were illuminated differently. However, they did not compare the appearance of achromatic surfaces in cases of transparency, where luminance values are perceptually divided into a background layer and transparent foreground layer. Our results, along with those of Ekroll and Faul (2013), suggest that asymmetric matching can lead to better matches when observers are able to vary the dimensions of lightness and transmittance.

If perceptual transparency does affect the lightness of test patches embedded in homogeneous surrounds, as suggested by some of the results in Chapter 4 and Ekroll and Faul’s (2013) findings, then it suggests that two separate mechanisms may be responsible for the perception of lightness in complex and homogeneous centre-surround displays: The effects of complex surface mesostructure and surface optics on lightness can be well explained by low-level contrast and luminance distributions across space and scale; conversely, lightness in homogeneous displays, which are descriptively simple, may involve segmentation of surfaces into layered image representations. Paradoxically, the homogeneous displays might evoke more complex scene representations than the ostensibly complex rocky displays. We suggest that if transparency or scission is responsible for the crispening effect, then representations of transparency need not be directly measurable or quantifiable, similar to observers' difficulty when matching illumination level (e.g. Rutherford & Brainard, 2002). Anderson and colleagues (Anderson, 1997, 1999, 2003a, 2003b; Anderson & Winawer, 2005, 2008) showed that geometric continuity of targets and surrounds, and consistent polarity relationships of the borders separating targets from the surrounds, determine how (or whether) scission is initiated and luminance values are separated into a foreground (see-through) and background layer. This in turn determines how (or if) luminance is partitioned between the different layers. Homogeneous centre-surround displays meet the conditions for this perceptual image decomposition. There is a "continuity" of the homogeneous surround and centre, although the distinct lack of geometry makes it ambiguous how luminance should be partitioned between the different layers. This ambiguity might be reflected in perception: homogeneous displays may evoke a weaker sense of transparency, and/or observers may not have perceptual access to this representation.

One clear result from Chapter 4 was that observers used transparent match displays to equate both the contrast of the edges and the lightness of the target in the homogeneous displays. Furthermore, observers preferred to match homogeneous displays to these transparent displays, over opaque matching displays. This suggests that observers require mid-level properties such as transparency to be varied in order to obtain more satisfactory matches to homogeneous displays, irrespective of any impressions of transparency. In these displays, low-level features such as edge contrast co-vary with lightness, and is something to which observers appear to have explicit perceptual access. Thus, observers may have been more effectively able to (and willing) to vary transmittance to match the contrast of the homogeneous centre-surround displays.

Relation to the simultaneous contrast effect

Regardless of whether lightness representations differ for homogeneous and inhomogeneous displays, the difference in the pattern of data between these two types of displays has important implications for a large body of literature investigating the simultaneous contrast (SC) effect. Previous studies have compared the size of the SC effect in uniform versus variegated centre-surround displays. It is often reported that SC is enhanced when surrounds are articulated (Adelson, 2000; Arend & Goldstein, 1987; Bressan & Actis-Grosso, 2006; Gilchrist et al., 1999; Schirillo, 1999a, 1999b), suggesting either poorer or better lightness constancy for articulated displays, depending on the experimental context. These displays are similar to the homogeneous (Experiment 1) and variegated (Experiment 2) displays used in the present study. Our results suggest that the enhanced SC effect for variegated displays may be a result of researchers evaluating lightness constancy using only a few test patch and surround values (another possibility is a failure to adequately equate the luminance of the surrounds). One advantage of the factorial combination of conditions used in Chapter 2 is that it reveals the complicated nature of the SC effect on homogeneous surrounds. For uniform displays, the SC effect clearly depends on the test patch's contrast against its surround (crispening). At low contrasts, these test patches seem to take on a qualitatively different appearance compared to test patches embedded in rocky or variegated displays. Therefore, it may not be meaningful to compare lightness constancy between uniform and articulated SC displays for single target and surround combinations.

We also note the large size of the SC effect in our experiments. In Chapter 1, perceived lightness of identical targets differed by as much as 4 Munsell steps when embedded in different surrounds. Even in the rocky glossy conditions, perceived lightness could differ by 2 Munsell steps or more. These effects are much larger than what is normally reported in the literature, which is about 0.5 to 1 Munsell step (Gilchrist, 2006). We are unsure why the effect is so strong for these stimuli, although we suspect it also has to do with the large range of test patch and surround reflectance values that were used.

The surfaces in the present study were different from simultaneous contrast (or centre-surround) displays that have previously dominated the literature. Previous studies have used physical experimental chambers with simple display setups and lighting (e.g. Blakeslee et al., 2008; Heinemann, 1955; Katona, 1935; Kraft & Brainard, 1999; Oyama, 1968), or computer generated displays with luminance values that are not separable into reflectance and illumination (e.g. Arend & Goldstein, 1987; Bäuml, 2001; Bressan & Actis-Grosso, 2006; Dixon & Shapiro, 2014; Laurinen et al., 1997; Schirillo, 1999a, 1999b; Soranzo & Agostini, 2004). There have been various reports that images presented on CRT monitors lead to different effects than directly viewed scenes (Agostini & Bruno, 1996; Brainard et al., 1997b; Kraft et al., 2002; Schirillo et al., 1990). The surfaces in the present experiments were computer rendered displays with simulated surface reflectance properties and complex illumination. We suggest that such displays could reduce perceptual differences between simulated and natural viewing because they allow for more information-rich conditions while maintaining controlled environments.

One limitation of rendered environments is that CRT monitors cannot capture the high dynamic range of luminance values experienced in natural scenes. In the present studies we compressed bright specular highlights to fit the luminance range of the monitor, which may have influenced their effectiveness as a potential cue to the illumination. However, using rendered scenes allowed us to test large factorial combinations of conditions, something that would not be feasible using real surfaces. It also allowed us to embed the scenes in a complex, natural light field, which is nearly impossible to control in physical laboratory setups. The findings in the present experiments are relevant to planar displays with surrounds containing one uniform reflectance value. The displays contained medium-scale surface relief but were globally planar surfaces viewed frontally. Future research should investigate whether the present findings extend to surfaces that vary in shape or orientation with respect to the light source, or surfaces with more than one albedo value in the surround. The experiments in the present thesis were intended as a first step to demonstrate the potential of using rendered environments to conduct controlled experiments with information-rich stimuli, while still being relatable to other lightness studies using centre-surround displays.

Conclusions

The experiments in the present thesis were unable to find support that the visual system computes surface lightness by forming a representation of the light field. Lightness perception in the complex displays used herein is well explained by low-level distributions of contrasts across space and scale (Chapter 2). Low-level explanations can also extend to complex displays under different levels of illumination (Chapter 3), but further investigation is required to determine whether this is true for cases of inhomogeneous illumination, such as scenes that contain illumination boundaries and/or gradients. Different mechanisms may play a role in the appearance of test patches embedded in homogeneous surrounds (Chapter 4). For such simple displays, it is arguably impossible to obtain a perfect control to rule out low-level effects on lightness. Dissociating mid-level transparency explanations from low-level contrast explanations will always be problematic, as by definition information is processed by the retina prior to higher visual processing areas responsible for the mid-level segmentation of surfaces. It is clear that targets embedded in homogeneous displays cannot be perfectly matched to targets on textured displays by varying lightness alone; a second perceptual dimension is required. Giving observers access to mid-level dimensions such as transparency provides a natural means for equating the contrast of the edges and the lightness of the target in homogeneous centre-surround displays. Thus, mid-level dimensions such as transmittance are at least required as methodological tools in order to tap into low-level perceptual dimensions such as perceived contrast.

References

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*(5142), 2042–2044.

Adelson, E. H. (1995). Checkershadow illusion. *Web.Mit.Edu*. Retrieved 2015, from http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga, *The New Cognitive Neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: The MIT Press.

Adelson, E. H., & Anandan, P. (1990). Ordinal characteristics of transparency (pp. 77–81). Presented at the AAAI-90 Workshop on Qualitative Vision, Boston, MA.

Adelson, E. H., & Pentland, A. P. (1996). The perception of shading and reflectance. In D. C. Knill & W. Richards, *Perception as Bayesian Inference* (pp. 409–423). New York: Cambridge University Press.

Agostini, T. A., & Bruno, N. (1996). Lightness contrast in CRT and paper-and-illuminant displays. *Perception & Psychophysics*, *58*(2), 250–258.

Agostini, T. A., & Galmonte, A. (2002). Perceptual organization overcomes the effects of local surround in determining simultaneous lightness contrast. *Psychological Science*, *13*(1), 89–93.

Agostini, T. A., & Proffitt, D. R. (1993). Perceptual organization evokes simultaneous lightness contrast. *Perception*, *22*(3), 263–272.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions. *Perception-London*, *26*(4), 419–453.

Anderson, B. L. (1999). Stereoscopic surface perception. *Neuron*, *24*, 919–928.

Anderson, B. L. (2003a). Perceptual organization in White's illusion. *Perception*, *32*, 269–284.

Anderson, B. L. (2003b). The role of occlusion in the perception of depth, lightness, and opacity. *Psychological Review*, *110*(4), 785–801.

Anderson, B. L., & Khang, B.-G. (2010). The role of scission in the perception of color and opacity. *Journal of Vision*, *10*(5), 26.

Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature*, *434*(7029), 79–83.

Anderson, B. L., & Winawer, J. (2008). Layered image representations and the computation of surface lightness. *Journal of Vision*, *8*(7), 18.1–22.

Anderson, B. L., Khang, B.-G., & Kim, J. (2011). Using color to understand perceived lightness. *Journal of Vision*, *11*(13), 19.

Anderson, B. L., Singh, M., & Meng, J. (2006). The perceived transmittance of inhomogeneous surfaces and media. *Vision Research*, *46*(12), 1982–1995.

Anderson, B. L., Whitebread, M., & de Silva, C. (2014). Lightness, brightness, and anchoring. *Journal of Vision*, *14*(9).

Annan, V., & Gilchrist, A. L. (2004). Lightness depends on immediately prior experience. *Perception & Psychophysics*, *66*, 943–952.

Arend, L. E., & Goldstein, R. (1987). Simultaneous constancy, lightness, and brightness. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *4*(12), 2281–2285.

Arend, L. E., & Spehar, B. (1993a). Lightness, brightness, and brightness contrast: 1. Illuminance variation. *Perception & Psychophysics*, *54*(4), 446–456.

Arend, L. E., & Spehar, B. (1993b). Lightness, brightness, and brightness contrast: 2. Reflectance variation. *Perception & Psychophysics*, *54*(4), 457–468.

Arend, L. E., Buehler, J. N., & Lockhead, G. R. (1971). Difference information in brightness perception. *Perception & Psychophysics*, *9*(3), 367–370.

Barrow, H. G., & Tenenbaum, J. M. (1978). Recovering intrinsic scene characteristics from images. In A. Hanson & E. Riseman, *Computer Vision Systems*. Academic Press.

Bäuml, K. H. (2001). Increments and decrements in color constancy. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *18*(10), 2419–2429.

Beck, J., Prazdny, S., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics*, *35*(5), 407–422.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, *39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2001). A multiscale spatial filtering account of the Wertheimer–Benary effect and the corrugated Mondrian. *Vision Research*, *41*, 2487–2502.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*(21), 2483–2503.

Blakeslee, B., & McCourt, M. E. (2012). When is spatial filtering enough? Investigation of brightness and lightness perception in stimuli containing a visible illumination component. *Vision Research*, *60*, 40–50.

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research*, *45*(5), 607–615.

Blakeslee, B., Reetz, D., & McCourt, M. E. (2008). Coming to terms with lightness and brightness: Effects of stimulus configuration and instructions on brightness and lightness judgments. *Journal of Vision*, *8*(11), 1–14.

Bloj, M. G., Ripamonti, C., Mitha, K., Hauck, R., Greenwald, S., & Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision*, *4*(9), 735–746.

Boyaci, H., Doerschner, K., & Maloney, L. T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision*, *4*(9), 664–679.

Boyaci, H., Doerschner, K., & Maloney, L. T. (2006a). Cues to an equivalent lighting model. *Journal of Vision*, *6*(2), 106–118.

Boyaci, H., Doerschner, K., Snyder, J. L., & Maloney, L. T. (2006b). Surface color perception in three-dimensional scenes. *Visual Neuroscience*, *23*(3-4), 311–321.

Boyaci, H., Maloney, L. T., & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision*, *3*(8), 541–553.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Brainard, D. H., & Maloney, L. T. (2011). Surface color perception and equivalent illumination models. *Journal of Vision*, *11*(5).

Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997a). Color constancy in the nearly natural image. I. Asymmetric matches. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *14*(9), 2091–2110.

Brainard, D. H., Rutherford, M. D., & Kraft, J. M. (1997b). Color constancy compared: experiments with real images and color monitors. *Investigative Opthalmology & Visual Science*, *38*(4), S476.

Bressan, P. (2001). Explaining lightness illusions. *Perception-London*, *30*(9), 1031–1046.

Bressan, P. (2006). The place of white in a world of grays: a double-anchoring theory of lightness perception. *Psychological Review*, *113*(3), 526–553.

Bressan, P., & Actis-Grosso, R. (2006). Simultaneous lightness contrast on plain and articulated surrounds. *Perception-London*, *35*(4), 445.

Cataliotti, J., & Gilchrist, A. L. (1995). Local and global processes in surface lightness perception. *Perception & Psychophysics*, *57*(2), 125–135.

Cohen, M. A., & Grossberg, S. (1984). Neural dynamics of brightness perception: features, boundaries, diffusion, and resonance. *Perception & Psychophysics*, *36*(5), 428–456.

Cornsweet, T. N. (1970). *Visual Perception*. New York: Academic Press.

Debevec, P. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography (pp. 189–198). Presented at the Computer Graphics (Proceedings of SIGGRAPH 98), Orlando, FL: ACM.

Devalois, R. L., & Devalois, K. K. (1988). *Spatial Vision*. New York: Oxford University Press.

Devalois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in the macaque visual cortex. *Vision Research*, *22*, 545–559.

Dixon, E., & Shapiro, A. G. (2014). Paradoxical effect of spatially homogenous transparent fields on simultaneous contrast illusions. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *31*(4), 307–313.

Doerschner, K., Boyaci, H., & Maloney, L. T. (2007). Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. *Vision Research*, *47*(28), 3409–3423.

Economou, E., Annan, V., & Gilchrist, A. L. (1998). Contrast depends on anchoring in perceptual groups. *Investigative Opthalmology & Visual Science*, *39*(4), S857.

Economou, E., Zdravković, S., & Gilchrist, A. L. (2007). Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *Journal of Vision*, *7*(12), 1–15.

Ekroll, V., & Faul, F. (2009). A simple model describes large individual differences in simultaneous colour contrast. *Vision Research*, *49*(18), 2261–2272.

Ekroll, V., & Faul, F. (2012a). Basic characteristics of simultaneous color contrast revisited. *Psychological Science*, *23*(10), 1246–1255.

Ekroll, V., & Faul, F. (2012b). New laws of simultaneous contrast? *Seeing and Perceiving*, *25*(2), 107–141.

Ekroll, V., & Faul, F. (2013). Transparency perception: the key to understanding simultaneous color contrast. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *30*(3), 342–352.

Ekroll, V., Faul, F., & Niederée, R. (2004). The peculiar nature of simultaneous colour contrast in uniform surrounds. *Vision Research*, *44*(15), 1765–1786.

Ekroll, V., Faul, F., & Wendt, G. (2011). The strengths of simultaneous colour contrast and the gamut expansion effect correlate across observers: evidence for a common mechanism. *Vision Research*, *51*(3), 311–322.

Evans, R. M. (1948). *An introduction to color*. New York: Wiley.

Faul, F., Ekroll, V., & Wendt, G. (2008). Color appearance: The limited role of chromatic surround variance in the gamut expansion effect. *Journal of Vision*, *8*(3), 3010–3020.

Gelb, A. (1929). Die Farbenkonstanz der Sehdinge. In A. Bethe, G. von Bergman, G. Embden, & A. Ellinger, *Handbuch der normalen und pathologischen Physiologie* (Vol. 12, pp. 594–687). Berlin: Springer.

Gelb, A. (1932). Die Erscheinungen des simultanen Kontrastes und der Eindruck der Feldbeleuchtung. *Zeitschrift Für Psychologie*, *127*, 42–59.

Gershun, A. (1939). The light field (P. Moon & G. Timoshenko, Trans.). *Journal of Mathematical Physics*, *18*, 51–151.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*(4274), 185–187.

Gilchrist, A. L. (1979). The perception of surface blacks and whites. *Scientific American*, *240*, 112–123.

Gilchrist, A. L. (1980). When does perceived lightness depend on perceived spatial arrangement? *Perception & Psychophysics*, *28*(6), 527–538.

Gilchrist, A. L. (1988). Lightness contrast and failures of constancy: a common explanation. *Perception & Psychophysics*, *43*(5), 415–424.

Gilchrist, A. L. (2006). *Seeing Black and White*. New York, NY: Oxford University Press.

Gilchrist, A. L., & Jacobsen, A. (1984). Perception of lightness and illumination in a world of one reflectance. *Perception-London*, *13*(1), 5–19.

Gilchrist, A. L., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, *33*(5), 425–436.

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T. A., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*(4), 795–834.

Grossberg, S. (1983). The quantized geometry of visual space: The coherent computation of depth, form, and lightness. *Behavioral and Brain Sciences*, *6*(4), 625–692.

Grossberg, S., & Mingolla, E. (1987). Neural dynamics of surface perception: Boundary completion, illusory figures, and neon color spreading. *Computer Vision, Graphics & Image Processing*, *37*, 116–165.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: a unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*(3), 241–277.

Hartline, H. K. (1940). The receptive fields of optic nerve fibers. *American Journal of Physiology*, *130*, 690–699.

Hartline, H. K., & Graham, C. H. (1932). Nerve impulses from single receptors in the eye. *Journal of Cellular Comparative Physiology*, *1*, 277–295.

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology*, *50*(2), 89–96.

Heinemann, E. G., & Chase, S. (1995). A quantitative model for simultaneous brightness induction. *Vision Research*, *35*(14), 2007–2020.

Helmholtz, H. V. (1962). *Treatise on Physiological Optics*. (J. P. L. Southall) (Vol. 2). New York: Dover.

Helson, H. (1938). Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. *Journal of Experimental Psychology*, *23*(5), 439.

Helson, H. (1943). Some factors and implications of color constancy. *Journal of the Optical Society of America*, 1–13.

Helson, H. (1964). *Adaptation-Level Theory*. New York: Harper & Row.

Helson, H., & Michels, W. C. (1948). The effect of chromatic adaptation on achromaticity. *Journal of the Optical Society of America*, *38*(12), 1025–1032.

Hering, E. (1964). *Outlines of a Theory of the Light Sense*. (L. M. Hurvich & D. Jameson). Cambridge, MA: Harvard University Press.

Hillis, J. M., & Brainard, D. H. (2007). Distinct mechanisms mediate visual detection and identification. *Current Biology*, *17*, 1714–1719.

Hochberg, J. E., & Beck, J. (1954). Apparent spatial arrangement and perceived brightness. *Journal of Experimental Psychology*, *47*(4), 263–266.

Hsia, Y. (1943). Whiteness constancy as a function of difference in illumination. *Archives of Psychology*, *248*.

Hurlbert, A. C., & Wolf, K. (2004). Color contrast: a contributory mechanism to color constancy. *Progress in Brain Research*, *144*, 147–160.

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, *64, Part 1*(6), 384–404.

Ikeda, M., Shinoda, H., & Mlzokami, Y. (1998). Three Dimensionality of the Recognized Visual Space of Illumination Proved by Hidden Illumination. *Optical Review*, *5*(3), 200–205.

Jaensch, E., & Müller, E. (1920). Über die Wahrnehmungen farbloser Helligkeiten und den helligkeitskontrast. *Zeitschrift Für Psychologie*, *83*.

Jameson, D., & Hurvich, L. M. (1964). Theory of brightness and color contrast in human vision. *Vision Research*, *4*(1), 135–154.

Kardos, L. (1934). Ding und Schatten [Object and shadow]. *Zeitschrift Für Psychologie*, *23*.

Katona, G. (1935). Color-contrast and color-constancy. *Journal of Experimental Psychology*, 49-63.

Katz, D. (1935). *The World of Colour*. London: Kegan Paul, Trench, Trubner & Co.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*(7), 652–673.

Kingdom, F. A. A., & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, *32*(8), 1565–1582.

Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2007a). The visual light field. *Perception-London*, *36*(11), 1595.

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shad(ow)ed random Gaussian surfaces. *Perception-London*, *33*(12), 1405–1420.

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2007b). Perception of illuminance flow in the case of anisotropic rough surfaces. *Perception & Psychophysics*, *69*(6), 895–903.

Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., Pas, te, S. F., & Pont, S. C. (2003). Illumination direction from texture shading. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *20*(6), 987–995.

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(1), 307–312.

Kraft, J. M., Maloney, S. I., & Brainard, D. H. (2002). Surface-illuminant ambiguity and color constancy: Effects of scene complexity and depth cues. *Perception-London*, *31*(2), 247–263.

Krauskopf, J., & Gegenfurtner, K. R. (1992). Color discrimination and adaptation. *Vision Research*, *32*(11), 2165–2175.

Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, *16*, 37–68.

Kuffler, S. W. (1973). The single-cell approach in the visual system and the study of receptive fields. *Investigative Opthalmology*, *12*, 794–813.

Lambert, J. H. (1760). *Photometria, sive de mensura et gradibus luminis, colorum et umbrae*. Augsburg: Eberhard Klett.

Land, E. H. (1986). An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 3078–3080.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America*, *61*(1), 1–11.

Landauer, A. A., & Rodger, R. S. (1964). Effect of "apparent" instructions on brightness judgments. *Journal of Experimental Psychology*, *68*, 80–84.

Laurinen, P. I., Olzak, L. A., & Peromaa, T. L. (1997). Early cortical influences in object segregation and the perception of surface lightness. *Psychological Science*, *8*(5), 386–390.

Lee, H.-C. (1986). Method for computing the scene-illuminant chromaticity from specular highlights. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *3*(10), 1694–1699.

Li, X., & Gilchrist, A. L. (1999). Relative area and relative luminance combine to anchor surface lightness values. *Perception & Psychophysics*, *61*(5), 771–785.

Logvinenko, A. D. (1999). Lightness induction revisited. *Perception-London*, *28*(7), 803–816.

Logvinenko, A. D. (2003). A fair test of the effect of a shadow-incompatible luminance gradient on the simultaneous lightness contrast. *Perception-London*, *32*(6), 717–730.

Logvinenko, A. D., & Maloney, L. T. (2006). The proximity structure of achromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, *68*(1), 76–83.

Logvinenko, A. D., & Ross, D. A. (2005). Adelson's tile and snake illusions: a Helmholtzian type of simultaneous lightness contrast. *Spatial Vision*, *18*(1), 25–72.

Logvinenko, A. D., & Tokunaga, R. (2011). Lightness constancy and illumination discounting. *Attention, Perception, & Psychophysics*, *73*(6), 1886–1902.

Logvinenko, A. D., Adelson, E. H., Ross, D. A., & Somers, D. (2005). Straightness as a cue for luminance edge interpretation. *Perception & Psychophysics*, *67*(1), 120–128.

Logvinenko, A. D., Petrini, K., & Maloney, L. T. (2008). A scaling analysis of the snake lightness illusion. *Perception & Psychophysics*, *70*(5), 828–840.

Mach, E. (1865). Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut. *Sitzungsberichte Der Mathematisch-Naturvissenshaftlichen Classe Der Kaiserlichen Akademic Der Wissenshaften*, *52*(2), 303–322.

Madigan, S. C., & Brainard, D. H. (2014). Scaling measurements of the effect of surface slant on perceived lightness. *I-Perception*, *5*(1), 53–72.

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In K. R. Gegenfurtner & L. T. Sharpe, *Color Vision: From Genes to Perception* (pp. 387–422). Cambridge, UK: Cambridge University Press.

Maloney, L. T. (2002). Illuminant estimation as cue combination. *Journal of Vision*, *2*(6), 493–504.

Maloney, L. T., Gerhard, H. E., Boyaci, H., & Doerschner, K. (2011). Surface color perception and light field estimation in 3D scenes. In L. R. Harris & M. Jenkin, *Vision in 3D Environments* (pp. 65–88). Cambridge, UK: Cambridge University Press.

Maloney, L. T., Wuerger, S. M., & Krauskopf, J. (1995). A method for testing Euclidian representations of proximity judgments in linear psychological spaces. In R. D. Luce, M. D'Zmura, D. Hoffman, G. J. Iverson, & A. K. Romney, *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow on his 70th birthday* (pp. 137–152). Mahwah, NJ: Erlbaum.

Marlow, P. J., Kim, J., & Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *Journal of Vision*, *11*(9).

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *207*(1167), 187–217.

McArthur, J. A., & Moulden, B. (1999). A two-dimensional model of brightness perception based on spatial filtering consistent with retinal processing. *Vision Research*, *39*(6), 1199–1219.

Metelli, F. (1970). An algebraic development of the theory of perceptual transparency. *Ergonomics*, *13*(1), 59–66.

Metelli, F. (1974a). Achromatic color conditions in the perception of transparency. In R. B. MacLeod & H. L. Pick, *Perception: Essays in Honor of J. J. Gibson*. Ithaca, NY: Cornell University Press.

Metelli, F. (1974b). The perception of transparency. *Scientific American, 230*. 90-98.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 221–245.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, *447*(7141), 206–209.

Olzak, L. A., & Thomas, J. P. (1991). When orthogonal orientations are not processed independently. *Vision Research*, *31*, 51–57.

Olzak, L. A., & Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision Research*, *32*, 1885–1898.

Oyama, T. (1968). Stimulus determinants of brightness constancy and the perception of illumination. *Journal of Psychological Research*, 1–10.

Pont, S. C., & Koenderink, J. J. (2003). Illuminance Flow. In *Computer analysis of images and patterns* (Vol. 2756, pp. 90–97). Berlin, Heidelberg: Springer Berlin Heidelberg.

Pont, S. C., & Koenderink, J. J. (2007). Matching illumination of solid objects. *Perception & Psychophysics*, *69*(3), 459–468.

Radonjić, A., & Gilchrist, A. L. (2013). Depth effect on lightness revisited: The role of articulation, proximity and fields of illumination. *I-Perception*, *4*(6), 437–455.

Ripamonti, C., Bloj, M. G., Hauck, R., Kiran, M., Greenwald, S., Maloney, S. I., & Brainard, D. H. (2004). Measurements of the effect of surface slant on perceived lightness. *Journal of Vision*, *4*(9), 747–763.

Rudd, M. E. (2001). Lightness computations by a neural filling-in mechanism. *Proceedings of the Society of Photo-Optical Engineers*, *4299*, 400–413.

Rudd, M. E. (2003). Progress on a computational model of human achromatic color processing. *Proceedings of the Society of Photo-Optical Engineers*, *170*, 181.

Rudd, M. E. (2013). Edge integration in achromatic color perception and the lightness-darkness asymmetry. *Journal of Vision*, *13*(14).

Rudd, M. E., & Arrington, K. F. (2001). Darkness filling-in: A neural model of darkness induction. *Vision Research*, *41*, 3649–3662.

Rudd, M. E., & Popa, D. (2007). Steven's brightness law, contrast gain control, and edge integration in achromatic color perception: A unified model. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *24*, 2766–2782.

Rudd, M. E., & Zemach, I. K. (2004). Quantitative properties of achromatic color induction: An edge integration analysis. *Vision Research*, *44*, 971–981.

Rudd, M. E., & Zemach, I. K. (2005). The highest luminance anchoring rule in achromatic color perception: Some counterexamples and an alternative theory. *Journal of Vision*, *5*(11), 5–5.

Rudd, M. E., & Zemach, I. K. (2007). Contrast polarity and edge integration in achromatic color perception. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *24*, 2134–2156.

Ruppertsberg, A. I., & Bloj, M. G. (2007). Reflecting on a room of one reflectance. *Journal of Vision*, *7*(13), 12–12.

Rutherford, M. D., & Brainard, D. H. (2002). Lightness Constancy: A Direct Test of the Illumination-Estimation Hypothesis. *Psychological Science*, *13*(2), 142–149.

Schirillo, J. A. (1999a). Surround articulation. I. Brightness judgments. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *16*(4), 793–803.

Schirillo, J. A. (1999b). Surround articulation. II. Lightness judgments. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *16*(4), 804.

Schirillo, J. A., Reeves, A. J., & Arend, L. E. (1990). Perceived lightness, but not brightness, of achromatic surfaces depends on perceived depth information. *Perception & Psychophysics*, *48*(1), 82–90.

Shapiro, A. G., & Lu, Z.-L. (2011). Relative brightness in natural images can be accounted for by removing blurry content. *Psychological Science*, *22*(11), 1452–1459.

Sharan, L., Li, Y., Motoyoshi, I., Nishida, S., & Adelson, E. H. (2008). Image statistics for surface reflectance perception. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *25*(4), 846–865.

Singh, M., & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review*, *109*(3), 492–519.

Snyder, J. L., Doerschner, K., & Maloney, L. T. (2005). Illumination estimation in three-dimensional scenes with and without specular cues. *Journal of Vision*, *5*(10), 863–877.

Soranzo, A., & Agostini, T. A. (2004). Impossible shadows and lightness constancy. *Perception*, *33*(11), 1359–1368.

Takasaki, H. (1966). Lightness change of grays induced by change in reflectance of gray background. *Journal of the Optical Society of America*, *56*(4), 504–509.

Thomas, J. P., & Olzak, L. A. (1996). Uncertainty experiments support the roles of second-order mechanisms in spatial frequency and orientation discriminations. *Journal of the Optical Society of America A. Optics, Image Science, and Vision*, *13*, 689–696.

Thompson, W. B., Fleming, R. W., Creem-Regehr, S. H., & Stefanucci, J. K. (2011). *Visual perception from a computer graphics perspective*. Boca Raton, FL: Taylor & Francis Group.

Todd, J. T., Norman, J. F., & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science*, *15*(1), 33–39.

Tokunaga, R., & Logvinenko, A. D. (2010a). Material and lighting dimensions of object colour. *Vision Research*, *50*(17), 1740–1747.

Tokunaga, R., & Logvinenko, A. D. (2010b). Material and lighting hues of object colour. *Ophthalmic and Physiological Optics*, *30*(5), 611–617.

Toscani, M., Valsecchi, M., & Gegenfurtner, K. R. (2013). Optimal sampling of visual information for lightness judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11163–11168.

Vladusich, T. (2012). Simultaneous contrast and gamut relativity in achromatic color perception. *Vision Research*, *69*, 49–63.

Vladusich, T. (2013). Gamut relativity: a new computational approach to brightness and lightness perception. *Journal of Vision*, *13*(1), 14.

Vladusich, T., Lucassen, M. P., & Cornelissen, F. W. (2007). Brightness and darkness as perceptual dimensions. *PLoS Computational Biology*, *3*(10), 1849-1858.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology*, *38*(3), 310–324.

Ward, G. J. (1994). The RADIANCE lighting simulation and rendering system (pp. 459–472). Presented at the Computer Graphics (Proceedings of SIGGRAPH 94), Orlando, FL: ACM.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, *25*(11), 1661–1674.

White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception*, *10*, 215–230.

Whittle, P. (1992). Brightness, discriminability and the “crispening effect.” *Vision Research*, *32*(8), 1493–1507.

Wiesel, T. N., & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, *29*, 1115–1156.

Williams, S. M., McCoy, A. N., & Purves, D. (1998). The influence of depicted illumination on brightness. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(22), 13296–13300.

Wilson, H. R., & Wilkinson, F. (2003). Spatial channels in vision and spatial pooling. In L. M. Chalupa & J. S. Werner, *The visual neurosciences* (pp. 1060–1068). Cambridge, MA: MIT Press.

Wollschläger, D., & Anderson, B. L. (2009). The role of layered scene representations in color appearance. *Current Biology*, *19*(5), 430–435.

Wuerger, S. M., Maloney, L. T., & Krauskopf, J. (1995). Proximity judgments in color space: Tests of a Euclidean color geometry. *Vision Research*, *35*, 827–835.

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Vision Research*, *41*(20), 2581–2600.