

Surface properties and the perception of colour

Zoey J. Isherwood^{1,2}, Quan Huynh-Thu^{3,4}, Matthew Arnison^{3,5}, David Monaghan^{3,5}, Matteo Toscani⁶, Stuart Perry⁷, Vanessa Honson¹, Juno Kim^{1*}

¹*School of Optometry and Vision Science, University of New South Wales (UNSW) Sydney, Australia*

²*School of Psychology, University of Nevada, Reno, Nevada, USA*

³*Canon Information Systems Research Australia (CiSRA), Macquarie Park, NSW, Australia*

⁴*Nearmap, Sydney, Barangaroo, NSW, Australia*

⁵*Bandicoot Imaging Sciences, NSW, Australia*

⁶*Abteilung Allgemeine Psychologie, Justus-Liebig-Universität Giessen, Giessen, Germany*

⁷*Faculty of Engineering and IT, University of Technology Sydney (UTS), NSW, Australia*

KEYWORDS: colour perception, saturation, hue, lightness, gloss, specular, 3D shape, virtual reality.

***Corresponding author:**

A/Prof Juno Kim

Sensory Processes Research Laboratory

School of Optometry and Vision Science

University of New South Wales

NSW 2052

AUSTRALIA

Email: juno.kim@unsw.edu.au

Tel: +61 2 9385 7474

Abstract

We examined whether perception of colour saturation and lightness depends on 3D shape and surface gloss of colour surfaces rendered to have different hues. In Experiment 1, we parametrically varied specular roughness of predominantly planar surfaces with different mesoscopic relief heights. The orientation of surfaces was varied relative to the light source and observer. Observers matched perceived lightness and chroma (effectively saturation) using spherical objects rendered using CIE LCH colour space. We observed strong interactions between perceived saturation and lightness with changes in surface orientation and surface properties (specular roughness and 3D relief height). Declines in saturation and increases in lightness were observed with increasing specular roughness. Changes in relief height had greater effects on perceived saturation and lightness for blue hues compared with reddish and greenish hues. Experiment 2 found inverse correlations between perceived gloss and specular roughness across conditions. Experiment 3 estimated perceived specular coverage and found that a weighted combination of perceived gloss and specular coverage could account for perceived colour saturation and lightness with different coefficients accounting for perceptual experience in for each of the three hue conditions. These findings suggest that perceived colour saturation and lightness depends on the separation of specular highlights from diffuse shading informative of chromatic surface reflectance.

Introduction

Surfaces reflect light that we use to infer the physical properties of objects (e.g., 3D shape, colour, lightness, and glossiness). The visual system must recover this information from a finite pattern of luminance variations contained in 2D retinal images. Understanding the underlying processes involved is challenging, because there are instances where the perception of any one surface attribute can be influenced by others (e.g., Marlow, Kim and Anderson, 2012). In this study, we consider the effects that surface gloss and relief height have on perceptual attributes of surface colour.

Evidence from previous literature suggests that colour and gloss are co-dependent perceptual constructs that, in combination, affect perceived glossiness, colour constancy, and colour estimation. For instance, perceived gloss can change depending on object colour, whereby brighter colours (e.g., yellow) are perceived as less glossy than darker colours (e.g., cyan)—a phenomenon known as *contrast gloss* (Hunter, 1937; Ferwerda et al. 2001; Baar et al., 2015). In other work, increasing the glossiness of an object has been found to improve colour constancy, showing that glossy objects have a more stable percept of colour under varying illuminations compared to matte objects (Granzier, Vergne and Gegenfurtner, 2014). Colour constancy has also been found to depend on the object's hue, where *blue* objects generate greater improvements in colour constancy between glossy and matte counterparts compared with *red* and *green* objects. Object shape has also been found to play a role whereby smooth shapes have higher colour constancy. However, this improvement was only found for the *blue* and *red* hues (Granzier et al., 2014). In terms of colour estimation, glossiness has also been found to affect the perceived lightness and colour saturation of objects (Xiao and Brainard, 2008). Objects appear darker and more saturated with increasing glossiness, with a more pronounced effect for darker colours (e.g., *purple*) than lighter colours (e.g., *yellowish green*). Although only a handful of studies have specifically investigated colour and shape perception in 3D objects with both real (Baar et al., 2015; Giesel and Gegenfurtner, 2010; Granzier et al., 2014) and rendered objects (Yang & Maloney, 2001; Boyaci et al., 2004; Fleming et al., 2004; Schultz et al., 2006; Boyaci et al., 2006; Xiao and Brainard, 2008; Xiao, Hurst, MacIntyre and Brainard, 2012; Matsumoto, Fukuda and Uchikawa, 2016), these studies have found that surface colour has a profound effect on the perception of other surface properties.

In more recent work on surface properties, Honson et al. (2020) investigated how object shape, specularity and surface orientation affect perceived colour and gloss. To investigate these material properties, they manipulated the mesoscopic relief height, specular roughness and orientation of surfaces relative to a primary overhead light source. They measured perceived colour saturation and value (of “green” surfaces rendered in HSV colour space) and identified complex interactions between the perception of these colour attributes and surface orientation, specular roughness and surface relief height. Transforming their data to a perceptually-uniform CIE LCH colour space preserved the effects of specular roughness and relief height on saturation (i.e., C^*/L^*) and lightness. The authors proposed that the visual misattribution of specular content to diffuse shading could explain the changes in perceived saturation and lightness. Perceived gloss and specular coverage were found to vary in complex ways with changes to specular roughness and mesoscopic relief height. For example, specular coverage was greater when surfaces (and their average surface normals) were oriented at 45° to bisect the angle formed between the viewing and lighting direction. When specular coverage was increased this way, perceived saturation was found to substantially decline. However, when surface relief height was increased, this decline in perceived saturation was limited. This effect of increasing relief on perceived saturation was attributed to reduced specular coverage caused by increased diversity in the range of surface normal orientations. Despite these complex interactions, the researchers showed that a weighted-linear combination of perceived gloss and specular coverage could account for perceived saturation and lightness.

Given the handful of studies conducted in this research area, the observed differences in perceived surface properties (e.g., glossiness, saturation and colour) as a function of surface hue are relatively misunderstood. While differences in perceived surface properties have been observed across a variety of hues, it is particularly notable in many of the aforementioned studies that significant differences were found for *bluish* hues compared to *red* and *green*. There is a large body of evidence indicating that differences in colour perception related to the colour blue may have to do with perceptual biases along the blue/yellow axis (Churma, 1994; Pearce et al., 2014, Winkler et al., 2015; Aston et al., 2019; Radonijic et al., 2016; Weiss et al., 2017). These biases are thought to reflect the presence of colours along the *daylight* locus in the environment since natural illumination throughout the day is dominated by blue and yellow light.

Indeed, it has been found that colour constancy is worse under blue illumination (Pearce et al., 2014; Winkler et al., 2015), suggesting that the visual system may have a *prior* that accounts for (or rather *discounts*) blue illumination when estimating object hue. This inability to disentangle a *blue* surface hue from the illuminant in certain situations has been noted across a handful of studies (Churma, 1994; Pearce et al., 2014, Winkler et al., 2015; Aston et al., 2019; Radonijic et al., 2016; Weiss et al., 2017). However, it is relatively unknown to what extent this bias for *blue* affects perceived surface properties for 3D objects. The majority of research on the blue/yellow perceptual bias has so far only used 2D images to study this phenomenon and/or focused solely on measuring colour constancy (Churma, 1994; Pearce et al., 2014; Winkler et al., 2015). While some studies investigating differences across hues have been conducted using real or rendered 3D objects (Xiao and Brainard, 2008; Giesel and Gegenfurtner, 2010; Xiao et al., 2012; Granzier et al., 2014), it is unclear the extent to which glossiness and specular coverage (as investigated previously in Honson et al. 2020) could account for perceptual differences in saturation and lightness judgments across a wider range of surface and material properties (e.g., mesoscopic relief height, specular roughness, orientation to light source, and hue).

To address this gap, here we further evaluated the findings of Honson et al. (2020) by assessing whether the effects of gloss on perceived colour saturation and lightness differ across a wider range of hues—red, green, and blue. We measured perceived lightness and saturation (Experiment 1), perceived gloss (Experiment 2), and perceived specular coverage (Experiment 3) to surfaces rendered across these hue conditions at different levels of relief height, specular roughness, and surface orientation relative to the light source. Given that the findings of Honson et al. (2020) may be explained by the inability to disentangle specular content from diffuse shading, we predicted that differences in the pattern of results for perceived lightness and saturation would be observed for *blue* surfaces, but be similar for *red* and *green* surfaces. This difference may be predicted by perceptual biases that exist along the blue/yellow axis that make it difficult for the visual system to disentangle a *blue* surface hue from the illuminant (Churma, 1994; Pearce et al., 2014, Winkler et al., 2015; Aston et al., 2019; Radonijic et al., 2016; Weiss et al., 2017).

Experiment 1

The purpose of Experiment 1 was to extend the findings of the original Honson et al. (2020) study to multiple hues. Honson et al. (2020) found there was large variation in perceived saturation and lightness as a function of changes in specular roughness, mesoscopic shape, and the orientation of the planar surfaces to the light source—which can profoundly change the way light interacts with/and is distributed across a surface. Specular coverage will be greater when orienting the surface at 45 degrees so its normals bisect the angle formed between the illumination and viewing directions. The increased specular coverage may make perceptual separation of specular and diffuse information more challenging, especially when specular roughness is increased. However, increasing relief height or slanting surfaces more frontally to reduce specular coverage should improve perceptual identification of diffuse colour.

Contrary to the view that perceived saturation depends on specular highlight roughness, Xiao and Brainard (2008) found that observer judgments of colour appearance were robust against changes in specular parameters between target and matching spheres. In addition, perceived colour saturation and lightness were found to be more robust for yellow-green hues compared with purple hues. When viewing objects with more complex structure, such as the tile stimuli used in Honson et al (2020), specular highlights (amongst other manipulated factors such as relief height and specular roughness) can greatly distort perceived colour appearance. Since Xiao & Brainard (2008) used a less complex stimulus (a smooth sphere), it is possible that this may have played a role in the observed robustness in perceived saturation across specular manipulations. It is also unknown whether the differences Xiao & Brainard (2008) observed between yellow-green and purple hues would also extend to more complex stimuli.

To ascertain whether biases in perceived lightness are constrained by hue in complex stimuli, we examined the effect of specular roughness and mesoscopic relief height on the perception of saturation and lightness along three distinctly different hue axes. Honson et al. (2020) found that transforming colour matching data from HSV to CIE LCH colour space generated no main effects on perceived chroma (C^*), but preserved the effects on perceived lightness and saturation (i.e., C^*/L^* – the ratio of chroma to lightness). In this instance, saturation is the estimated *colourfulness* of a surface patch that is proportional to

the patch's perceived brightness (Fairchild, 2013). In CIE LCH space, saturation can be computed as C^*/L^* (see Schiller & Gegenfurtner, 2016; Schiller, Valsecchi & Gegenfurtner, 2018). We used the CIE LCH colour space to obtain perceptual estimates of chroma, lightness and saturation. The CIE LCH colour space was used to control the dynamic ranges in saturation and lightness across different hue axes. This allowed us to determine whether there were any differences in the effect of specular roughness and mesoscopic relief height on the appearance of surface colours represented on a standard RGB display.

Materials and Method

Observers

Nine observers with normal or corrected-to-normal vision participated in the experiment, four of whom were authors of the present study (QHT, MA and ZI, DM). All procedures were approved by the Human Research Ethics Advisory Panel (HREA) at the University of New South Wales (UNSW Sydney).

Stimuli

The general process for generating stimuli followed the similar procedures as Honson et al. (2020). Initially, the upper face of a cube was subdivided into a 203 x 203 vertex mesh. The remaining four vertices that formed the other five faces of the cube were moved towards the upper face to model the profile of a 3D tile 10 cm × 10 cm × 3cm (width × height × thickness). Mesoscale shape was introduced to the upper face by displacing each vertex along the orthogonal z-axis according to a base cloud noise texture map (Size: 1.00, Nbla: 0.03 and Depth: 0) that was generated in Blender 3D (Blender Foundation, Version 2.77). The depth of 0 generated comparatively smoother surfaces than the depth of 1 used by Honson et al. (2020). The values of the displacement map were scaled by different amounts to vary the amplitude of undulations in mesoscopic surface shape along an orthogonal axis to the surface plane. Additional smoothing was performed using the Corrective Smooth modifier in Blender (Factor: 1.0 and Repeat: 12). This smoothing improved the quality of the 3D modelling following displacement mapping. Stimulus images were rendered using an opaque BSDF implemented using Cycles render in Blender 3D. The diffuse component was rendered using a Lambertian BSDF, and the specular component was rendered using the microfacet Beckmann function. The number of rays was 5 for the rendering of each stimulus image.

To investigate how perceived colour lightness and chroma are affected by a surface’s hue, surfaces were generated across three different hue angles corresponding to red, green, and blue respectively: 39.999°, 136.016°, and 306.285°. These hue angles were chosen because they fell on the three primary axes of sRGB space. The perceptually-uniform CIE-LCH colour space was used to ensure each hue was perceptually uniform in lightness and chroma—i.e., increases in lightness and chroma are perceptually equal across hues. Furthermore, we chose this space as past research has shown that the Lab colour space (of which LCH is the polar version) accurately predicts perceived saturation (C^*/L^*) (Schiller, Valsecchi & Gegenfurter, 2018). For luminance around $\sim 40 \text{ cd/m}^2$ (middle range of a typical computer monitor), it is the best predicting colour space compared to other spaces such as HSV and LUV, and perceived saturation for different hues is on average closer than 1 JND to predictions.

The 3D tiles were generated to have a lightness and chroma value fixed at a value of 60 across the three hue angles. Since Blender 3D does not use LCH colour space, LCH values were converted to sRGB values prior to rendering. These values are listed in Table 1 below, and were converted to sRGB space using custom Python scripts (Version 2.7.0) with embedded calls to the *colormath* library (Version 2.2.0; <https://github.com/gtaylor/python-colormath>).

Table 1. LCH and corresponding sRGB values used in Experiment 1

Hue	LCH	sRGB (0-1)
Red	60, 60, 39.999°	0.91, 0.42, 0.31
Green	60, 60, 136.016°	0.32, 0.64, 0.26
Blue	60, 60, 306.285°	0.63, 0.49, 0.90

Surfaces were centred within a simulated lighting environment that was similar to standard viewing chambers used in real-world psychophysical experiments on material appearance. Figure 1 provides an overview of the setup for simulated viewing and lighting conditions used in this experiment. The simulated room was a cube (3.28 m^3) with completely matte walls and floor. We used a large overhead rectangular emitter ($2.5 \text{ m} \times 1.0 \text{ m}$) containing an additional two rectangular area lights $6 \text{ cm} \times 120 \text{ cm}$ (Energy: 100) to generate natural primary lighting of surfaces embedded in our viewing chamber (Figure 2A).

The colour of the light was the neutral colour of the display (i.e., $r=g=b$). A camera with a focal length of 35 mm was situated 60 cm from the midpoint of the surface. This distance was appropriate to ensure the surfaces remained in full view across changes in their angular orientation of θ around the horizontal x-axis (Figure 1B). Figure 1C shows rendered sample images obtained for θ values of 15°, 30° and 45°.

[INSERT FIGURE 1 ABOUT HERE]

At each of the three surface orientations, we parametrically varied mesoscopic relief height and specular roughness, as exemplified in the 30° condition for red (Figure 2), green (Figure 3) and blue surfaces (Figure 4). The complete set of conditions across different surface orientations is provided in Appendix A. We varied mesoscopic relief height over four levels using the vertex displacement modifier in Blender (0.025, 0.050, 0.100, and 0.200). These values scaled the intensity range of the displacement map and generated undulations in mesoscopic shape with peak-to-peak amplitudes that approximately corresponded to 2.5%, 5%, 10%, and 20% of the surface's width. We also parametrically varied specular roughness over five levels (0.010, 0.100, 0.200, 0.300 and 0.400) while holding specular amplitude at 0.2 in the simulation. This range of specular roughness levels was chosen to be similar to ranges used in previous research (Honson et al., 2020; Mooney and Anderson, 2014). Specular amplitude was held constant at 0.20, as used previously to generate the realistic glossy appearance of common natural materials (e.g., Marlow, Kim and Anderson, 2012).

Images were rendered for each stimulus conditions at a resolution of 2000 × 2000 pixels in 24-bit RGB bitmap format. The rendered images were generated using the Cycles Render for Blender 2.77, which was controlled by a custom python script running on a Dell Precision 5510 with Intel (R) Core i7-6820HQ CPU@2.70GHz computer. Path tracing was used with 128 render samples per pixel. The simulated light paths were set with default parameters for full global illumination. These rendering parameters were appropriate for generating images that could be sub-sampled to 800 × 800 using the Lanczos filter method to eliminate noise. Custom stimulus presentation software was used to present these

images on an Eizo CG275W monitor (27-inch diagonal with resolution 1920 x 1200, sRGB colour mode, 8 bits per channel, 2.2 gamma, D65 colour temperature). Images were viewed at a distance of approximately 70 cm for an effective size of approximately $\pm 10^\circ$ visual angle (horizontal and vertical). The estimated luminance from sRGB values computed over a square region of the surface images for the three different hues was roughly equated: 40.6 (Red), 41.9 (Green) and 42.4 (Blue).

[INSERT FIGURES 2, 3 & 4 ABOUT HERE]

Procedure

Prior to participating, observers were informed that they would be required to make perceptual matches of surface colour for planar surface images that were presented in a random order on a computer monitor. Training was offered for some observers to gain familiarity with what they were required to match. The pre-rendered images used in training were of a smooth planar surface devoid of mesoscopic surface changes presented on the left side of the display. Most of the observers were confident they understood the task after completing several trials. For the actual experiment, three blocks (each with a different hue condition) of 60 images were presented in a randomised order on the left side of the display (4 levels of relief height, 5 levels of specular roughness, and 3 levels of orientation relative to the light source). The order of blocks was pseudorandomised across observers for performing the matches to each of the three hue conditions. A short rest break of up to 5 minutes was provided between blocks. The entire experiment took each of the participants no longer than 1 hour to complete.

Perceptual matches of lightness and chroma were made using pre-rendered images of a matte sphere devoid of specular reflections that was presented on the right side of the display (Figure 5). We used a sphere to ensure that the distribution of surface orientations was compatible with all three surface orientations of target planes. A sample scenario of the matching task is shown in Figure 6. Observers used the arrow keys on a standard keyboard to move through a pre-rendered 22x22 matrix of images (22 levels of lightness and 22 levels

of chroma). The hue of matching spheres was the same as the target surface. The LCH space was used to control the matching sphere's changes in lightness and chroma to ensure each step would be perceptually uniform across hues. Horizontal keypresses increased or decreased lightness (ranging from 18.750 to 97.500, step size: 3.75). Vertical keypresses increased or decreased chroma (ranging from 22.500 to 101.250, step size: 3.75). The observer pressed the spacebar to record the setting that appeared to most closely match the lightness and chroma of the target plane. Responses were recorded to ASCII file for subsequent data analysis. We used a higher lightness (97.5) and chroma (101.25) than the values used to generate the tiles (60.0) because Honson et al. (2020) found saturation to be perceived up to ~10% higher than the simulated saturation when matching with matte spheres, depending on orientation, relief height, and specular roughness. We verified that the monitor could display these extended LCH combinations by converting the values to an sRGB colour space to and ensuring these values were within sRGB bounds.

[INSERT FIGURES 5 & 6 ABOUT HERE]

Data analysis

Observer settings of chroma and lightness were separately averaged across observers for plotting purposes. A four-way repeated-measures Analysis of Variance (ANOVA) was used to test for main effects of specular roughness, relief height and surface orientation on perceived chroma and lightness. The analysis was performed separately on the three different hue conditions.

Results and discussion

Figures 7 and 8 show data for perceived chroma and lightness, respectively. We observed a very slight decline in chroma with increased specular roughness across the three hue conditions. There appears to be a clear increase in lightness with increasing specular roughness across all hue conditions.

Perceived saturation (C/L*)*

A four-way repeated-measures ANOVA found a significant main effect of hue on perceived saturation ($F_{2,16}=4.09$, $p < .05$). There were also main effects of surface orientation ($F_{2,16} = 82.99$, $p < .00001$) and relief height on perceived saturation ($F_{3,24} = 71.10$, $p < .00001$). There was also a significant main effect of specular roughness of perceived saturation ($F_{4,32} = 87.12$, $p < .00001$). There was a significant two-way interaction effect between hue and relief height on perceived saturation ($F_{6,48} = 21.89$, $p < .00001$). There was a significant two-way interaction effect between surface orientation and relief height on perceived saturation ($F_{6,48} = 4.62$, $p < .001$). There was also a significant two-way interaction effect between relief height and specular roughness on perceived saturation ($F_{12,96} = 3.09$, $p < .001$). We also found a significant three-way interaction effect across hues in the relationship between relief height and specular roughness on perceived saturation ($F_{12,96} = 2.16$, $p < .01$). There were no other interaction effects on perceived saturation.

The results suggest that viewing orientation plays an important role in the perception of colour saturation with obliquely oriented surfaces generating lower levels of perceived saturation than more frontally oriented surfaces. There was a clear and consistent effect of relief height on perceived saturation across all conditions, with greater relief generating higher perceived saturation. Overall, there was a consistent inverse association between specular roughness and perceived saturation; increasing specular roughness was found to reduce perceived saturation. This effect of specular roughness on perceived saturation was found to be consistent across the three hues, but the effect of relief height was differential; effects of relief height on perceived saturation were greatest for the blue hue and lowest for the green hue. This was confirmed by the significant two-way interaction effect between hue and relief height on perceived saturation. Further, the three-way interaction effect suggests that the effect of specular roughness of perceived saturation varies non-linearly with relief height and the extent of this non-linearity differs across hues.

[INSERT FIGURE 7 ABOUT HERE]

Perceived lightness

A four-way repeated-measures ANOVA found a significant main effect of hue on perceived lightness ($F_{2,16} = 5.58, p < .05$). There were also main effects of surface orientation ($F_{2,16} = 91.65, p < .00001$) and relief height on perceived lightness ($F_{3,24} = 21.30, p < .00001$). There was also a significant main effect of specular roughness on perceived lightness ($F_{4,32} = 47.65, p < .00001$). There was a significant two-way interaction effect between hue and relief height on perceived lightness ($F_{6,48} = 3.58, p < .01$). There was a significant two-way interaction effect between surface orientation and relief height on perceived lightness ($F_{6,48} = 5.76, p < .0005$). There was a significant two-way interaction effect between surface orientation and specular roughness on perceived lightness ($F_{8,64} = 3.90, p < .001$). There was also a significant two-way interaction effect between relief height and specular roughness on perceived lightness ($F_{12,96} = 2.47, p < .01$). We also found a significant three-way interaction effect across hues in the relationship between surface orientation and specular roughness on perceived lightness ($F_{16,128} = 1.81, p < .05$). There was also a significant three-way interaction effect across surface orientations in the relationship between surface relief height and specular roughness on perceived lightness ($F_{24,192} = 1.72, p < .05$). There was no other interaction effects on perceived lightness found.

For lightness matches, across the three hues we observe a consistent linear increase in matched lightness with increasing specular roughness, following the findings of Honson et al. (2020). This increase in matched lightness is consistent across all relief heights and orientation conditions, the only difference being the finding that the magnitude of the effect monotonically increased with higher roughness and orientation values (e.g., 0.4 roughness at 45°). The magnitude of this monotonic increase in matched lightness significantly differed across hues. This difference was observed in the significant two-way interactions between relief height and specular roughness, as well as between relief height and surface orientation. The significant three-way interactions suggest that the effect of specular roughness on perceived lightness varies as a function of both relief height, and this relationship varies with either surface orientation or hue. Perceived lightness was more consistent across changes in relief height and specular roughness for red surfaces and lowest in consistency across these conditions for blue surfaces. These effects depended on

viewing orientation with more frontally aligned surfaces generating lower variation in perceived lightness with relief height across levels of specular roughness for all hues.

[INSERT FIGURE 8 ABOUT HERE]

Taken together, these data suggest that perceived saturation and lightness were most distorted when relief height was low and specular roughness was high, whereby the stimulus is perceived to have lower saturation and higher lightness than the true values (Lightness 60, Chroma 60). Interestingly, the magnitude of this effect was largest for *blue* surfaces and was relatively similar between *red* and *green*. Thus, the question arises: *why do large differences only occur only for blue?*

One may argue that the differences we observe may be due to sRGB gamut limits, i.e. observers would choose settings that are out of gamut, and thus artificially create hue differences, since the limits for L^* and C^* depend on H^* . To figure out whether this was the case, we have plotted in the left panel of Appendix B the C^* limit (max C^*) for each L^* level, for the three H^* values we used, i.e. the largest C^* value at which the corresponding RGB triplet has values between 0 and 1. The red hue has the smallest range (Appendix B, left panel) to select from, which would predict large ceiling effects for red when estimating Lightness—however, this was not the case in our results (see Figure 8). On the right panel of Appendix B we have plotted the saturation limit (max S) as a function of L^* for the three hues. The borders differ, potentially creating gamut artifacts. We have also plotted the maximum saturation settings (from the data plotted in Figure 7) with corresponding L^* in the right panel of Appendix B (from the data plotted in Figure 8; green circle for the green hue, red x, blue +). The three maxima nearly overlap, indicating that differences between hues are not caused by different upper limits of saturation. To compute the limit, for each L^* level (0 to 100), and each of the three H^* values used in the present study, we have input in a broad range of C^* values, converted to sRGB and looked at the maximum C^* values that produces within sRGB gamut. Although the green hue gets close to the maxima, since the

highest saturation settings overlap and do not touch the upper limits, this suggests that our results were not driven by differences in the upper limits of different hues.

So, what else can account for differences in lightness and saturation judgments observed for *blue*? It is possible that the different pattern of results we observe for *blue* may be due to perceptual biases along the blue/yellow axis (Churma, 1994; Pearce et al., 2014, Winkler et al., 2015; Aston et al., 2019; Radonijic et al., 2016; Weiss et al., 2017). Past research has found that colour constancy is worse when a scene is illuminated by blue light (Pearce et al., 2014; Winkler et al., 2015), and that it is difficult for the visual system to disentangle blue surface hues from the illuminant in certain scenarios (Churma, 1994; Pearce et al., 2014; Winkler et al., 2015). For instance, Winkler et al. (2015) found that people tend to attribute bluish tints to the illuminant rather than the object itself.

The findings of these past studies suggest that the visual system has a *prior* for discounting/accounting for *blue* illumination when estimating surface hue. In the present study, given the large differences in the way the light source illuminates the tile stimuli across orientations, mesoscopic relief heights, and specular roughness levels, it is likely that the process of disentangling the illuminant from the tiles when judging lightness and saturation would substantially differ for *blue* hues compared to *red* and *green*. When estimating lightness this inability to disentangle the illuminant from the tiles themselves may be the reason lightness judgments were inflated for blue tiles with low relief heights. Specifically, it has been demonstrated by Toscani et al. (2013) that the lightness of an object is determined by what is deemed as the brightest part of the object—which can be a problem if the illuminant cannot be correctly parsed from the object hue.

Although perceived saturation was variable across the three hues, the magnitude of the effect of decreasing saturation with increasing specular roughness is extremely subtle. These data are consistent with the dominant response in lightness and saturation based on colour matching data transformed from HSV to CIE LCH space by the Honson et al. (2020) study. Across all surface images, we found a systematic effect of relief height on perceived saturation and lightness. This effect would appear to suggest that increases in the distribution of specular reflections and not only their clarity (i.e., roughness) is important for estimating surface saturation and lightness. The effect of surface relief height on perceived

colour attributes is consistent with our previous research (Honson et al., 2020). We therefore examined the effect of relief height and specular roughness on perceived gloss and specular coverage in the experiments that follow.

Experiment 2

The previous experiments found that chroma and lightness judgments were strongly dependent on specular roughness, mesoscopic relief, and surface orientation. Although perceived lightness was found to increase significantly with increasing specular roughness, perceived saturation declined more for the short wavelength surface (blue hue) and less for red and even less for green. The results of Experiment 1 followed the findings of Honson et al. (2020), whereby perceived lightness was furthest from veridical when specular roughness was high and mesoscopic relief height was low. It may be possible these differences arise due to the incomplete separation of specular reflections from diffuse shading when specular edges have increased roughness. If some of the specular energy were (mis-)classified as diffuse shading in this way, then proportionally less classifiable specular content would be available for generating the experience of gloss. While differences in perceived lightness were subtle across hue, it may be possible that differences in this incomplete separation may explain these effects on perceived colour. In Experiment 2, we investigated whether hue affects perceived gloss across changes in specular roughness, mesoscopic relief, and surface orientation.

Materials and Method

Observers

Four observers with normal or corrected-to-normal vision were recruited. All participants were authors of this study, but were well trained in performing psychophysical tasks. All procedures received local ethical clearance and followed the principles outlined in the Declaration of Helsinki.

Stimuli

We used planar surface images that were identical to those used in Experiment 1. However, because we used the paired-comparisons method here, we eliminated the 0.200 specular roughness value to reduce the number of trials. We eliminated this particular specular roughness value based on the findings of Honson et al. (2020), since non-linearities in perceived gloss were not present between specular roughness values 0.010 and 0.300. As

such, the conditions for specular roughness were 0.010, 0.100, 0.300, 0.400; and for relief height were 0.025, 0.050, 0.100, 0.200. The same orientation (15°, 30° and 45°) and hue (red, green, blue) conditions were used as in Experiment 1.

Procedure

Following Honson et al. (2020), we measured perceived surface gloss using the paired-comparisons method. Observers were asked to select which of two images presented side-by-side appeared glossier or shinier using the LEFT/RIGHT arrow keys on the keyboard. The terms glossier or shinier were used to ensure that observers were attending to the overall specular component and not merely the material quality of the surface. Observer responses were recorded to an ASCII file for subsequent analysis. To minimise the number of trials in a session, we broke up the experiment into nine blocks; one for each level of surface orientation (15°, 30° and 45°) and hue (red, green, blue). Within each block there were 240 trials based on the 4 levels of specular roughness and 4 levels of relief height (16 x 16 – 16 conditions). We focussed our comparisons on relief height and specular roughness as these conditions were found to generate the strongest main effects in Experiment 1. Image pairs were fully randomised, and we pseudorandomised the order for performing blocks of trials at each surface orientation and hue across observers. Although the paired images were presented for an unlimited period of time, observers tended to make their judgments within approximately 5 s. Observers took between 1.5 to 2.0 hours to complete the experiment, and all 9 blocks were either completed in one session with short breaks, or over a period of two days.

Data analysis

As in Honson et al. (2020), we computed probability estimates of perceived gloss for each image in each condition by dividing the number of times the image was selected as glossier by the number of times it was presented on the display. Probability estimates of perceived gloss were analysed using a series of repeated-measures two-way ANOVAs in the open-access statistical package R.

Results and discussion

Perceived gloss was plotted as a function of specular roughness for each of the levels of relief height and viewing orientation (Figure 9). Separate rows of axes show data obtained for surfaces having each of the three different diffuse colours. We analysed the results separately for each of the nine conditions below.

[INSERT FIGURE 9 ABOUT HERE]

Red surfaces were analysed for effects on perceived gloss at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was a significant main effect of specular roughness ($F_{3,9}=244.2$, $p < 0.00001$) and relief height ($F_{3,9}=281.7$, $p < 0.0005$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=77.44$, $p < 0.0005$). For the 30° condition, there was a significant main effect of specular roughness ($F_{3,9}=320.6$, $p < 0.00001$) and relief height ($F_{3,9}=24.52$, $p < 0.0005$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=5.65$, $p < 0.0005$). For the 45° condition, there was a significant main effect of specular roughness ($F_{3,9}=180.7$, $p < 0.00001$) and relief height ($F_{3,9}=9.01$, $p < 0.005$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=5.65$, $p < 0.0005$).

Next, green surfaces were analysed for effects on perceived gloss at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was a significant main effect of specular roughness ($F_{3,9}=114.8$, $p < 0.00001$) and relief height ($F_{3,9}=165.1$, $p < 0.00001$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=32.97$, $p < 0.00001$). For the 30° condition, there was a significant main effect of specular roughness ($F_{3,9}=159.9$, $p < 0.00001$) and relief height ($F_{3,9}=13.88$, $p < 0.005$). However, there was no significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=1.22$, $p < 0.32$). For the 45° condition, there was a significant main effect of specular roughness ($F_{3,9}=693$, $p <$

0.00001) and relief height ($F_{3,9}=24.29$, $p < 0.0005$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=4.47$, $p < 0.005$).

Finally, blue surfaces were analysed for effects on perceived gloss at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was a significant main effect of specular roughness ($F_{3,9}=438$, $p < 0.00001$) and relief height ($F_{3,9}=165.4$, $p < 0.00001$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=70.25$, $p < 0.00001$). For the 30° condition, there was a significant main effect of specular roughness ($F_{3,9}=169.2$, $p < 0.00001$) and relief height ($F_{3,9}=33.22$, $p < 0.00005$). However, there was no significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=2.07$, $p < 0.07$). For the 45° condition, there was a significant main effect of specular roughness ($F_{3,9}=404.1$, $p < 0.00001$) and relief height ($F_{3,9}=5.89$, $p < 0.05$). There was also a significant interaction effect between specular roughness and relief height on perceived gloss ($F_{9,27}=7.37$, $p < 0.00005$).

We find the same consistent decline in perceived gloss with increasing specular roughness, especially for surfaces oriented more frontally to the observer and obliquely relative to the primary lighting direction, which other studies have reported (Honson et al., 2020; Marlow, Kim and Anderson, 2012). We also found the same pattern of results as Honson et al. (2020) in that the effect of varying relief height on perceived gloss was more variable for surfaces that were oriented frontally (i.e., 15°). This can be observed in Figure 10, where there is larger displacement between curves plotted for the 15° condition compared to the increasingly slanted surfaces across all hues. An inverse pattern of responses across specular roughness values can also be observed for the lowest relief height value (0.025) across all hues (Figure 9); increases in perceived gloss are observed with increasing specular roughness values that asymptote between values 0.300 and 0.400. We found very little difference in the effects of relief height and specular roughness on perceived gloss across different hues. These results suggest that perceived gloss alone cannot account for the differences we observed in Experiment 1 for perceived lightness and saturation.

Experiment 3

In Experiment 1, the magnitude of changes in perceived lightness and chroma were different across the three hues. Experiment 2 results suggest these differences are unlikely to be attributed to differences in perceived gloss, as percepts of gloss were somewhat invariant across changes in hue. For the 15° condition, there was a consistent decline in perceived gloss with increasing specular roughness in all relief heights except for the 0.025 condition which had the opposite pattern – an increase in perceived gloss with increasing specular roughness which plateaued between 0.300 and 0.400. Given these findings, perceived gloss per se does not appear to account for differences in perceived colour attributes (chroma and lightness) observed across the three different hues we considered.

An alternative explanation for perceived changes in lightness and chroma could be specular coverage (Honson et al., 2020). Increasing relief height from a flat plane will increase the local curvature of the surface, which in turn increases the range of surface normal orientations across the same surface regions. This increase in range of surface normal orientations will increase the number of surface regions with normals that bisect that angle between the viewing and illumination vectors, generating specular highlights. The variation in the distribution of specular highlights that results (specular highlight coverage) may potentially be used as a cue to differentiating diffuse from specular regions for assessing surface lightness and saturation. This specular coverage cue has been found in previous work to account for perceived gloss across a variety of viewing conditions (Marlow, Kim and Anderson, 2012) and has been further extended to assess its covariance with perceived colour saturation (Honson et al., 2020). In Experiment 3, we determined whether perceived specular coverage of our planar surfaces could account for the observed changes in perceived chroma and lightness with changes in specular roughness and relief height.

Materials and Method

Observers

Four observers with normal or corrected-to-normal vision were recruited. All participants were authors of this study, but were well trained in performing these

psychophysical tasks. All procedures were approved for ethical clearance from CiSRA and followed the principles outlined in the Declaration of Helsinki.

Stimuli

All experimental stimuli were the same as those used in Experiment 2.

Procedure

The procedure for the current experiment was identical to the previous paired-comparisons experiment, except for a change in instruction. Here, the task posed to the observers was to “select which of the two images appears to have greater surface area covered by specular highlights”. Responses were recorded using the same procedures as used in Experiment 2.

Data analysis

Data were analysed in the same way as described in Experiment 2. We computed probability estimates of perceived coverage for each image in each condition by dividing the number of times the image was selected as having more specular coverage by the number of times it was presented. We also performed Pearson’s correlations to determine whether a weighted linear combination of perceived gloss and coverage estimates could account for perceived colour attributes of lightness and chroma.

Results and discussion

Perceived specular highlight coverage was plotted as a function of specular roughness for each of the levels of relief height and viewing orientations (Figure 10). Separate rows of axes show data obtained for surfaces having each of the three different diffuse colours. We analysed the result separately for each of the nine conditions below.

Red surfaces were analysed for effects on perceived highlight coverage at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was no significant main effect of specular roughness ($F_{3,9}=1.28$, $p = 0.34$), but there was a significant main effect of relief height ($F_{3,9}=17.05$, $p < 0.0005$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=4.54$, $p < 0.005$). For the 30° condition, there was no significant main effect of specular

roughness ($F_{3,9}=1.46$, $p = 0.29$), but there was significant main effect of relief height ($F_{3,9}=24.78$, $p < 0.0005$). There was no significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=1.85$, $p = 0.11$). For the 45° condition, there was a significant main effect of specular roughness ($F_{3,9}=19.02$, $p < 0.0005$) and relief height ($F_{3,9}=14.38$, $p < 0.001$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=9.69$, $p < 0.00001$).

[INSERT FIGURE 10 ABOUT HERE]

Green surfaces were analysed for effects on perceived highlight coverage at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was no significant main effect of specular roughness ($F_{3,9}=0.86$, $p = 0.49$), but there was a significant main effect of relief height ($F_{3,9}=43.87$, $p < 0.0001$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=9.69$, $p < 0.00001$). For the 30° condition, there was no significant main effect of specular roughness ($F_{3,9}=2.64$, $p = 0.11$), but there was a significant main effect of relief height ($F_{3,9}=17.23$, $p < 0.0005$). There was a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=2.90$, $p < 0.05$). For the 45° condition, there was a significant main effect of specular roughness ($F_{3,9}=5.43$, $p < 0.05$) and relief height ($F_{3,9}=7.93$, $p < 0.01$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=6.48$, $p < 0.0001$).

Blue surfaces were analysed for effects on perceived highlight coverage at each viewing orientation using separate repeated-measures ANOVAs. For the 15° condition, there was no significant main effect of specular roughness ($F_{3,9}=0.23$, $p = 0.87$), but there was a significant main effect of relief height ($F_{3,9}=43.76$, $p < 0.0001$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=9.46$, $p < 0.00001$). For the 30° condition, there was no significant main effect of specular roughness ($F_{3,9}=1.33$, $p = 0.32$), but there was a significant main effect of relief height ($F_{3,9}=17.23$, $p < 0.0005$). There was a significant interaction effect between specular

roughness and relief height on perceived coverage ($F_{9,27}=3.31$, $p < 0.01$). For the 45° condition, there was no significant main effect of specular roughness ($F_{3,9}=2.18$, $p = 0.16$), but there was a significant main effect of relief height ($F_{3,9}=8.71$, $p < 0.01$). There was also a significant interaction effect between specular roughness and relief height on perceived coverage ($F_{9,27}=7.47$, $p < 0.0001$).

Across all three hues, we find the same complex, systematic differences in perceived specular coverage across different orientations, specular roughnesses, and relief heights as Honson et al. (2020). With low relief height values, coverage was predicted to be incrementally greater with increasing the orientation of the stimulus away from the observer and towards the light source. With high relief height the opposite was predicted, whereby coverage would be incrementally lower with increasing surface orientation away from the observer and towards the light source. Although these predictions are supported by the pattern of responses we observed in the present study, the magnitude of curves for relief heights across specular roughness values seem to differ quite considerably between the present experiment and Honson et al. (2020). For example, the curve plotting values for a relief height of 0.200 at an orientation of 15° had the highest probability of being chosen to have higher coverage in Honson et al. (2020) (about 75% of the time), but was chosen only about 50% of the time in the present experiment. These differences in magnitude may be reflective of the lightness and chroma values chosen for the stimuli in the present experiment (60 and 60 respectively), which were considerably lower compared to the values used in Honson et al. (2020) (71.679 and 101.237 respectively).

We find that relief height and specular roughness generated consistent variations in perceived specular highlight coverage across hues. Similar to Honson et al. (2020), we observed an interaction effect demonstrated by the apparent “travelling wave” in the plots across surface orientation conditions. The amount of perceived specular coverage tended to increase with increasing specular roughness for the 15° viewing condition, but decreased with increasing specular roughness for the 45° viewing condition. Increasing relief height increased coverage for the 15° condition where coverage was initially very low, but reduced it for the 45° orientation where initial coverage was higher. The strong interaction effects we found here were consistent across the different hue conditions.

Despite the similarity across hues in the effects of surface properties on perceived gloss (Experiment 2) and specular highlight coverage (Experiment 3), we attempted to ascertain whether some weighted linear combination of perceived gloss and coverage could account for variations in perceived chroma and lightness we observed in Experiment 1. This analysis was based on the three observers who participated in these three relevant experiments, all of whom were authors (ZI, QH, MA). We separately fit the following weighted linear models:

$$P'_{saturation} = W * P_{coverage} + (1 - W) * (1 - P_{gloss})$$

$$P'_{lightness} = W * P_{coverage} + (1 - W) * (1 - P_{gloss})$$

where, $P'_{saturation}$ and $P'_{lightness}$ refers to the *estimated* perceived saturation and lightness, respectively, from $P_{coverage}$ and P_{gloss} corresponding to the perceived coverage and gloss estimated through psychophysical experiments (Experiment 2 and 3). The model weight (W) was estimated empirically based on linear least-squares fits between the model output and actual psychophysical data of perceived saturation and lightness. We allowed the weight of the model to vary parametrically between -1 and 1 at 0.01 steps to identify the correlation that accounted for the most variability in perceptual data on saturation and lightness. A value for W of 0 indicates no weighting on perceived coverage and 100% on perceived gloss. A value for W of 1 indicates 100% weighting on perceived coverage and no weighting on perceived gloss. For example, if surfaces are perceived as glossier, then it is assumed that the visual system has intrinsically attributed more of the specular image structure to glossiness. If the visual system attempts to attribute image content to physical causes in this way (Barrow and Tenenbaum, 1978), then the perception of colour saturation and lightness should be less biased by residual specular image structure not attributed to glossiness.

Figure 11 shows the correlations for all surface colours and viewing orientations between perceived saturation (based on C^*/L^* results of Experiment 1) and a weighted linear combination of perceived gloss (Experiment 2) and specular highlight coverage (Experiment 3). Data were averaged across all participants within each of the three experiments. Data for different relief heights and specular roughness levels were then pooled for assessing the performance of model predictors in each of the nine plots. The legend in each plot shows the value of W (with the correlation coefficient and p-value) that

was found to account for the greatest variability in perceived saturation. The average model across all conditions pooled returned a weight of $W = 0.41$, weighting gloss 59% and coverage 41%. Eyeballing the separate conditional plots suggests that model weights for the blue hue were similar to the average model response. Variability in model weighting was greater for the red surfaces and greatest for green surfaces. All viewing orientations and surface colours generated a high level of prediction of perceived saturation using the optimal weighted linear combination of perceived gloss and coverage. Indeed, in Experiment 1, we found consistent dependence of perceived saturation on relief height and specular roughness for these conditions.

On close inspection of the data pooled across all conditions, the correlation between perceived coverage and gloss was found to be relatively low, though significant ($r = 0.32$, $p < 0.0001$). Using these pooled data, we found that the optimum weight for fitting our linear model to account for perceived saturation was $W = 0.41$ ($r = -0.77$, $p < .000001$), which is very close to the weights required to optimally fit responses to estimates of perceived saturation of blue surfaces across each viewing orientation. Forcing the weight to 0.41 in all sub-conditions resulted in slightly different fits as shown in Table 2. We find that the blue surfaces generated the same model correlations when optimized using a variable W or when forced to the 0.41 weighting. Red and green surfaces generated relatively high correlations across all conditions, but fits were consistently lower when forced to the 0.41 weighting. All correlations were significant ($p \leq 0.001$). The results of simple linear regression models suggest that both gloss and coverage are together useful for explaining the perception of colour saturation across hues with emphasis on gloss over coverage.

[INSERT FIGURE 11 ABOUT HERE]

Table 2. Model fits for perceived saturation using weights for individual and pooled data. Significant correlation indicated by asterisk ($p \leq 0.001$)

Hue	Slant	w	r_w	$r_{0.41}$
Red	15	+0.56	-0.85*	-0.75*
	30	+0.32	-0.94*	-0.93*
	45	+0.48	-0.87*	-0.85*
Green	15	+0.48	-0.78*	-0.75*
	30	-0.02	-0.87*	-0.77*
	45	+0.15	-0.92*	-0.84*
Blue	15	+0.43	-0.73*	-0.73*
	30	+0.47	-0.91*	-0.91*
	45	+0.42	-0.88*	-0.88*

We also considered the ability of our model in accounting for perceived lightness. Figure 12 shows the correlations for all surface colours and viewing orientations between perceived lightness (based on results of Experiment 1) and a weighted linear combination of perceived gloss (Experiment 2) and specular highlight coverage (Experiment 3). Here, the model was found to account very well for the perception of lightness across all the three different surface colours we used. Significant correlations were found for conditions where surfaces were oriented at 30° or 45° only. No significant correlations were found between perceived lightness and our weighted linear model for the 15° viewing condition. When we forced values to the optimum model fit to the pooled data ($W = 0.34$), there was little change in the overall strength of the model correlations with perceived lightness across hues (see Table 3).

[INSERT FIGURE 12 ABOUT HERE]

The model generated significant negative correlations with the perceived saturation of all three hues for surfaces viewed at 15° ($p \leq 0.001$). Although the model accounted for perceived lightness at more oblique viewing orientations, this was not the case in the more frontal viewing condition of 15°. Model predictions were therefore more reliable in predicting saturation than lightness. Hence, despite the variability in perceived colour

saturation found in Experiment 1, we find here that a weighted linear combination of gloss and specular coverage largely accounts for the appearance of colour saturation. Emphasis on gloss and coverage was more consistent for blue hues, as greater dependence on gloss was found in green conditions. We examine the broader implications of these findings further in the general discussion.

Table 3. Model fits for perceived lightness using weights for individual and pooled data. Significant correlation indicated by asterisk ($p \leq 0.005$)

Hue	Slant	w	r_w	$r_{0.34}$
Red	15	+0.47	+0.41	+0.38
	30	0	+0.86*	+0.78*
	45	+0.26	+0.84*	+0.84*
Green	15	+0.11	+0.18	+0.17*
	30	+0.09	+0.86*	+0.82*
	45	+0.25	+0.94*	+0.93*
Blue	15	+0.81	-0.08	+0.02*
	30	+0.35	+0.92*	+0.92*
	45	+0.45	+0.72*	+0.7*

General discussion

The present study sought to investigate and model how hue, surface orientation, mesoscopic and microscale shape can affect the perceived lightness and saturation (estimated from chroma and lightness) of an object. To this end, we generated planar surfaces and varied the surface hue, the orientation of the surface towards the light source (from above), as well as the mesoscopic relief height and specular roughness. Under each of these conditions, we parametrically varied specular roughness.

In Experiment 1, we found the perception of an object's lightness and saturation could be biased by changes imposed in the object's orientation to the light source and its surface properties. The magnitude in which a surface's perceived lightness increased and saturation decreased systematically varied as a function of the parameters we manipulated. Perceived saturation declined, and perceived lightness increased with increasing specular roughness, replicating the findings of previous studies using HSV colour matching (Honson et al. 2020). Differences in the magnitude of these effects were observed across the three different hues we tested (red, green and blue), whereby the magnitude in which perceived saturation decreased across specular roughness was greatest for the shorter wavelength blue hue. These effects do not appear to be attributed to ordinal differences in luminance, as we roughly equated luminance across the three hues and the mean luminance of blue surfaces was close to the luminance of red and green surfaces. However, there were strong dependencies of perceived lightness on specular roughness across the three hue conditions. The magnitude of the difference in perceived lightness and saturation observed for blue, but not red and green, may indeed be due to perceptual biases that lie along the blue/yellow axis (Churma, 1994; Pearce et al., 2014; Winkler et al., 2015; Weiss et al., 2017).

Given that blue and yellow are the dominant illuminant colours along the *daylight* locus, it is thought that the visual system has specific *priors* that account for/discount the colour of illuminant along this locus when making colour judgments. It has indeed been found that colour constancy is worse under blue illumination (Pearce et al., 2014; Winkler et al., 2015), and that bluish tints are attributed to the illuminant rather than the object itself (Churma, 1994; Pearce et al., 2014; Winkler et al., 2015). Along these lines of reasoning, we propose that the specular highlights in our blue tile stimuli are more difficult to segment and therefore may be attributed to the illuminant rather than the tile's colour per se. Due to this

inability to disentangle specular highlights from tile hue, lightness judgments can be artificially inflated (especially for tiles with low relief heights), as lightness judgments have been found to be based on the brightest part of the object itself (Toscani et al., 2013).

Interestingly, despite the fact that specular coverage estimates across all three hues were essentially the same (Experiment 3), the fact that saturation and lightness judgments for *blue* significantly differed from *red* and *green* suggests that these judgments are not influenced by top down knowledge. Rather, the bias for *blue* we observe is bottom up and cannot be “overwritten”—at least in the case of our experimental manipulation(s). It would be interesting in future work to see if we would observe similar sorts of biases for *yellow* tiles, as the visual system has evolved not only to *blue* illuminants along the daylight locus, but also to *yellow* illuminants. It is possible that we would observe a different pattern of results as it has been demonstrated that *yellow* illuminants are more strongly attributed to the hue of objects in a scene compared to *blue* illuminants (Winkler et al., 2015). However, more research is needed to extend these findings to more complex stimuli such as the stimuli used in the present study.

Although we did not test many different hues to assess whether any effect of changing hue is continuous or categorical, research on constancy suggests that hue effects may involve a gradual continuous profile (Weiss, Witzel & Gegenfurtner, 2017). That study found colour constancy was preserved at the borders of different colour categories, rather than generating large differences relative to the centre of categorical zones. It would however be useful to explore the potential effects of colour we report across a wider range of hues.

In Experiment 2, we found that perceived gloss generally declined with increasing specular roughness across all viewing orientations and relief heights. Similar to Honson et al. (2020), the only condition that followed the inverse was the low relief height viewed at the 15° viewing angle. There were no clear differences in the pattern of perceived gloss data across the three different surface hues, suggesting that additional information would be required to account for differences in perceived lightness and saturation we observed across hues in Experiment 1. Hence, Experiment 3 considered the role of *perceived* specular highlight coverage. We obtained estimates of perceived specular highlight coverage from the same observers who participated in Experiment 2. We observed a pattern of data that

can be described as a “travelling wave” with increasing surface orientation away from the viewing direction. The wave consisted of a displacement in peak perceived coverage towards lower specular roughness at more oblique viewing orientations. This wave-like pattern was primarily observed for low relief heights. Greater relief heights tended to maintain an inverted-U shaped pattern of perceived coverage across changes in surface orientation.

Applying a similar weighted linear regression model as used in Honson et al. (2020) to our data, we found that perceived coverage and gloss could account very well for the perceived lightness of surfaces viewed at oblique angles of 30° and 45°. However, the weighted linear model performed better at accounting for perceived saturation at more frontal viewing orientation of 15°. These complex dependencies of perceived saturation and lightness on viewing orientation were on the whole observed across the different hues used to render the diffuse component of our 3D planar surfaces. Hence, we find that perceived lightness and saturation depends not only on viewing orientation, but also on surface properties of hue, relief height and specular roughness.

Although perceived gloss and coverage measurements are essentially very similar across hues, these properties may be weighted differently across hues when judging the perceived saturation and lightness of our surface stimuli. For example, the model weighting on gloss and coverage was found to generally diagnostic of perceived saturation across hues, but was found to emphasise gloss more for green surfaces. The model weighting was more consistent for blue surfaces. This result suggests that additional information may be useful for constraining perceived saturation for green surfaces, which may not be the case for blue surfaces. This finding may suggest that different priors may be at play when judging saturation and lightness (e.g. difficulty parsing surface hue from the illuminant) compared to gloss and surface coverage judgments. Irrespective of the differences however, the variations in perceived lightness and saturation appear to be largely explained by perceptual indices of perceived gloss and specular highlight coverage.

The apparent variability in the model explanatory power of gloss and specular coverage over perceived lightness across different hues might be attributed to the separability of specular and diffuse shading components (Anderson & Kim, 2009). Indeed, surfaces with higher specular roughness were estimated to have higher lightness and lower

saturation than surfaces with sharp specular reflections. The blurring of specular edge contours appears to have made it perceptually challenging to separate the diffuse component from the specular component for colour saturation and lightness analysis. Indeed, we found the increasing specular roughness generally increased perceived lightness and reduced perceived saturation. In support of the view that the visual system separates diffuse from specular shading for colour perception, evidence suggests that surface regions covered by specular highlights are ignored when participants estimate a surface's body colour or lightness (Kim et al., 2012; Toscani et al., 2017). The co-dependence of perceived saturation on not just perceived gloss but also perceived specular coverage supports the view that participants selectively attend to surface regions that are not covered in specular highlights when estimating colour saturation and lightness. Due to the tendency of specular highlights to appear near brighter regions of diffuse shading in natural scenes (Kim, Marlow & Anderson, 2011), observers will tend to be biased towards estimating colour using darker regions of diffuse shading (Toscani et al., 2013; Toscani and Valsecchi, 2019). Consistent with this view, we found that surfaces with greater relief height were perceived as darker overall, compared with surfaces with lower relief height, presumably due to the greater variability in orientation of surface normal pointing away from the primary lighting direction.

In earlier work, Marlow et al. (2012) found that perceived gloss could be explained by a weighted linear combination of specular sharpness, contrast and coverage in monocular images. The model we used in the present study was proposed to explain the perceived lightness and saturation of surfaces, which was assumed to depend on the brain's (in-)ability to separate diffuse from specular content. Our model was successful in accounting for variations in perceived lightness (and saturation). It is possible the model could be improved by obtaining additional data on perceived specular sharpness and contrast (in addition to coverage). However, this approach involves a purely perceptual model that requires psychophysical information to be obtained from human observers on the appearance of the specular properties of surfaces they see in images. It would be useful to improve the utility of this model by automating the measurement of specular highlight coverage and other specular attributes computationally.

In summary, we find complex dependencies of perceived saturation and lightness on the physical properties of surfaces. Although some of the differences might be attributed to

residual inaccuracies in representing LCH positions in sRGB space, our findings show there exists complex non-linear perceptual dependencies of perceived lightness and saturation on: 1) the intended hues of 3D surfaces rendered; 2) surface orientation relative to the light source; 3) mesoscopic surface relief height; 4) microscale specular roughness. It would be beneficial in future research to determine what aspects of 3D rendering and the simulated lighting environment might best account for these perceptual variations so that the representation of colour (and the perception of it) can be most appropriately managed.

Acknowledgements

This work was supported in part by an Australian Research Council (ARC) Future Fellowship awarded to JK (FT140100535) and was aided by the Sensory Processes Innovation Network (SPINet). We also thank Michael A. Webster for comments on early iterations of the manuscript.

Figure legends

Figure 1. The simulated lighting environment. A: The lighting environment was an enclosed viewing chamber fitted with an overhead rectangular area light. Note that the near corner of the viewing chamber has been cut away for the purposes of showing the arrangement of overhead lighting, the planar 3D surface (red hue), viewing camera, and internal matte walls (Ref=0.4) and floor (Ref=0.1). B: Surfaces were slanted obliquely, relative to the camera's viewing direction, by angular rotations (θ) around the horizontal axis (shown in grey). C: Sample images showing views of the surface slanted away from frontoparallel to the viewing direction by angles of 15°, 30°, or 45°.

Figure 2. Rendered images of red surfaces varying multi-parametrically in mesoscopic shape (across rows) and specular roughness (across columns). Images shown for the surface oriented at 30° in slant.

Figure 3. Rendered images of green surfaces varying multi-parametrically in mesoscopic shape (across rows) and specular roughness (across columns). Images shown for the surface oriented at 30° in slant.

Figure 4. Rendered images of blue surfaces varying multi-parametrically in mesoscopic shape (across rows) and specular roughness (across columns). Images shown for the surface oriented at 30° in slant.

Figure 5. Spherical surfaces were used in the asymmetric colour matching task. A: The surfaces were situated within the same light field as the bumpy planar surfaces, but with a mural of random brick texturing situated on the far wall behind the spheres. The sphere was also rendered on a flat plane that provided some ambient lighting to the underside of the sphere to increase the realism of the display and accurate perception of shape. B: Selections were varied by pre-set steps in chroma (along columns) and lightness (along rows).

Figure 6. Stimulus presentation and feedback. A: Example of a single experimental trial. A randomly presented plane is situated on the left of the display. The observer's task was to adjust the lightness and chroma of the matte sphere on the right to match the plane.

Figure 7. Mean ratings of perceived saturation (C^*/L^*) for red (upper row), green (middle row) and blue surfaces (lower row) plotted as a function of specular roughness. Different lines are used to represent each relief height and separate axes (left, middle, right) respectively show data for the 15°, 30° and 45° surface orientations.

Figure 8. Mean ratings for perceived lightness for red (upper row), green (middle row) and blue surfaces (lower row) plotted as a function of specular roughness. Different lines are used to represent each relief height and separate axes (left, middle, right) respectively show data for the 15°, 30° and 45° surface orientations.

Figure 9. Means and standard errors of probability estimates for perceived gloss of planar surfaces varying in specular roughness and relief height across 3 different hues: red, green, and blue (top, middle and bottom rows, respectively). Data plotted in each column represent the 3 surface orientations used in the experiment (15°, 30° and 45°, respectively). Different line types and symbols are shown in the legend and correspond to data across the different relief heights.

Figure 10. Means and standard errors for perceived coverage of surfaces by specular highlights across 3 different hues: red, green, and blue (top to bottom row respectively). Data plotted in each column represent the 3 surface orientations used in the experiment (15°, 30° and 45°). Different line types and symbols are shown in the legend and correspond to data across the different relief heights.

Figure 11. Mean perceived saturation (C^*/L^*) plotted as a function of our weighted linear model based on perceived coverage and gloss. Results shown for the 3 different hues: red, green, and blue (top, middle and bottom rows, respectively). Data plotted in each column represent the 3 surface orientations (15°, 30° and 45°). Legends show the weighting (W) for the best model fit with the associated correlation coefficient and p-value. Data based on averages across three observers for sixteen conditions pooled for each plot (4 levels relief height x 4 levels specular roughness).

Figure 12. Mean perceived lightness plotted as a function of our weighted linear model based on perceived coverage and gloss. Results shown for the 3 different hues: red, green, and blue (top, middle and bottom rows, respectively). Data plotted in each column represent the 3 surface orientations (15°, 30° and 45°). Legends show the weighting (W) for the best model fit with the associated correlation coefficient and p-value. Data based on averages across three observers for sixteen conditions pooled for each plot (4 levels relief height x 4 levels specular roughness).

References

- Anderson, B.L., Kim, J. (2009) Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11), 10, 1-17.
- Aston, S., Radonjić, A., Brainard, D. H., & Hurlbert, A. C. (2019). Illumination discrimination for chromatically biased illuminations: Implications for color constancy. *Journal of Vision*, 19(3):15, 1–23, <https://doi.org/10.1167/19.3.15>.
- Baar, T., Brettel, H., Ortiz Segovia, M.V. (2015). Towards gloss control in fine art reproduction. Proc.SPIE, 9398.
- Barrow, H. G., Tenenbaum, J., M. Hanson, A., Riseman, R. (1978). Recovering intrinsic scene characteristics from images. Computer vision systems. (pp. 3–26). New York: Academic Press.
- 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), 1-1.
- Boyaci, H., Doerschner, K., Snyder, J. L., & Maloney, L. T. (2006). Surface color perception in three-dimensional scenes. *Visual neuroscience*, 23(3-4), 311.
- Churma, M. E. (1994). Blue shadows: physical, physiological, and psychological causes. *Appl Opt*, 33(21), 4719-4722. doi:10.1364/AO.33.004719
- Fairchild, M.D. (2013). Color appearance models. John Wiley & Sons.
- Fleming, R. W., Torralba, A., Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, 4(9), 798-820.
- Giesel, M., Gegenfurtner, K.R. (2010) Color appearance of real objects varying in material, hue, and shape. *Journal of Vision*, 10(9), 10, 1–21.
- Granzier, J.J.M., Vergne, R., Gegenfurtner, K.R. (2014) The effects of surface gloss and roughness on color constancy for real 3-D objects. *Journal of Vision*, 14(2), 16, 1–20.

Honson, V., Huynh-Thu, Q., Arnison, M., Monaghan, D., Isherwood Z., Kim, J. (2020) Effects of shape, roughness and gloss on the perceived reflectance of coloured surfaces. *Frontiers in Psychology*, Accepted 02 March 2020.

Kim, J., Marlow, P., Anderson, B.L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision*, 11, 1 – 19.

Kim, J., Marlow, P., Anderson, B.L. (2012). The dark side of gloss. *Nature Neuroscience*, 15(11), 1590-1595.

Marlow, P.J., Kim, J., Anderson, B.L. (2012) The perception and misperception of specular reflectance. *Current Biology*, 22, 1909-1913.

Matsumoto, T., Fukuda, K., Uchikawa, K. (2016) Appearance of Gold, Silver and Copper Colors of Glossy Object Surface. *International Journal of Affective Engineering*, 15(3), 239-247.

Mooney, S.W.J., Anderson, B.L. (2014) Specular image structure modulates the perception of three-dimensional shape. *Current Biology*, 24, 2737-2742.

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G.D., Hurlbert, A. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PLoS One*, 9, e87989.

Radonjić, A., Pearce, B., Aston, S., Krieger, A., Dubin, H., Cottaris, N. P., Brainard, D. H., &

Schiller, F. & Gegenfurtner, K.R. (2016). Perception of saturation in natural scenes. *JOSA A*, 33(3), A194-A206.

Schiller, F., Valsecchi, M. & Gegenfurtner, K.R. (2018). An evaluation of different measures of color saturation. *Vision research*, 151, 117-134.

Schultz, S., Doerschner, K., & Maloney, L. T. (2006). Color constancy and hue scaling. *Journal of Vision*, 6(10), 10-10.

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

Toscani, M., Valsecchi, M. & Gegenfurtner, K. R. (2017). Lightness perception for matte and glossy complex shapes. *Vision Research*, 131, 82-95.

Toscani, M. & Valsecchi, M. (2019). Lightness discrimination depends more on bright rather than shaded regions of three-dimensional objects. *i-Perception*, 10(6), 2041669519884335.

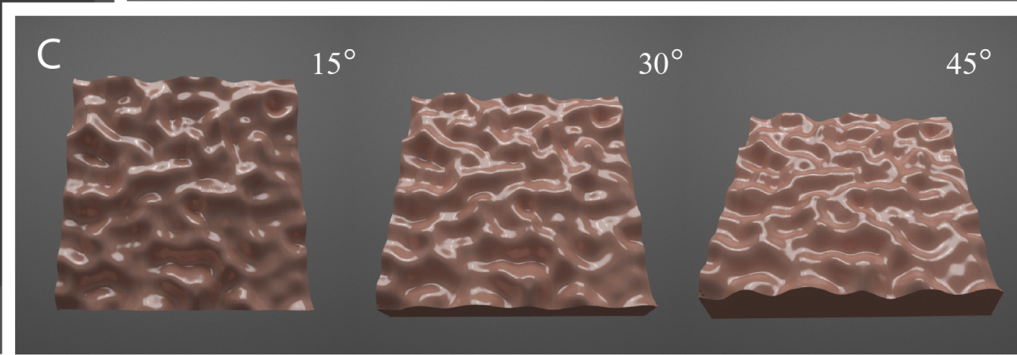
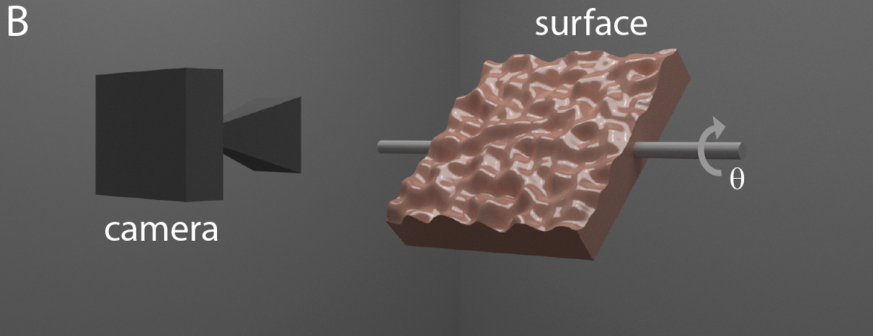
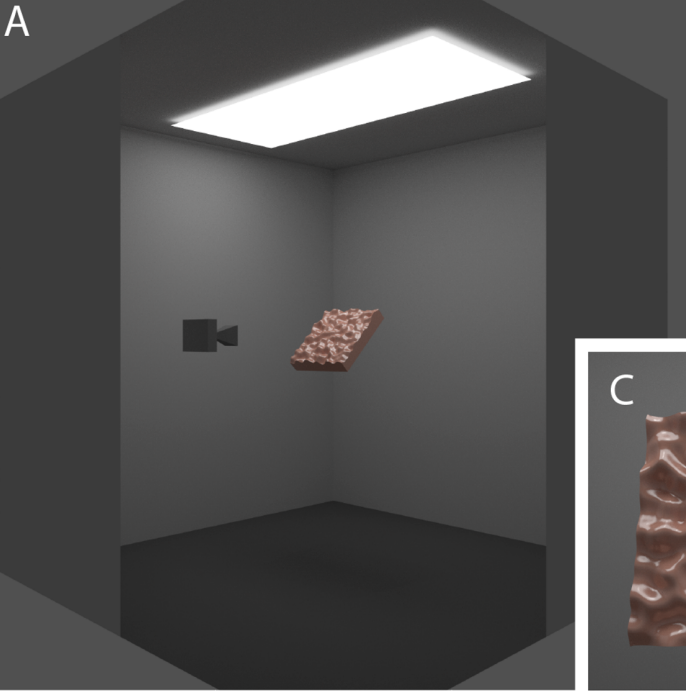
Weiss, D., Witzel, C., Gegenfurtner, K., 2017. Determinants of Colour Constancy and the Blue Bias. *Iperception* 8, 2041669517739635.

Winkler, A. D., Spillmann, L., Werner, J. S., & Webster, M. A. (2015). Asymmetries in blue-yellow color perception and in the color of 'the dress'. *Curr Biol*, 25(13), R547-548.
doi:10.1016/j.cub.2015.05.004

Xiao, B., Brainard, D. H. (2008). Surface gloss and color perception of 3D objects. *Visual Neuroscience*, 25, 371–385.

Xiao, B., Hurst, B., MacIntyre, L., Brainard, D.H. (2012) The color constancy of three-dimensional objects. *Journal of Vision*, 12 (4), 6-6.

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



Specular Roughness

0.010

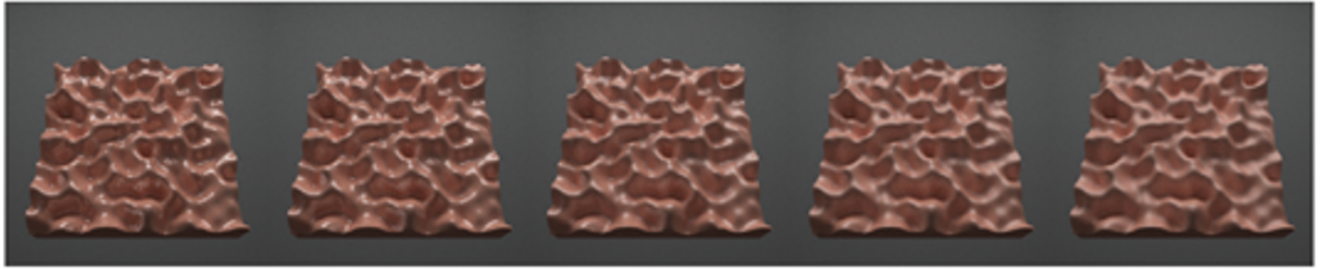
0.100

0.200

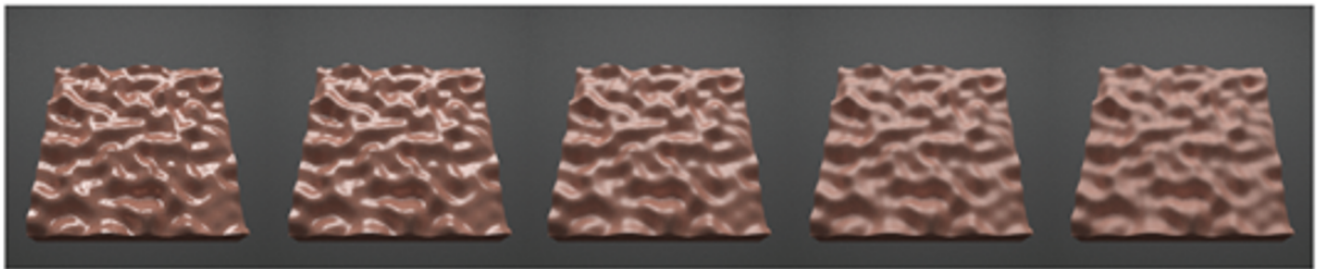
0.300

0.400

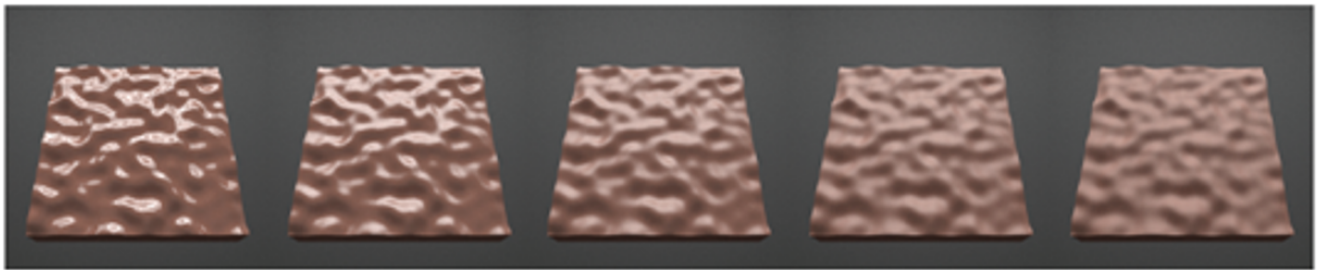
0.200



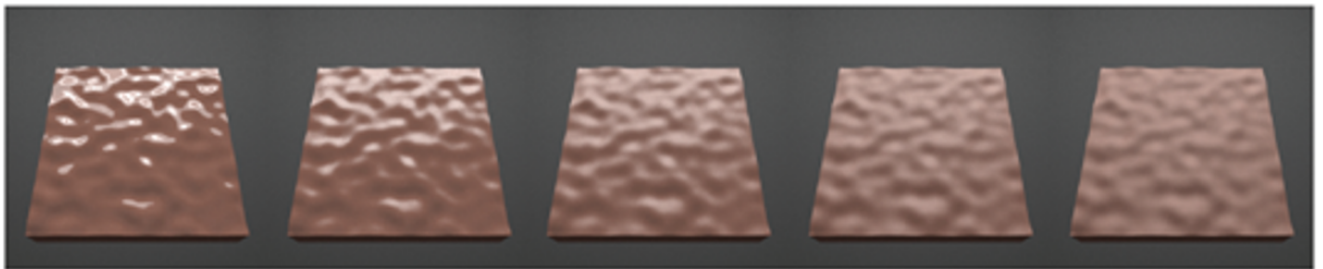
0.100



0.050



0.025



Mesoscopic relief height

Specular Roughness

0.010

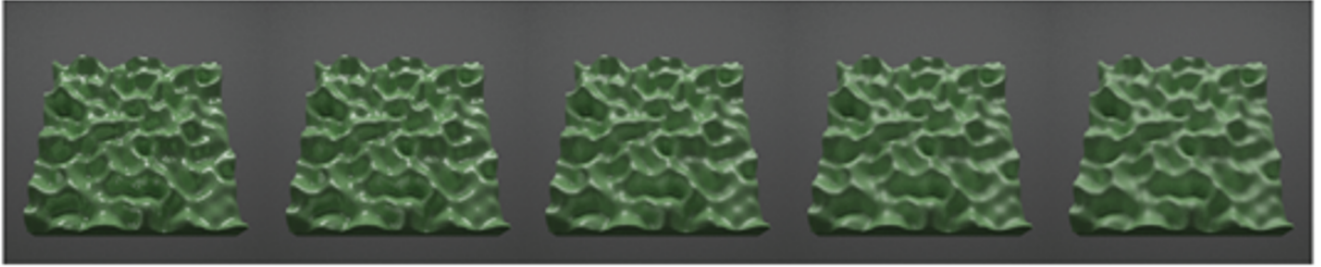
0.100

0.200

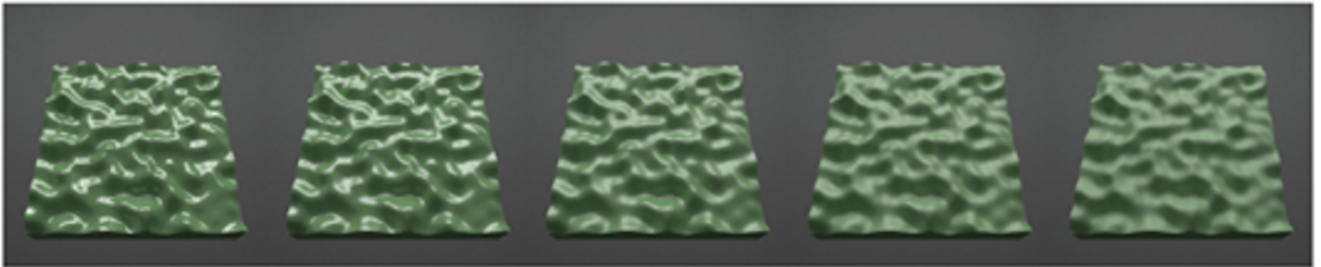
0.300

0.400

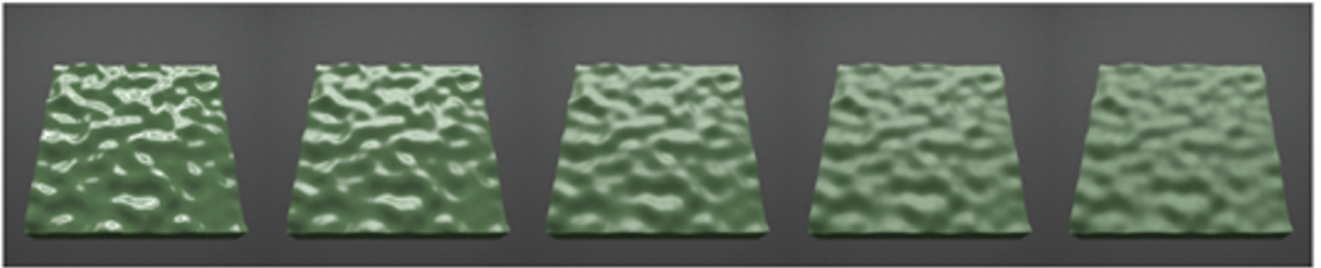
0.200



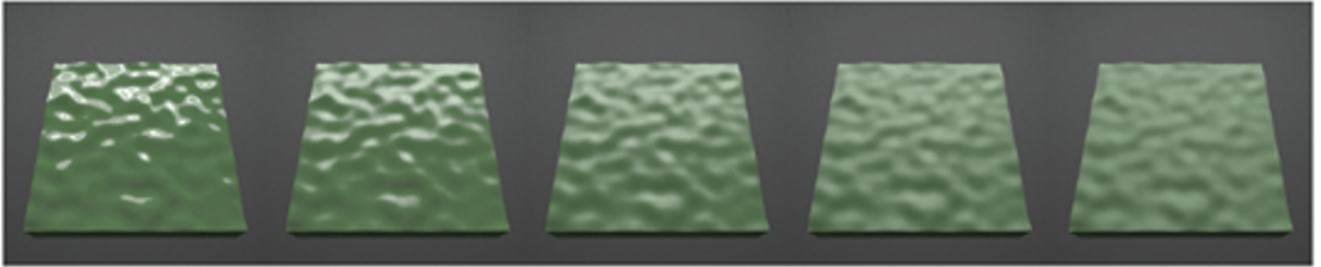
0.100



0.050



0.025



Mesoscopic relief height

Specular Roughness

0.010

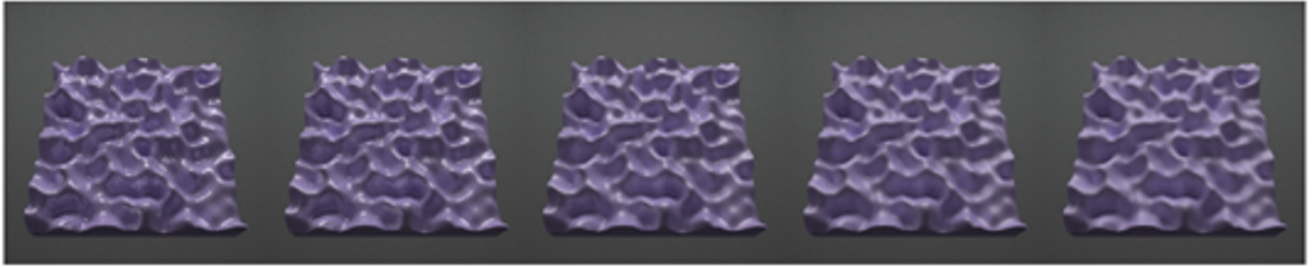
0.100

0.200

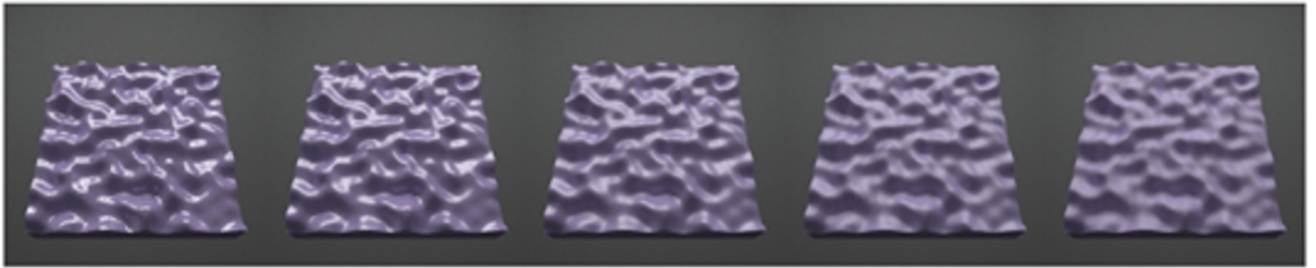
0.300

0.400

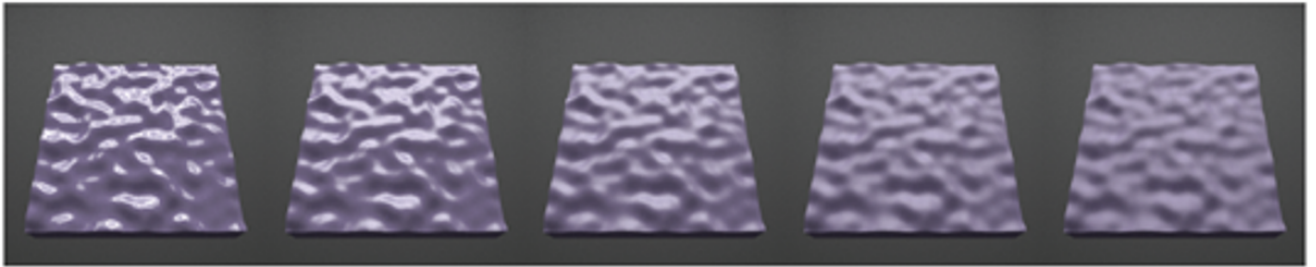
0.200



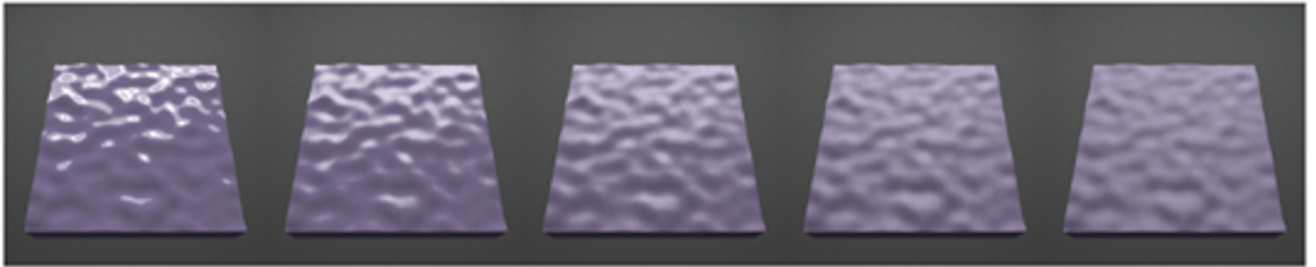
0.100



0.050

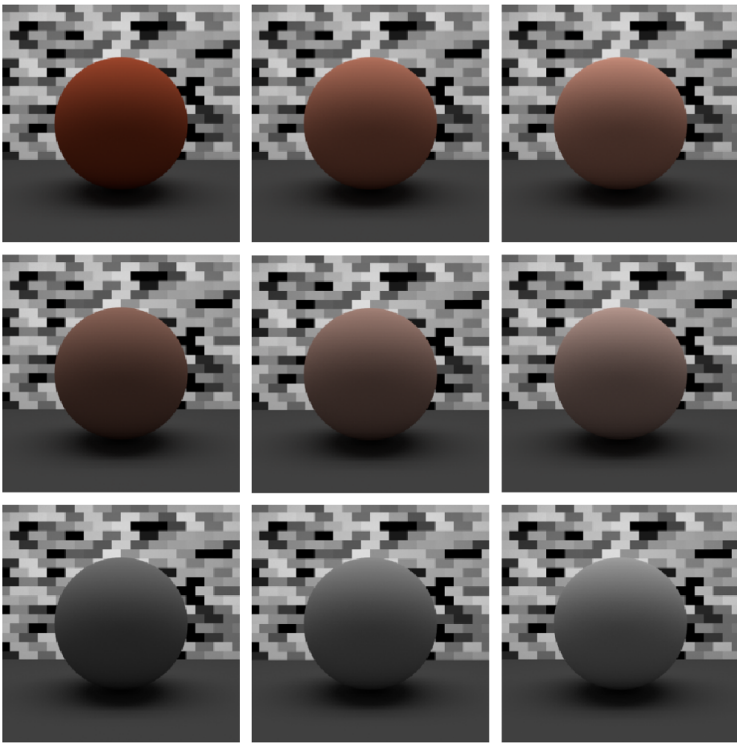


0.025

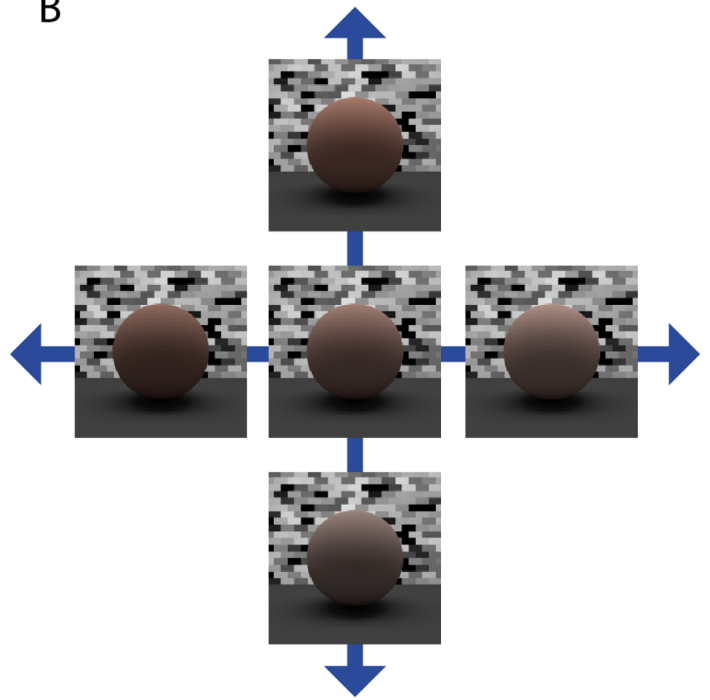


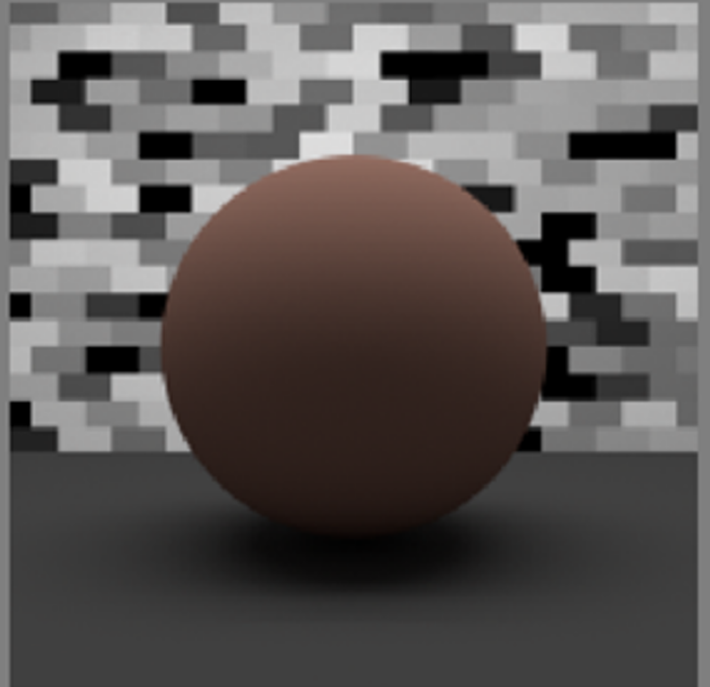
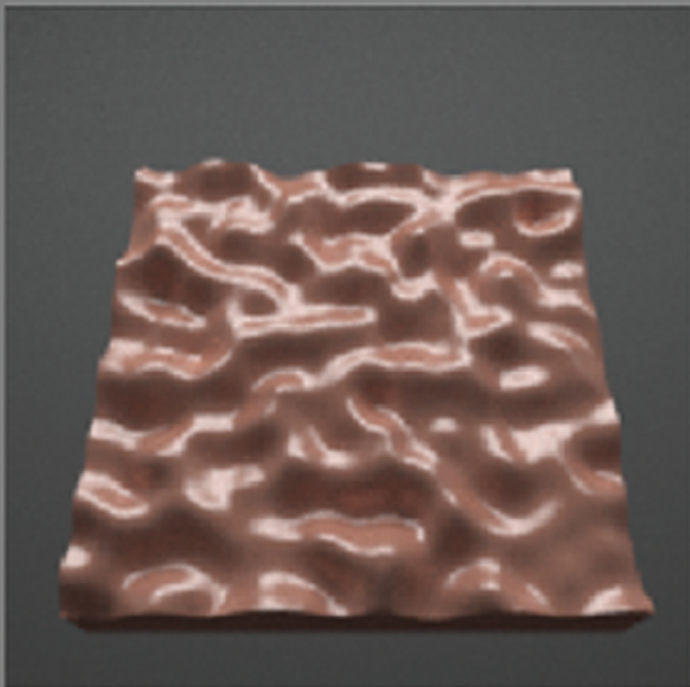
Mesoscopic relief height

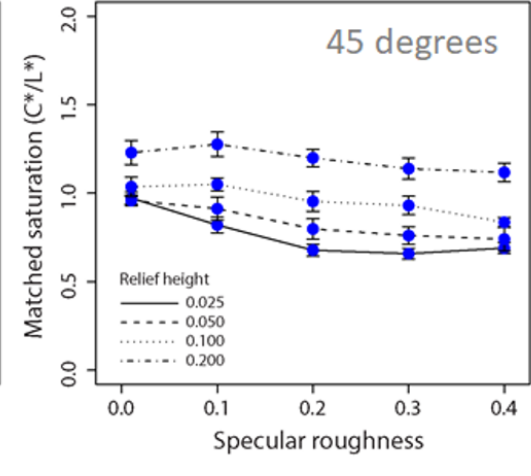
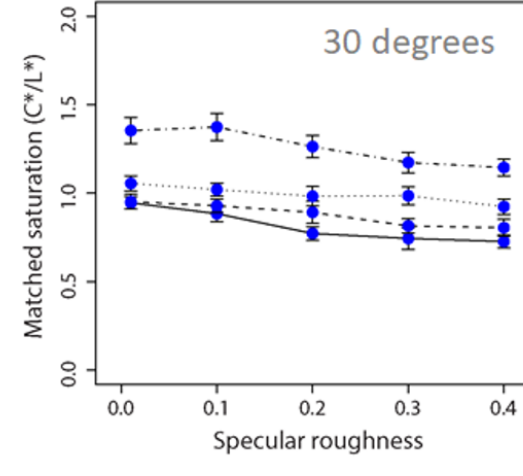
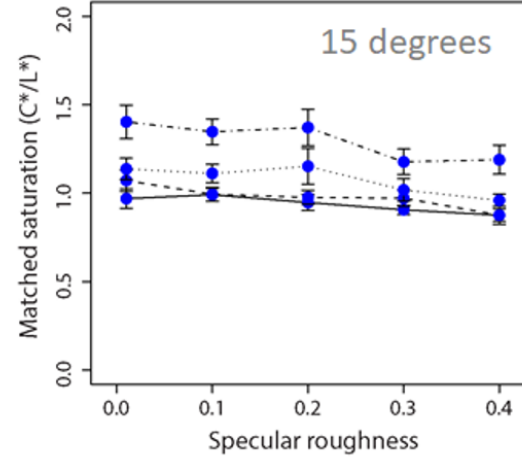
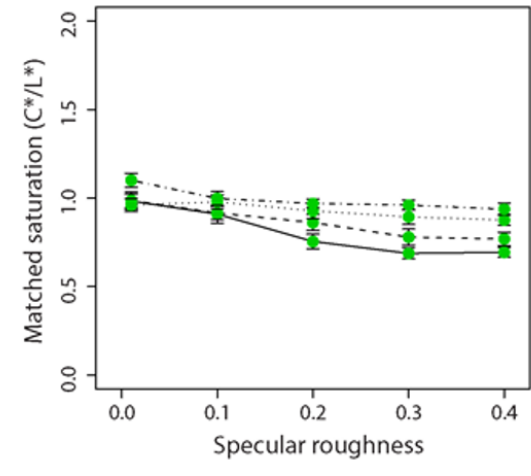
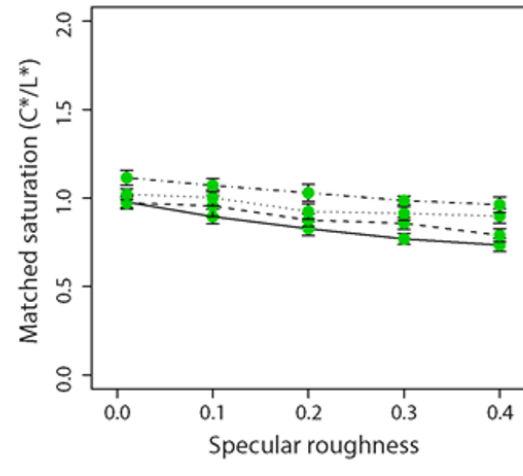
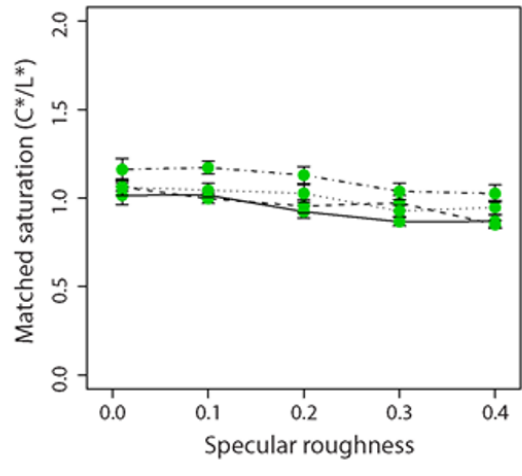
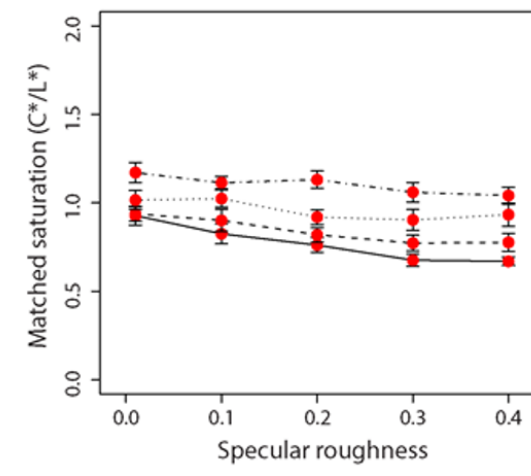
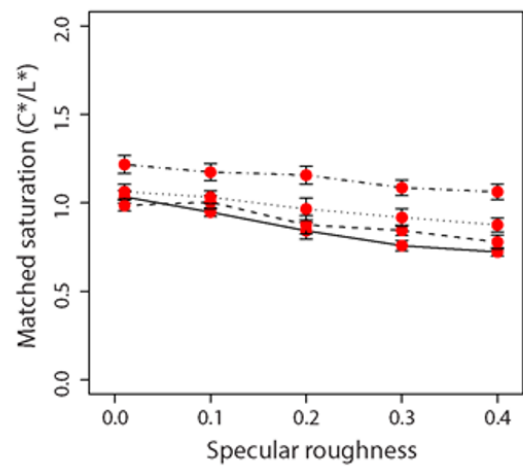
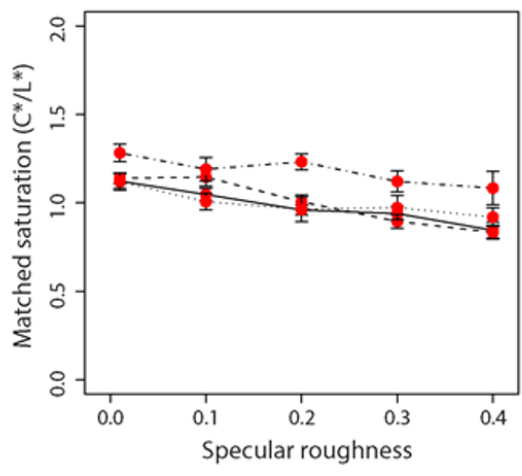
A

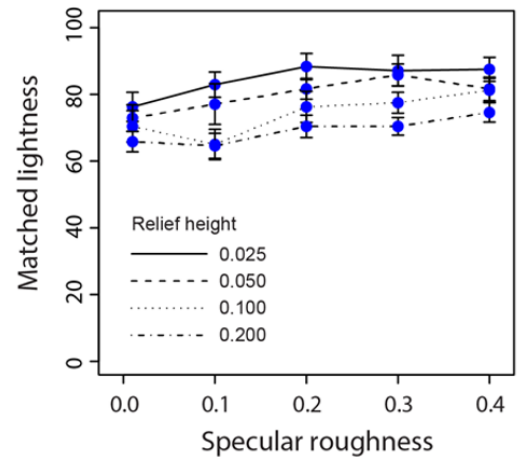
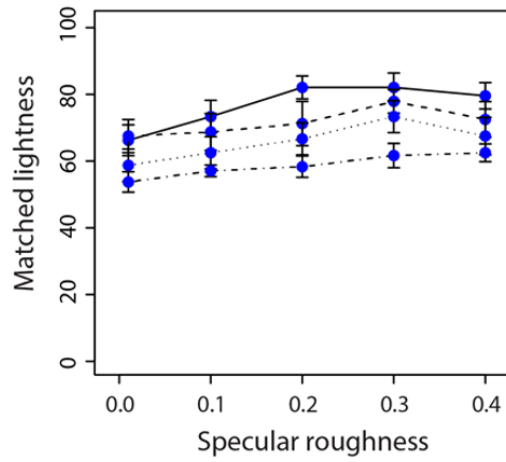
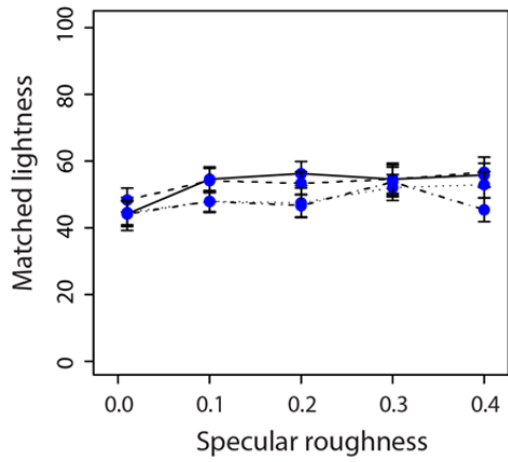
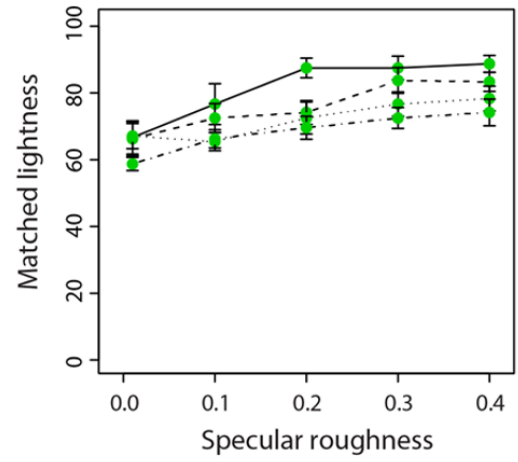
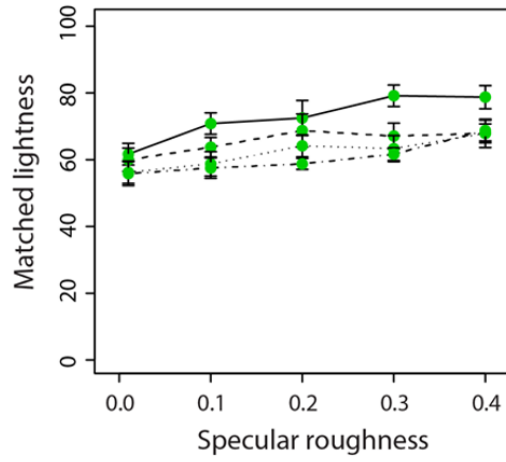
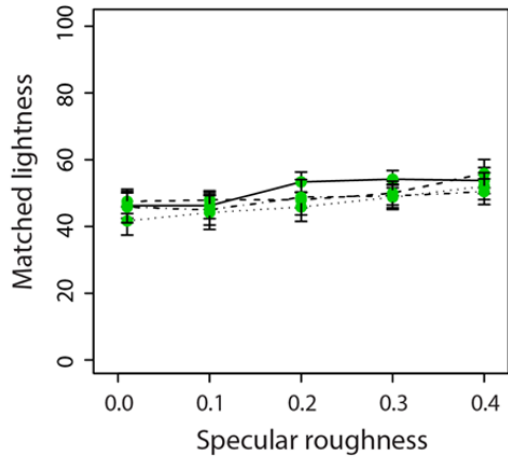
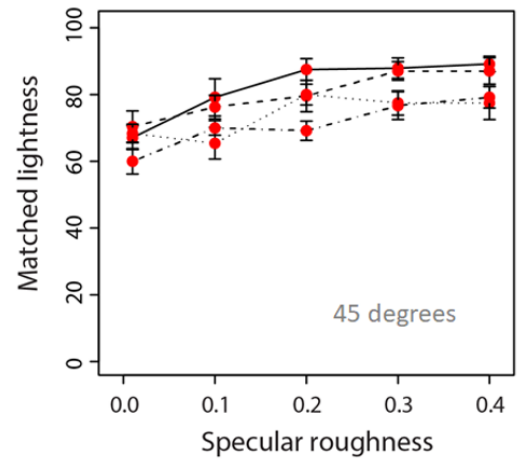
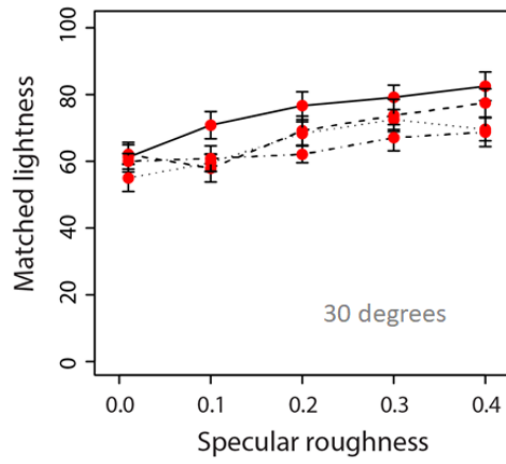
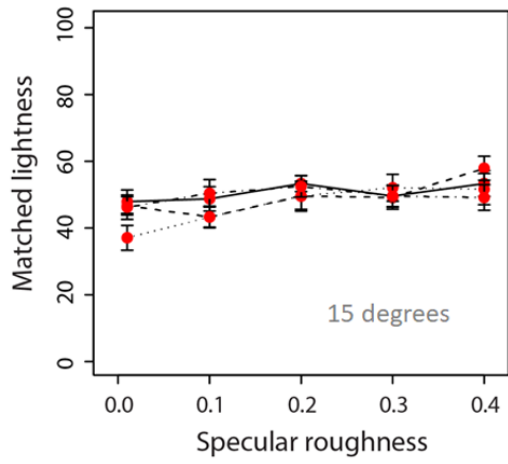


B









Relief height

- 0.025
- - - 0.050
- ... 0.100
- · - 0.200

