

# Interactions between Material Roughness, Perceived Gloss, and Color Appearance for 3D-Rendered Faces

Yuan Tian<sup>1</sup>, Mekides Assefa Abebe<sup>1</sup>, and Christopher A. Thorstenson<sup>1</sup>

<sup>1</sup>Munsell Color Science Laboratory, Rochester Institute of Technology, Rochester, NY 14623  
E-mail: yt6580@rit.edu

**Abstract.** Human faces are considered an important type of stimuli integral to social interaction. Faces occupy a substantial share of digital content, and their appearance can meaningfully impact how they are perceived and evaluated. In particular, past work has shown that facial color appearance can directly influence such perceptions. However, little is known regarding the perception of facial gloss and its influence on facial skin color appearance. The current work investigates how skin roughness influences perceived facial gloss and how these in turn affect facial color appearance for 3D rendered faces. Here, “roughness” refers to a parameter of the microfacet function modeling the microscopic surface. Two psychophysical experiments were conducted to model the interaction among skin roughness, perceived facial gloss, and perceived facial color appearance using varied facial skin tones. The results indicated an exponential relationship between skin roughness and perceived facial gloss, which was consistent across different skin tones. Additionally, gloss appearance influenced the perceived lightness of faces, a pattern not observed to the same extent among non-face objects included in the experiment. We expect that these results might partially be explained by discounting specular components for surface color perception to infer color attributes and by simultaneous contrast induced by a concentrated specular highlight. The current findings provide guidance for predicting visual appearance of face and non-face objects and will be useful for gloss and color reproduction of rendered digital faces.

**Keywords:** appearance, gloss, color, face

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[DOI: 10.2352/J.Percept.Imaging.2025.8.000502]

## 1. INTRODUCTION

Face perception is a crucial task that informs daily social interaction. We frequently process the visible characteristics of faces to form rapid and often accurate inferences about other people. These inferences meaningfully guide social interactions and include information about familiarity, age, sex, personality, and emotion [5, 60]. Facial processing is likely reinforced early on, beginning with infants recognizing caregivers [25]. Facial processing continues to become more specialized, likely because of the heightened ecological value of face information [26] relative to some other kinds of visual stimuli. Among facial characteristics that contribute to face processing, skin color and gloss play a significant role [23, 36, 46]. Past work has shown that dissociated neural pathways

are involved when processing faces separate from other types of objects and that color perception of faces is impacted more by memory color than other kinds of objects [18]. Several biophysical properties contribute to skin color and gloss, including hydration, sebum, erythema, carotenoids, hemoglobin, and melanin [31, 35, 52]. The concentration and distribution of these substances can vary in the skin due to many factors, including genetic characteristics, health, diet, emotion, and environmental factors. While facial processing occurs and is reinforced naturally, faces (including artificially rendered faces) are increasingly perceived among digital content, such as in gaming, graphics, and other visual media. Because face perception is crucial to social interaction among digital environments, it is important to better understand how material properties of rendered faces influence their gloss and consequent color appearance.

Perceived gloss comprises several attributes. For instance, refractive index, Fresnel reflection, and roughness in reflection models can vary different gloss attributes, including specular gloss, contrast gloss, sheen at grazing angles, haze, distinctness of reflected images, and absence of surface texture [10, 22]. For images, contrast, sharpness, and coverage of specular reflection are cues that predict gloss appearance [32]. The skin has particularly complex optical properties, which further complicate gloss reproduction. The outer layers of the skin are translucent and partially reflect incident light, and chromophores within the skin absorb and scatter light, causing it to exit in random directions. Therefore, skin reflection involves both surface and subsurface reflection [30]. Previous work has investigated gloss perception of faces at the macro- and mesoscale, where pores and wrinkles are visible, via surface and subsurface reflections, showing that the contrast of surface reflection intensity significantly improved skin gloss perception at the macroscale [53]. Most of the previous studies evaluating facial gloss appearance do not capture the complex optical properties of skin, including physical parameters like microscale skin roughness.

Apart from material characteristics, gloss and color are also affected by illumination and object shapes. Surfaces appear more glossy under more directional lighting or higher contrast lighting [34, 62]. Gloss constancy remains better under natural illumination [12]. Other cues such as color, motion, and disparity improve gloss constancy [54]. For shapes, perceived gloss first increases and then decreases

Received May 20, 2024; accepted for publication Jan. 30, 2025; published online July 04, 2025. Associate Editor: Benjamin Balas.

2575-8144/2025/8/000502/12/\$00.00

with bump magnitude [19, 32]. Spheres have lower discrimination accuracy for gloss than irregular shapes [50]. In color perception, illumination with higher luminance makes surfaces appear more colorful and have more contrast [9]. The stability of perceived color under spatially and spectrally inhomogeneous light fields could be improved with visual cues providing light field information [4]. Color appearance is stable with varying shapes and textures for real 3D objects. Perceived chroma is affected by the lightness of objects [14]. Previous studies also found that color appearance was affected by other surface properties [33, 57, 58].

The perceptual appearance of gloss interacts with color perception [21]. Lighter surfaces tend to appear less glossy than darker ones [11, 22, 47], and gloss tends to make surfaces appear more chromatic and darker due to the concentration of light at specular reflection [1, 3, 13]. Considering simultaneous contrast, a glossy background slightly decreases the perceived lightness of the center while enhancing its perceived gloss. Lighter backgrounds diminish both the perceived gloss and the perceived lightness [16]. Additionally, a steeper luminance gradient makes diffuse areas appear darker [29], and perceived lightness is impacted more by diffuse reflection than by spatial average reflection [49, 57]. Furthermore, specular roughness (surface roughness in the microfacet function) impacts perceived saturation and value more than relief height [20]. Finally, the appearance of gloss relies on visual cues that similarly affect color perception. Perceived color saturation and value depend on both specular coverage and surface orientation. Color appearance along these dimensions is impacted when it is difficult to separate diffuse from specular highlights, for example when the specular reflectance component is concentrated on relief-shaped surfaces [20].

Sensory and cognitive mechanisms are involved in the interaction between gloss and color appearance. The neural processing of color perception involves photoreceptors, retinal ganglion cells, optic nerve, lateral geniculate nucleus, and visual areas in the cortex [9]. In macaques, the processing of specular images is correlated with regions in the inferior temporal (IT) cortex [38]. Neural selectivity is associated with different aspects of gloss perception. The mechanism of combining local features to form gloss perception is represented in the neurons of the IT cortex [37]. Color processing occurs earlier than surface texture in cortical areas [6]. There are distinct pathways for processing color, texture, and form/shape. Moreover, the processing of surface properties is separated from shape/form [6, 7]. The perception of gloss is not entirely reliant on the same brain regions involved in processing color and texture [28]. Regarding cognitive mechanisms, previous research investigated how memory color affected color appearance and color constancy [9, 15, 51]. Previous research found that measured gloss increased with increasing sebum-moisture ratio in faces and reported thresholds of skin sebum-moisture ratio that observers perceived as shiny or greasy [24].

Altogether, both gloss and color appearance are informed by material optical properties as well as perceptual and cognitive mechanisms. Much research has been

conducted to investigate the interaction between gloss and color appearance, mostly focused on non-face stimuli. However, there is limited work towards modeling these interactions for facial stimuli, whose uniqueness lies in skin optical properties, sensory and cognitive mechanisms involved in face perception, and distinct neural pathways recruited for facial processing. The current work aims to fill this gap by investigating the interaction among material roughness, perceived gloss, and color appearance while perceiving faces. The study explores how roughness and consequent gloss appearance influence perceived lightness and chroma via physically based rendering to simulate skin composition and facial structure. The research also explores whether facial stimuli elicit distinct perceptual behaviors relative to non-face stimuli. Specifically, the current work aims to (1) investigate the influence of material roughness on perceived gloss for facial stimuli, (2) investigate the influence of perceived gloss on color appearance along the dimensions of lightness and chroma, and (3) investigate the extent to which this influence differs between facial and non-facial objects.

Two psychophysical experiments were conducted to achieve these aims. First, we investigated the influence of skin roughness on perceived gloss for rendered faces comprising different skin tones (Experiment 1). The results from Experiment 1 were then used to create a scale of perceived gloss as a function of skin roughness. Next, a second experiment was conducted to evaluate the influence of gloss appearance on summary color appearance for both face and non-face object (Experiment 2). This paper is an extension of work previously published in Ref. [48]. We expand on this work here by reporting on an additional experiment (Experiment 1) and offering extended analyses and discussion for the current work as a whole.

## 2. EXPERIMENT 1: GLOSS PERCEPTION

To address the first research question, we investigate the influence of material roughness on perceived gloss for rendered faces having various skin tones. An experiment is conducted to measure the perceived gloss of faces rendered with different material roughness values using a graphical ruler method. The experimental results are intended to create a perceptual scale for face gloss, which will also inform the stimulus selection for subsequent Experiment 2.

### 2.1 Methods

Roughness in the microfacet distribution function for computer graphics can control how smooth the surface is, contributing to how glossy the surface appears. This experiment is conducted to establish a quantitative relationship between roughness (a parameter in the material model) and perceived gloss (derived from observer judgments of the resulting models). The experiment employs a graphical ruler method where a linear perceptual scale is provided, and observers are tasked with populating the scale using a range of simultaneously presented stimuli. Each trial included 21 images for a single skin tone, resulting in 4 total trials for

4 skin tones. Each trial took approximately 10 minutes to complete.

### 2.1.1 Stimuli

The steps to create face images include generating geometric face shapes and rendering face appearance with 4 skin tones and 21 roughness levels. To render images, a physically based ray-tracing algorithm was used. The algorithm simulated a camera to generate an image film. From a point of the image film, a generated ray traveled and intersected with a surface. To obtain the color values at each image point, the surface normal, the geometric and radiometric information of the illuminant, and material optical properties were the inputs of the rendering engine. The geometric shape of the face providing surface normals was the initial model of FaceBuilder [27], which is a gender-neutral 3D face model without textures and shadings. To illuminate the environment, light-map rendering was utilized with a high dynamic range image shown in Figure 1. The light map had outdoor natural illumination and buildings in the scene. Thus, it simulates natural illumination in the real world, avoiding excessively high luminance that causes overexposure and overly low luminance that leads to a lack of contrast. The buildings in the scene introduced some texture into the images, which provided visual cues for gloss judgment [56, 61]. The material optical properties were also defined by a bidirectional scattering surface reflectance distribution function (BSSRDF). In this study, the skin BSSRDF was estimated with a given diffuse scattering coefficient, transmission coefficient (set at 1), microfacet roughness, refractive index (set at 1.55), and mean free path ([0.0013, 0.0009, 0.0006] for RGB channels). The mean free path, the mean distance of light path in a medium before scattering, is the reciprocal of attenuation coefficient, ranging from 0 to 1. As shown in Figure 2 and Table A.1, 21 levels of roughness were used to generate the experimental stimuli. The selected physically based rendering engine used the Beckmann–Spizzichino model as the microfacet distribution reflection function, and the roughness in the model represents the variation in slopes of microfacets in a unit area [41]. To represent the diffuse optical property of the material, the engine approximated spectral reflection with three-channel RGB values [41, 45]. Other than the material optical parameters, the camera and environmental parameters were kept unchanged.

Using the ray-tracing technique, four different skin type face models were generated. For each skin type, 21 face objects with varying roughness levels were created. The four skin tones were selected through clustering the Pantone SkinTone Guide [40] color sets (measured with an i1Pro spectrophotometer) by using the k-means clustering algorithm. We converted the reflectance of the skin set into a perceptual color space (CAM16 UCS) for clustering. From the resulting eight cluster centers, four representative and distinct colors were chosen. Then, the Jab of CAM16 UCS was converted into sRGB as the input of rendering (Figure 4). The rendered face models at various roughness levels and



Figure 1. Tone-mapped low dynamic range image of background. A high dynamic range image is used in rendering.

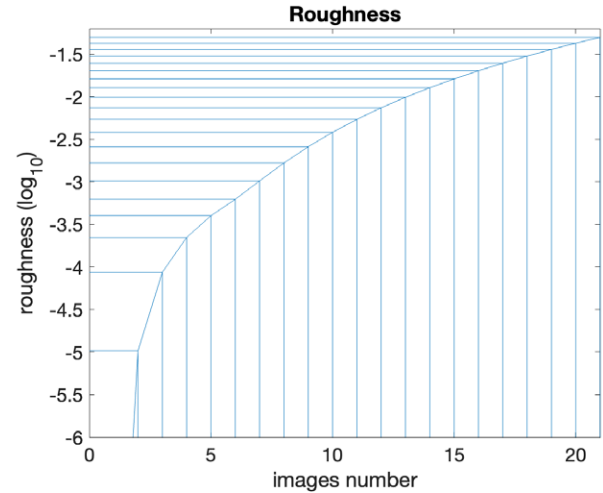


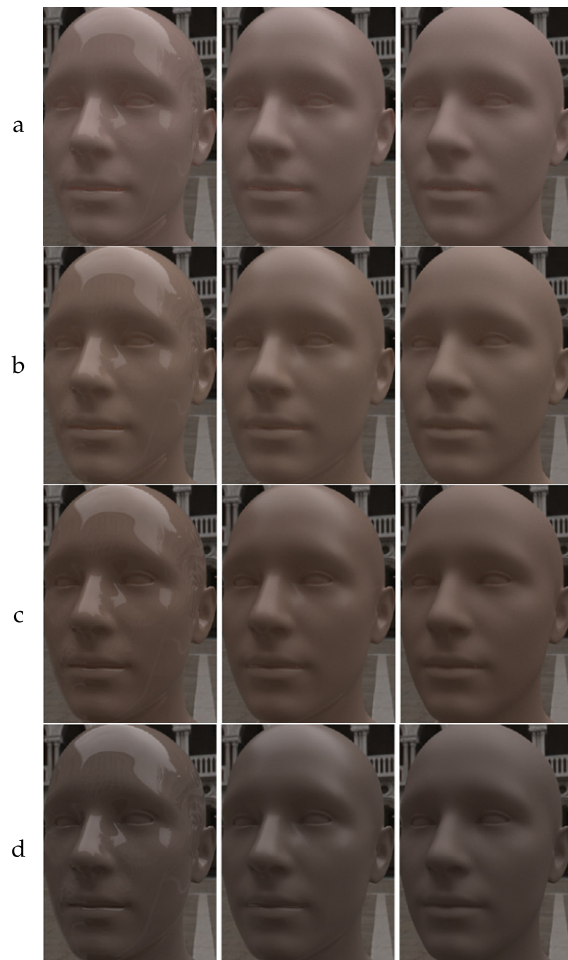
Figure 2. For each skin tone, there are 21 images with various roughness inputs to the renderer.

selected skin colors are shown in Figures 3 and 5, respectively. It should be noted that the colors in Fig. 5 are from the final renderings. The colors selected from the clustering results and used as the diffuse RGB parameters do not exactly match the final renderings, as the rendering process involves additional parameters beyond diffuse characteristics. The final stimulus colors in Fig. 5 are labeled “lightest,” “light,” “dark,” and “darkest” to indicate relative lightness differences but do not correspond to any specific skin typology (e.g., Fitzpatrick, Monk).

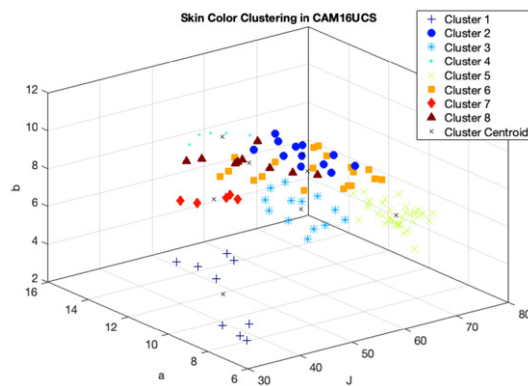
### 2.1.2 Procedure

Eleven observers participated in this experiment (all had normal or correct-to-normal visual acuity, and normal color vision; 4 males and 7 females of age from 20 to 35). Observers were presented with one skin color stimulus set at a time and asked to arrange the 21 roughness level images along a horizontal scale by dragging and dropping them with their mouse, aligning them with their perceptions of gloss relative to the left and right anchor images at the bottom of the experimental interface (see Figure 6). Observers were instructed that the relative distance between the images along the horizontal direction should be linear with respect to their perceived gloss (e.g., images could overlap if they were perceptually close, and images could have unequal spacing between them). The positions of the images along the vertical direction were not recorded, but observers could use the





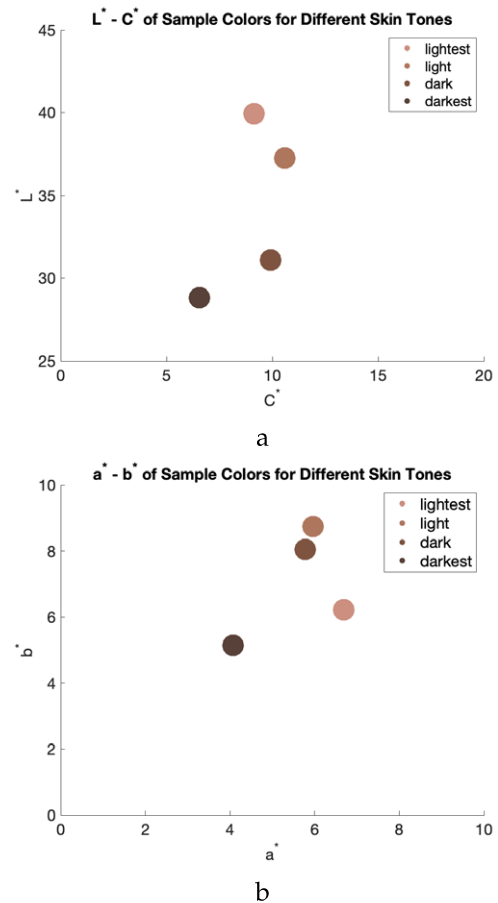
**Figure 3.** Examples of face stimuli used in Experiment 1. Examples include roughness values 0, 0.0054, and 0.05 (left to right) and skin tones “lightest,” “light,” “dark,” and “darkest” (top to bottom).



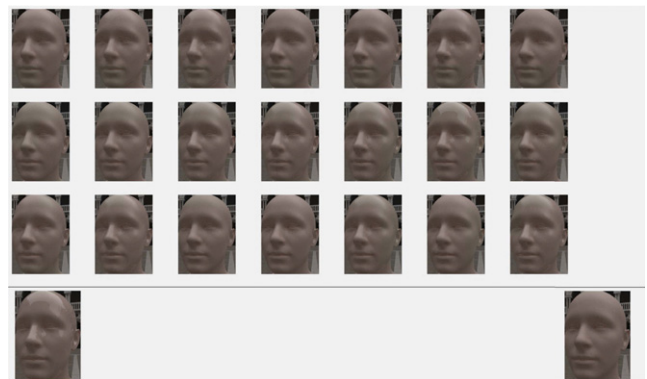
**Figure 4.** The k-means clustering of Pantone skin set in CAM16 UCS.

vertical dimension to avoid image overlap, which could be useful for comparing images simultaneously. The images used for left and right anchors were chosen to be those having minimum and maximum roughness values, respectively. The images are resized so that the field of view of each image subtends  $12^\circ$  in width and  $16^\circ$  in height,  $320 \times 440$  in pixel width and height. An example trial is shown in Fig. 6.

Corresponding to the four skin color types, each observer performed four trials. In each trial, 21 images with



**Figure 5.** CIELAB  $L^*$  (lightness)  $\times$   $C^*$  (chroma) and  $a^*$  (red-green)  $\times$   $b^*$  (yellow-blue) of skin colors used in Experiment 1, selected from a point sample of a diffuse region of the stimuli. Lightness values decrease across the four color types labeled “lightest,” “light,” “dark,” and “darkest,” respectively.



**Figure 6.** Experimental setup of Experiment 1. The 21 images shown on the top of the experimental interface are to be scaled by observers in between the left and right anchor images shown at the bottom left and right corners. The two anchor images have the lowest and highest roughness levels, representing the most glossy and matte samples. Note that for illustration purposes, this example is condensed horizontally, but observers completed the task with double the horizontal space.

various roughness levels were shown simultaneously. The starting position of the images before being moved to the scale were displayed in random order. Observers arranged

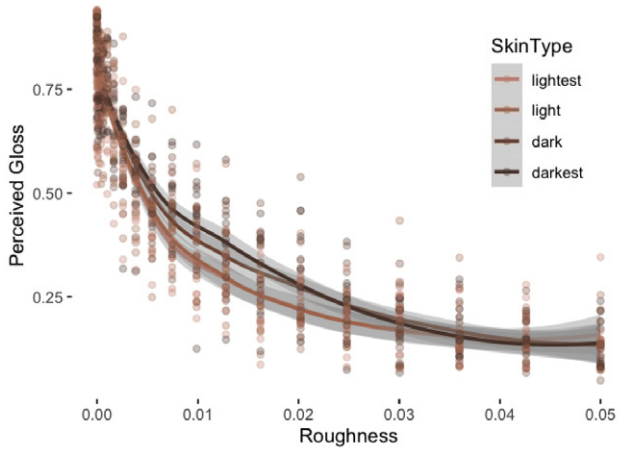


Figure 7. Summary of results from Experiment 1. Perceived gloss as a function of roughness and skin color. Individual (points) responses shown for each skin color and roughness value. Solid lines show loss fit predicted values and shaded error bars show 95% confidence intervals for fits.

each of the 21 images before moving on to the next trial. The order of trials for a given skin color was also randomized for different observers.

The spatial coordinates of each image placed by observers along the horizontal axis in the experimental interface were recorded and converted to gloss scales. The left anchor image representing the most glossy was assigned a value of 1, and the right anchor image representing the most matte was assigned a value of 0. The collected perceived gloss was used in this experiment as the dependent variable. Given the large number of images, two monitors were used (side by side) for the experiment. Both monitors were calibrated to the sRGB color profile with D65 white points (luminance:  $100 \text{ cd/m}^2$ ) with primaries R (0.64, 0.33), G (0.3, 0.6), and B (0.15, 0.06) in CIExy. Observers completed the experiment under a dim lighting condition, and their heads were at a fixed position facing the center of two monitors. Anti-glare monitors were used with the glare shield on. Observers were not able to see highlights and reflections from monitor surfaces under the dim lighting.

## 2.2 Results and Discussion

The observers' perceived gloss values recorded in this experiment are presented in Figure 7 for each skin color type and roughness levels. As expected, the glossiness of the faces reduced with increasing roughness.

A linear mixed-effect model was used on original data (Fig. 7) to evaluate the influence of roughness, skin color, and their interaction on perceived gloss. The results indicated a significant effect of roughness ( $F(1,10) = 880.99, p < 0.001$ ) such that perceived gloss decreased as roughness increased ( $B = -14.17, SE = 0.784$ ). There was no significant effect of skin color ( $F(3,8) = 1.46, p = 0.30$ ) or no significant effect of skin color and roughness interaction ( $F(3,8) = 0.13, p = 0.94$ ), suggesting that the effect of roughness on perceived gloss did not vary as a function of skin color.

One aim of this experiment is to establish a perceptual gloss scale to aid in stimulus selection for subsequent Experiment 2 detailed in the following section. As shown

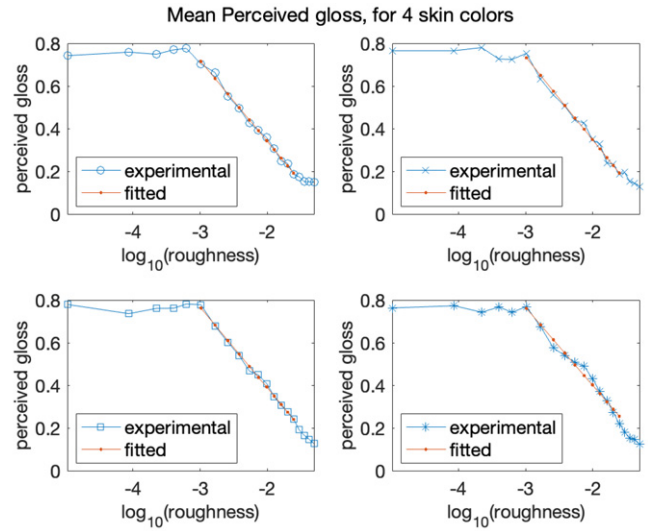


Figure 8. Linear fitting to the relationship between mean perceived gloss and logarithm of roughness. From top left to bottom right, four plots are for four color types: "lightest," "light," "dark," and "darkest."

in Fig. 7, perceived gloss responses follow an exponential function of roughness, suggesting a linear relationship in logarithmic space. Therefore, linear regression was implemented to fit perceived gloss to logarithmic roughness. As shown in Figure 8, the roughness values corresponding to perceived gloss from 0.2 to 0.7 remain linear. Therefore, a decision was made to fit a linear function within this range and with an interval of 0.08 roughness increment as shown in Fig. 8. The resulting function was then utilized to generate perceptually linear glossy samples for each of the four skin color types to be used in the next experiment.

## 3. EXPERIMENT 2: GLOSS AND COLOR APPEARANCE

The previous experiment demonstrated the exponential relationship between material roughness and perceived gloss in facial objects. The resulting linear regression model is employed in this experiment to generate perceptually spaced stimuli. The experiment is designed to investigate the influence of perceived gloss on the color appearance of faces. Previous research has shown that the color representation of faces involves different neural processing pathways and cognitive factors compared to other objects [18]. Accordingly, two additional non-face object shapes are included to determine if perceived gloss affects the color appearance of non-face objects in the same way. Similar to Experiment 1, the experiment is conducted for four different baseline skin color types.

### 3.1 Methods

Experiment 2 employs a "method of adjustment" experimental methodology. Observers were presented with a single 3D object on the left side of the screen and asked to match its color appearance by adjusting a uniform color patch displayed on the right side of the screen. An

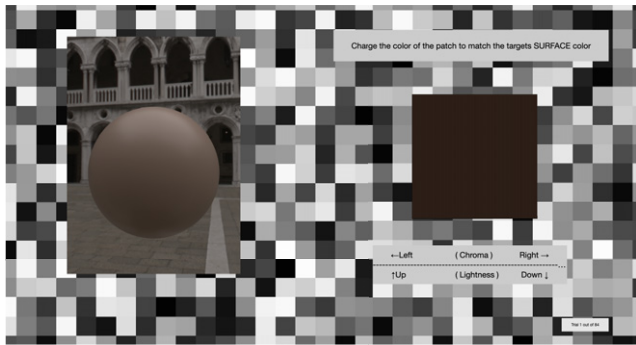


Figure 9. Experimental interface for an example trial from Experiment 2.

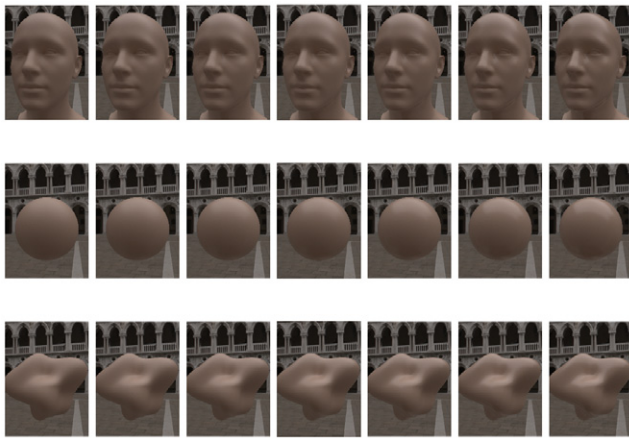


Figure 10. Example stimuli used in Experiment 2. Face, sphere, and blob-shaped objects with seven gloss levels for the “lightest” color type.

experimental interface for a sample trial is shown in Figure 9. The experiment included a total of 84 3D object stimuli, consisting of 3 different shapes, 4 baseline skin color types, and 7 perceptual gloss levels as illustrated in Figures 10 and 11. For the color appearance matching task, observers adjusted the lightness and chroma of the color patch. It was expected that varying levels of gloss would influence the perceived color appearance of the objects, particularly in terms of lightness. Additionally, we investigated whether this influence varied depending on the object’s shape and baseline color.

### 3.1.1 Stimuli

The experiment contains 84 stimuli. On the left side of the experimental interface, an object image would appear with a randomly selected shape (face, sphere, or blob), baseline color type (“lightest,” “light,” “dark,” “darkest”), and gloss level (seven total levels). The face-shaped objects were the same as those used in Experiment 1. We additionally chose to include the sphere-shaped object to represent a simple geometric shape with smooth, predictable material and light interactions. The blob-shaped object was chosen to represent a more complex shape with less predictable light interaction (edges and contours similar to a face) but with less familiarity

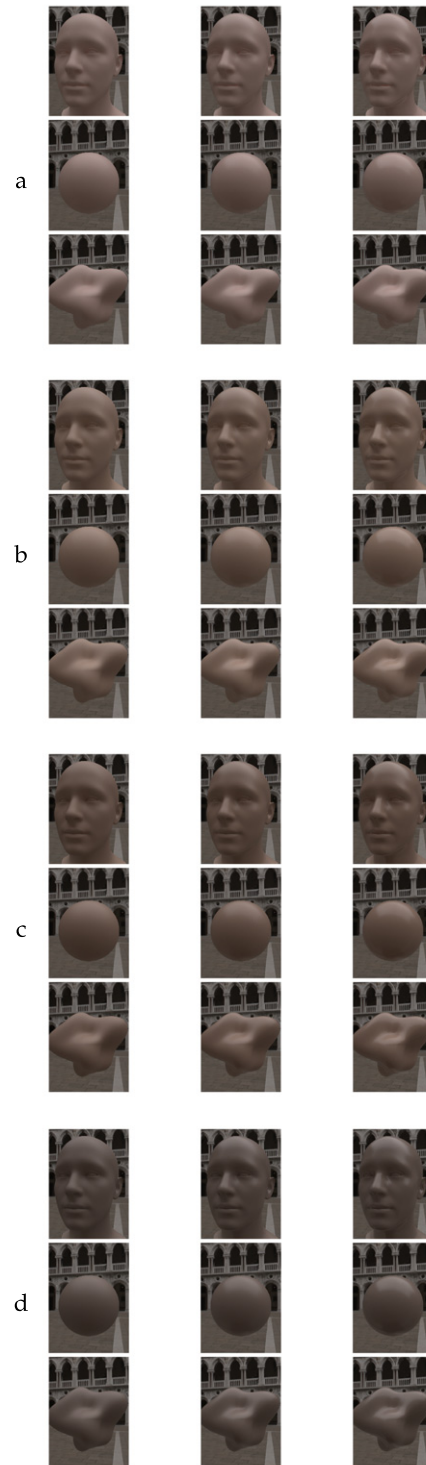


Figure 11. Example stimuli used in Experiment 2: from left to right—gloss levels of 1, 4, and 7; from top to bottom—face, sphere, and shapes; from (a) to (d)—“lightest,” “light,” “dark,” and “darkest” skin tones.

and social relevance than a face. The seven gloss levels of the objects were determined by the roughness function established in Experiment 1, so the roughness values used in Experiment 2 increase approximately linearly with perceived gloss (Figs. 7, 8, and 11). The gloss levels used represent ranges of perceived gloss from 0.2 to 0.7 with an interval



of 0.08. The four baseline color types were the same as those used as in Experiment 1.

### 3.1.2 Procedure

Fifteen observers participated in the experiment (age from 25 to 50; 9 females and 6 males). They were familiar with color space dimensions. All had normal or correct to normal visual acuity, and normal color vision. Eleven of them had previously participated in the first experiment. One object target image was displayed on the left side of the screen for each trial. This target image had a randomly selected shape type, baseline color type, and gloss level. The color patch was displayed on the right side of the screen simultaneously. Observers were asked to adjust the color of the patch to match the target's surface color. Observers were able to adjust the color patches by using a keyboard along the dimensions of lightness (CIELAB  $L^*$ , using up/down keys) and chroma (CIELAB  $C^*$ , using left/right keys). The base colors of the patch were those extracted from a point sample of the forehead region of the rendered face images with 0 roughness. These point samples were selected as they did not include either specular highlights or shadows, and they were judged by the authors as having the most diffuse area of the images. The full adjustment range was  $\pm 20 L^*$  and  $\pm 10 C^*$  of the base color of patches. The initial color was randomized. The color path could be adjusted to either increase or decrease along the  $L^*$  and  $C^*$  dimensions independently, and the step size of each adjustment (key press) was 1 unit for each dimension (Figs. 9 and 10). When observers were satisfied with their color match, their response was recorded. The responses included the final CIELAB  $L^*$ ,  $C^*$ ,  $a^*$ , and  $b^*$  values of the color patch as well as the  $L^*$  and  $C^*$  indices (representing the steps of  $L^*$  and  $C^*$  adjustment, respectively). Each observer completed a trial for each combination of target shape type (3), baseline color type (4), and gloss level (7) for a total of 84 trials.

## 3.2 Results and Discussion

The  $L^*$  and  $C^*$  indices were used in subsequent analyses. These indices represent a 1-unit step change in lightness and chroma adjustment made by observers, respectively. The analyses were conducted using these indices in order to normalize the adjustment responses so that relative changes along these dimensions could be compared across the different color types. The base, point-sample colors (described in Experiment 1 and used as starting points in Experiment 2) had an  $L^*$  index of 21 and a  $C^*$  index of 11, which are indices at the center of adjusting matrices. Therefore, indices above (or below) these values represent an increase (or decrease) in lightness and chroma adjustments made to the color patches to match the target object.

Linear mixed-effect models with random participant slopes and intercepts were used to evaluate the effects of gloss level (seven levels of a continuous predictor), shape type (three factors: faces, spheres, blobs), color type (four factors: "lightest," "light," "dark," "darkest"), and their interactions on adjusted  $L^*$  and  $C^*$  indices. We report the results

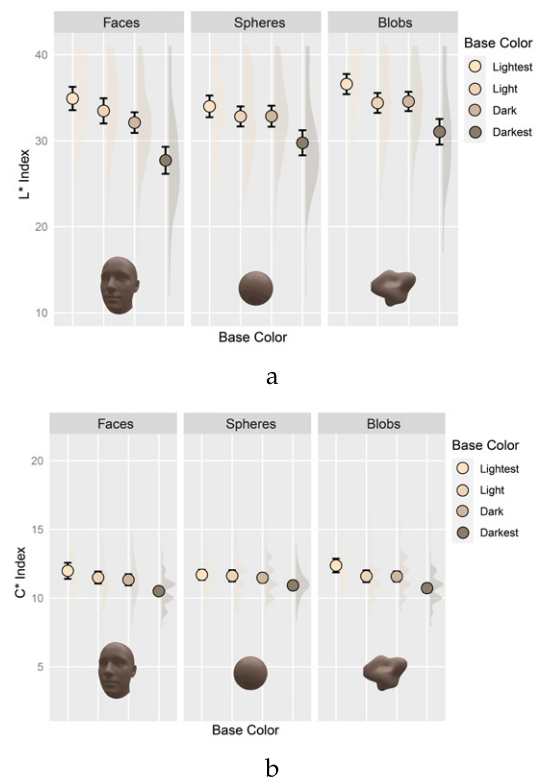


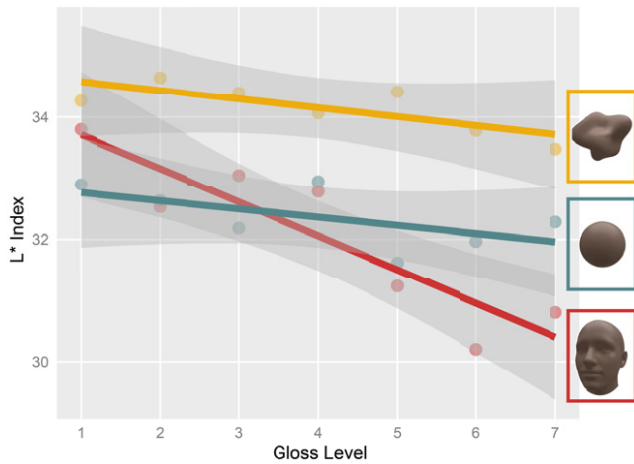
Figure 12. The mean (circles) of  $L^*$  and  $C^*$  indices for various shape and color types. Error bars represent standard error. Distributions represent the frequency of  $L^*$  or  $C^*$  index of individual responses.

from the statistical analyses separately for lightness and chroma adjustments. The reported values were returned by F statistics [2]. The responses of lightness and chroma adjustment were also predicted using linear models in R [55].

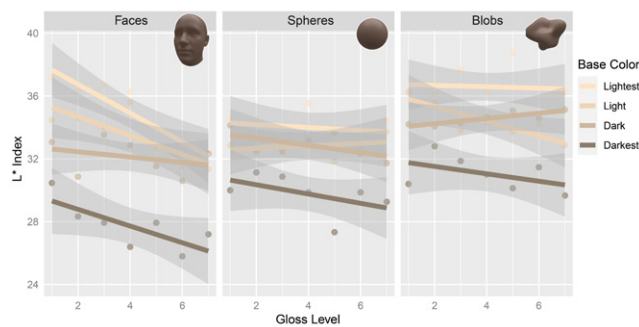
### 3.2.1 Lightness Adjustment

There was a significant main effect of gloss level on lightness adjustment ( $F(1,14) = 14.54$ ,  $p = 0.002$ ), indicating that perceived lightness generally decreased as gloss increased ( $B = -0.91$ ,  $SE = 0.24$ ). The main effect of shape type did not reach statistical significance ( $F(2,13) = 3.21$ ,  $p = 0.074$ ). However, followup  $t$ -tests indicated that faces ( $M = 32.1$ ,  $SE = 0.83$ ) were generally perceived as darker than blobs ( $M = 34.2$ ,  $SE = 0.66$ ;  $t(14) = 3.46$ ,  $p = 0.004$ ) but not spheres ( $M = 32.4$ ,  $SE = 0.81$ ;  $t(14) = 0.42$ ,  $p = 0.68$ ). There was a significant main effect of color type on lightness adjustments ( $F(3,12) = 10.07$ ,  $p = 0.001$ ). Generally, participants increased lightness more for "lightest" colors ( $M = 35.2$ ,  $SE = 0.65$ ), followed by "light" ( $M = 33.6$ ,  $SE = 0.66$ ), "dark" ( $M = 33.2$ ,  $SE = 0.70$ ), and "darkest" ( $M = 29.5$ ,  $SE = 1.03$ ) color types. All pairwise differences among the color types were significant ( $p < 0.005$ ) except for the difference between "light" and "dark" ( $t(14) = 1.21$ ,  $p = 0.248$ ); see Figures 12(a) and 13.

There was also a significant gloss level and shape type interaction ( $F(2,13) = 4.052$ ,  $p = 0.043$ ), indicating that the influence of gloss on lightness varied as a function of shape type. Perceived lightness decreased as gloss increased



**Figure 13.** Predicted  $L^*$  (lines) and mean of responses (dots) for three shapes as a function of gloss level. Shaded area represents 95% confidence intervals. Observers' perceptions of surface lightness generally decrease as gloss increases. This pattern is particularly pronounced for face-shaped objects.



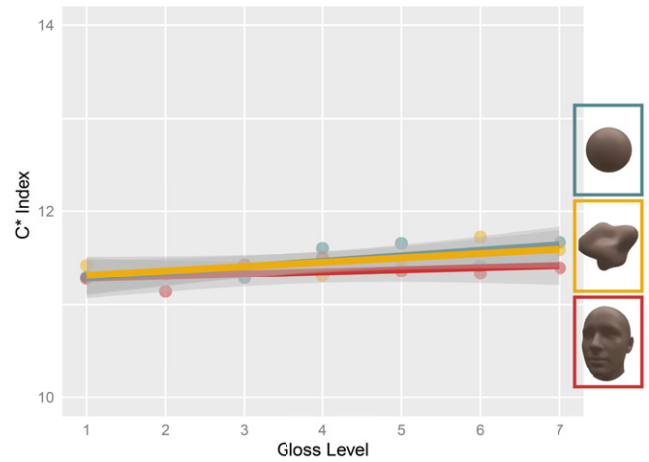
**Figure 14.** Predicted  $L^*$  index (lines) and mean responses (dots) for four color types and three shapes as a function of gloss level. On average, observers increase lightness more in order from “lightest” to “darkest” color types. These differences are larger for faces relative to other shapes.

for face-shaped objects ( $B = -0.55$ ,  $SE = 0.13$ ;  $t(14) = 4.29$ ,  $p < 0.001$ ), but this pattern was not statistically significant for spheres ( $B = -0.13$ ,  $SE = 0.12$ ;  $t(14) = 1.12$ ,  $p = 0.28$ ) or blobs ( $B = -0.15$ ,  $SE = 0.12$ ;  $t(14) = 1.21$ ,  $p = 0.25$ ); see Figures 13 and 14.

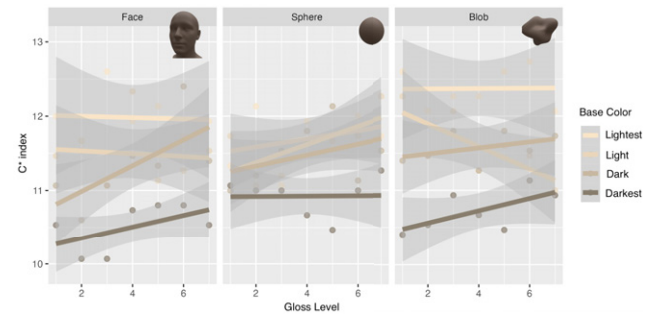
There were no statistically significant interactions between gloss level and color type ( $F(3,12) = 1.27$ ,  $p = 0.33$ ), between shape type and color type ( $F(6,9) = 1.32$ ,  $p = 0.34$ ), or the three-way interaction among them ( $F(6,9) = 0.74$ ,  $p = 0.63$ ); see Fig. 14.

### 3.2.2 Chroma Adjustment

Our analysis also shows that there is no significant effect of gloss level ( $F(1,14) = 0.01$ ,  $p = 0.92$ ) and shape type ( $F(2,13) = 1.32$ ,  $p = 0.30$ ) on chroma adjustments. But a significant main effect on chroma ( $F(3,12) = 3.56$ ,  $p = 0.047$ ) due to color type is observed. Generally, perceived chroma decreased from “lightest” ( $M = 12.0$ ,  $SE = 0.31$ ) to “light” ( $M = 11.6$ ,  $SE = 0.22$ ), “dark” ( $M = 11.5$ ,



**Figure 15.** Predicted  $C^*$  index (lines) and mean responses (dots) for three shapes. The shaded area represents 95% confidence intervals. Gloss did not significantly influence perceptions of object chroma.



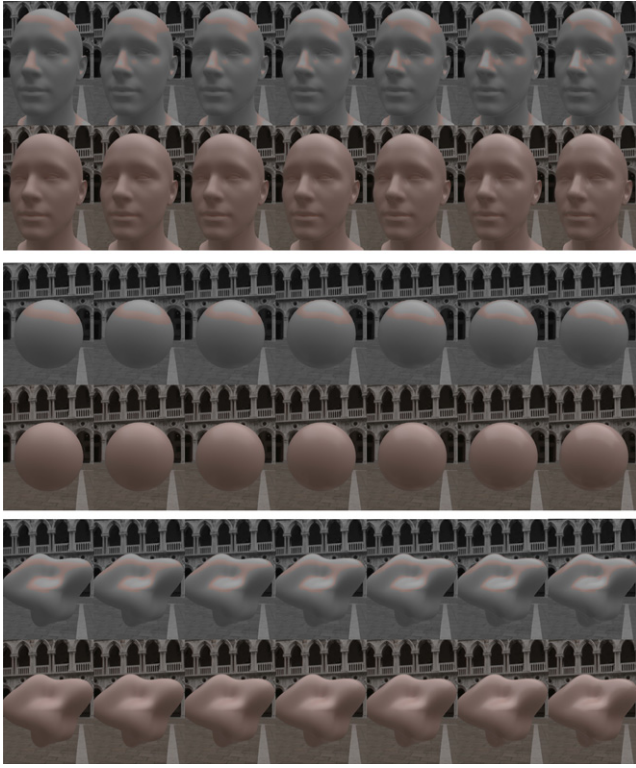
**Figure 16.** Predicted  $C^*$  index (lines) and mean responses (dots) for four color types and three shapes as a function of gloss level.

$SE = 0.21$ ), and “darkest” ( $M = 10.7$ ,  $SE = 0.14$ ) color types. All pairwise differences among the color types were significant ( $p < 0.013$ ) except for the difference between “light” and “dark” ( $t(14) = 1.10$ ,  $p = 0.288$ ); see Fig. 12(b). Moreover, no statistically significant two-way or three-way interactions were observed among these variables ( $p > 0.41$ ); see Figures 15, 12(b), and 16).

### 3.2.3 Image Color Analysis

In addition to gloss level analysis on color appearance, we explored whether observer color adjustments correspond to systematic patterns among regions of the images themselves. To do this, we identified the image pixels that comprised CIELAB values having small color differences ( $\Delta E \leq 5$ ) from the average CIELAB values of observers' responses. These pixels are shown in color with pixels outside this range colored gray in Figure 17. For each shape at each gloss level,  $\Delta E$  is computed and averaged across the four color types such that the images shown identify the objects' regions of interest rather than the precise color of each object. This analysis speculates about the strategy observers might have used to determine their representative color matches. However, it is worth noting that their decisions may not





**Figure 17.** Image analysis corresponding to observer color matches. Top rows: pixels containing color values within  $\Delta E \leq 5$  from mean observer responses are shown with all other pixels colored gray. Bottom rows: original images for reference. Images increase in gloss from left to right. Note that images for only one color type are shown, but these include spatial areas corresponding to  $\Delta E \leq 5$  averaged across all color types.

have been determined according to specific areas of the objects necessarily. Additional data such as from eye-tracking or pixel-selection methods in future work may be needed to confirm these strategies. Because perceived chroma was not impacted by changes in gloss, this section focuses on discussions regarding perceptions of object lightness. These analyses indicate that the image areas containing similar colors as observer responses are largely near the edges surrounding the specular highlight, suggesting that observers tend to use the brightest region of the objects (but excluding the highlight area) when judging colors of the surfaces. This strategy likely becomes easier as objects become more glossy as the specular highlight edge becomes more defined. This possibility is likely supported by the general decrease in lightness adjustments as gloss increased, as highlight regions are more likely to be avoided and considered distinct from perceptions of the surface color. The colored area is larger when gloss level decreases for spheres and blobs but not faces (Fig. 17). This indicates that the area representing surface colors by observers is larger with decreasing gloss, possibly because the boundary between diffuse reflection and specular highlight is less sharp for more matte surfaces. It has been previously shown that perceived lightness of objects is determined by diffuse reflection instead of mean lightness of objects [14, 57] and that observers judge

the lightness of an object according to the area neighboring the highlight region [14], which is supported by the current work.

#### 4. GENERAL DISCUSSION

Color appearance plays an important role in face processing and social perception. For digital graphics used in gaming and animation, faces are rendered by simulating interactions between light and material properties. This means that the appearance of digital faces is influenced by material properties like roughness, which alter perceptions of gloss and color. Therefore, it is crucial to better understand how material properties influence the color appearance of objects, and faces in particular, to facilitate interactions with digital social characters.

The current study evaluated the relationship between material roughness and perceived gloss in face rendering for four skin color types. The non-linear relationship did not account for the “other race effect,” which could be studied in the future. Additionally, we explored the influence of perceived gloss on color appearance for faces and non-face objects having different baseline colors. Notably, we found that perceptions of object lightness decreased as gloss increased but that this pattern was most evident for face-shaped objects, supporting the notion that processing of facial stimuli operates in different ways than other kinds of objects, most likely because faces are particularly familiar and contain important social information relative to abstract shapes. Previous research also found that gloss–lightness interactions were affected by shapes [17, 39] and gloss perception was affected by shapes at the micro-, meso-, and macroscale [19, 43, 44]. Therefore, it is useful to study the gloss–lightness interactions among shapes with face-like curvatures and ridges while not being perceived as faces. This would help confirm whether the observed decrease in perceived lightness of faces as gloss increases was due to the unique role faces play in daily life. We did not find evidence that perceptions of chroma were impacted by gloss, and therefore we focus the discussion on perceptions of lightness.

Previous research has demonstrated that the color of an object is influenced by perceived gloss and depends on the perceptual separability of the specular highlight from the diffuse area. People tend to ignore specular highlights when identifying surface body color [57]. It could also be the case that specular highlights are misattributed to diffuse areas when surface colors are lighter and more matte. Consequently, the perceived colors of matte and lighter surfaces are expected to have higher perceived lightness than more glossy surfaces [20]. These expectations were largely supported by the current findings, as perceived lightness generally decreased as gloss increased. Furthermore, our image color analyses indicate that the image areas containing similar colors as observer responses were largely near the edges surrounding the specular highlight, suggesting that observers tend to use the brightest region of the objects (while excluding the highlight area) when judging colors of the surfaces. This strategy likely becomes easier as

objects become more glossy, as the specular highlight edge becomes more defined. Conversely, highlight regions for more matte objects are more difficult to avoid, becoming less distinct from perceptions of the surface color. However, discounting the highlights can only partially explain this strategy, as it appears that the color matches still included some highlight information rather than darker, more diffuse areas of the objects. Another potential explanation may lie in simultaneous contrast, a perceptual phenomenon whereby very light colors seen adjacent to a target will make the target appear darker in comparison. Likewise, as objects become glossier the concentration of the specular highlight becomes more pronounced and appears lighter, which might make adjacent areas appear darker in comparison.

For the effect of color types on lightness adjustment, the significant differences between “lightest” and “light” and between “dark” and “darkest” indicate that observers exaggerate lightness adjustments with increasing skin luminance. However, there was no significant shape type by color type interaction on lightness adjustment, indicating that this exaggeration happened for all shapes. The lightness differences between “lightest” and “light” and between “dark” and “darkest” are smaller than the difference between “light” and “dark” (Fig. 5). However, the visual difference between “light” and “dark” is approximately equal to the base lightness difference. The other color type pairs (“lightest” and “light”; “dark” and “darkest”) have larger visual lightness differences than base lightness difference (Fig. 12b). A similar pattern occurs in chroma adjustment responses. The reasons for these visual difference patterns in lightness and chroma responses found in the study should be further investigated by future work.

The current study has some limitations that could be addressed by future work. Future work could simulate facial skin with more realism and naturalism using actual optical parameters of skin; for example, setting up wavelength-dependent refractive index separately for epidermis and dermis. Additionally, the face objects were rendered without realistic textures for skin and other properties of real faces (e.g., eyes, hair). This was done to facilitate comparisons with non-face objects that do not have these properties. However, it would be worthwhile to evaluate the current results in the context of more realistic rendered faces as stimuli. In addition, the degree of the viewing angle of face stimuli could affect gloss perception and gloss-color interaction, especially combining with real face textures such as pores and wrinkles. Although material roughness was the primary input used to generate perceptions of gloss, there are several other dimensions that would impact gloss and color appearance, which were not evaluated in the current work (e.g., distinctness of image and contrast gloss are functions of diffuse reflectance component, the energy of specular component, and the spread of specular lobe as described by Ward’s model [11]). In addition, the skin color selection was based on the measured reflectance of a commercial skin set. However, the printed color set may be less representative than real human skin measurement databases [59]. Furthermore, this approach summarizes

spectral data with RGB values to represent skin color, which certainly impacts how skin material would interact in a light simulation. For the lighting condition, we used outdoor urban lighting, but the illumination and scene dynamic range affected gloss appearance as found in previous research [8, 42]. Therefore, it is worthwhile to study the effect of gloss on color appearance under various lighting conditions in the future.

## 5. CONCLUSION

The present study conducted two psychophysical experiments to evaluate the relationships among material roughness, perceived gloss, and color appearance for faces and non-face objects having four baseline skin tones. The results indicated that perceived lightness (but not chroma) was influenced by gloss, which was most notable on face-shaped stimuli. We also speculate that people tend to judge surface color by largely (but not entirely) discounting specular highlights and that simultaneous contrast may also play a role in these judgments. The current work demonstrates the value of evaluating the influence of material properties on color appearance for digitally rendered objects and particularly highlights the need to consider face stimuli independent of other kinds of objects.

## APPENDIX

**Table A.1.** Roughness values used in Experiment 1.

Image number	Roughness	$\log_{10}$ (roughness)
1	0	n/a
2	1.04E−05	−4.99
3	8.66E−05	−4.06
4	2.23E−04	−3.65
5	4.01E−04	−3.40
6	6.25E−04	−3.20
7	1.03E−03	−2.99
8	1.67E−03	−2.78
9	2.60E−03	−2.59
10	3.83E−03	−2.42
11	5.43E−03	−2.27
12	7.42E−03	−2.13
13	9.86E−03	−2.01
14	1.28E−02	−1.89
15	1.62E−02	−1.79
16	2.02E−02	−1.69
17	2.48E−02	−1.61
18	3.00E−02	−1.52
19	3.60E−02	−1.44
20	4.26E−02	−1.37
21	5.00E−02	−1.30

## REFERENCES

- <sup>1</sup> T. Baar, S. Samadzadegan, H. Brettel, P. Urban, and M. V. O. Segovia, “Printing gloss effects in a 2.5 d system,” *Proc. SPIE* **9018**, 160–167 (2014).

- <sup>2</sup> D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *J. Stat. Software* **67**, 1–48 (2015).
- <sup>3</sup> H. Boyaci, K. Doerschner, and L. T. Maloney, "Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity," *J. Vision* **4**, 1–1 (2004).
- <sup>4</sup> H. Boyaci, K. Doerschner, J. L. Snyder, and L. T. Maloney, "Surface color perception in three-dimensional scenes," *Visual Neurosci.* **23**, 311–321 (2006).
- <sup>5</sup> V. Bruce and A. Young, "Understanding face recognition," *British J. Psychol.* **77**, 305–327 (1986).
- <sup>6</sup> J. S. Cant and M. A. Goodale, "Attention to form or surface properties modulates different regions of human occipitotemporal cortex," *Cereb. Cortex* **17**, 713–731 (2007).
- <sup>7</sup> C. Cavina-Pratesi, R. Kentridge, C. Heywood, and A. Milner, "Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia," *Cereb. Cortex* **20**, 2319–2332 (2010).
- <sup>8</sup> B. Chen, M. Piovarczy, C. Wang, H.-P. Seidel, P. Didyk, K. Myszkowski, and A. Serrano, "Gloss management for consistent reproduction of real and virtual objects," *SIGGRAPH Asia 2022 Conf. Papers* (ACM, New York, NY, USA, 2022), pp. 1–9.
- <sup>9</sup> M. D. Fairchild, *Color Appearance Models* (John Wiley & Sons, Ltd, Chichester, UK, 2013).
- <sup>10</sup> F. Faul, "The influence of Fresnel effects on gloss perception," *J. Vis.* **19**, 1–1 (2019).
- <sup>11</sup> J. A. Ferwerda, F. Pellacini, and D. P. Greenberg, "Psychophysically based model of surface gloss perception," *Human Vision Electron. Imaging VI* **4299**, 291–301 (2001).
- <sup>12</sup> R. W. Fleming, R. O. Dror, and E. H. Adelson, "Real-world illumination and the perception of surface reflectance properties," *J. Vision* **3**, 3–3 (2003).
- <sup>13</sup> R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflections and the perception of shape," *J. Vision* **4**, 10–10 (2004).
- <sup>14</sup> M. Giesel and K. R. Gegenfurtner, "Color appearance of real objects varying in material, hue, and shape," *J. Vision* **10**, 10–10 (2010).
- <sup>15</sup> J. J. Granzier and K. R. Gegenfurtner, "Effects of memory colour on colour constancy for unknown coloured objects," *i-Percept.* **3**, 190–215 (2012).
- <sup>16</sup> S. Hansmann-Roth and P. Mamassian, "A glossy simultaneous contrast: Conjoint measurements of gloss and lightness," *i-Percept.* **8**, 2041669516687770 (2017).
- <sup>17</sup> S. Hansmann-Roth, S. C. Pont, and P. Mamassian, "Contextual effects in human gloss perception," *Electron. Imaging* **30**, 1–7 (2018).
- <sup>18</sup> M. Hasantash, R. Lafer-Sousa, A. Afraz, and B. R. Conway, "Paradoxical impact of memory on color appearance of faces," *Nat. Commun.* **10**, 3010 (2019).
- <sup>19</sup> Y.-X. Ho, M. S. Landy, and L. T. Maloney, "Conjoint measurement of gloss and surface texture," *Psychol. Sci.* **19**, 196–204 (2008).
- <sup>20</sup> V. Honson, Q. Huynh-Thu, M. Arnison, D. Monaghan, Z. J. Isherwood, and J. Kim, "Effects of shape, roughness and gloss on the perceived reflectance of colored surfaces," *Front. Psychol.* **11** (2020).
- <sup>21</sup> R. W. G. Hunt and M. R. Pointer, *Measuring Colour* (John Wiley & Sons, Chichester, UK, 2011).
- <sup>22</sup> R. S. Hunter, "Methods of determining gloss," *J. Res. National Bureau of Standards* **18**, 19–281 (1937).
- <sup>23</sup> N. G. Jablonski and G. Chaplin, "The evolution of human skin coloration," *J. Human Evol.* **39**, 57–106 (2000).
- <sup>24</sup> Y. Jang, B. Kim, T. K. Moon, N. S. Kim, S. H. Lee, and H.-j. Lee, "The differentiation criteria between greyness and shininess on the face using mechanical evaluation and image," *J. Soc. Cosmet. Sci. Korea* **44**, 231–238 (2018).
- <sup>25</sup> M. H. Johnson, S. Dziurawiec, H. Ellis, and J. Morton, "Newborns' preferential tracking of face-like stimuli and its subsequent decline," *Cognition* **40**, 1–19 (1991).
- <sup>26</sup> N. Kanwisher and G. Yovel, "The fusiform face area: a cortical region specialized for the perception of faces," *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 2109–2128 (2006).
- <sup>27</sup> KeenTools, 2022 Facebuilder for blender. [Online]. Available: <https://keentools.io/products/facebuilder-for-blender>.
- <sup>28</sup> R. W. Kentridge, R. Thomson, and C. A. Heywood, "Glossiness perception can be mediated independently of cortical processing of colour or texture," *Cortex* **48**, 1244–1246 (2012).
- <sup>29</sup> J. Kim, P. J. Marlow, and B. L. Anderson, "The dark side of gloss," *Nat. Neurosci.* **15**, 1590–1595 (2012).
- <sup>30</sup> T. Lister, P. A. Wright, and P. H. Chappell, "Optical properties of human skin," *J. Biomed. Opt.* **17**, 090901 (2012).
- <sup>31</sup> L. Machková, D. Švadlák, and I. Dolečková, "A comprehensive in vivo study of Caucasian facial skin parameters on 442 women," *Arch. Dermatol. Res.* **310**, 691–699 (2018).
- <sup>32</sup> P. J. Marlow, J. Kim, and B. L. Anderson, "The perception and misperception of specular surface reflectance," *Curr. Biol.* **22**, 1909–1913 (2012).
- <sup>33</sup> T. Matsumoto, K. Fukuda, and K. Uchikawa, "Appearance of gold, silver and copper colors of glossy object surface," *Int. J. Affective Eng.* **15**, 239–247 (2016).
- <sup>34</sup> I. Motoyoshi and H. Matoba, "Variability in constancy of the perceived surface reflectance across different illumination statistics," *Vis. Res.* **53**, 30–39 (2012).
- <sup>35</sup> G. W. Nam, J. H. Baek, J. S. Koh, and J. K. Hwang, "The seasonal variation in skin hydration, sebum, scaliness, brightness and elasticity in Korean females," *Skin Res. Technol.* **21**, 1–8 (2015).
- <sup>36</sup> D. Nemrodov, M. Behrmann, M. Niemeier, N. Drobotenko, and A. Nestor, "Multimodal evidence on shape and surface information in individual face processing," *NeuroImage* **184**, 813–825 (2019). Available: <https://www.sciencedirect.com/science/article/pii/S1053811918319591>.
- <sup>37</sup> A. Nishio, N. Goda, and H. Komatsu, "P1-24: Neural representation of gloss in the macaque inferior temporal cortex," *i-Percept.* **3**, 638–638 (2012).
- <sup>38</sup> G. Okazawa, N. Goda, and H. Komatsu, "Selective responses to specular surfaces in the macaque visual cortex revealed by fMRI," *NeuroImage* **63**, 1321–1333 (2012).
- <sup>39</sup> M. Olkkonen and D. H. Brainard, "Joint effects of illumination geometry and object shape in the perception of surface reflectance," *i-Percept.* **2**, 1014–1034 (2011).
- <sup>40</sup> Pantone LLC, *Skintone Guide: Revealing the new PANTONE SkinTone Guide*. Available: <https://www.pantone.com/articles/product-spotlight/skintone-guide-revealing-the-new-pantone-skintone-guide>.
- <sup>41</sup> M. Pharr, W. Jakob, and G. Humphreys, *Physically Based Rendering: From Theory to Implementation* (MIT Press, Cambridge, MA, 2023).
- <sup>42</sup> J. B. Phillips, J. A. Ferwerda, and S. Luka, "Effects of image dynamic range on apparent surface gloss," *Proc. IS&T/SID CIC17: Seventeenth Color Imaging Conf.* (IS&T, Springfield, VA, 2009), Vol. 17, pp. 193–197.
- <sup>43</sup> L. Qi, M. J. Chantler, J. P. Siebert, and J. Dong, "How mesoscale and microscale roughness affect perceived gloss," *Proc. 3rd Int'l. Conf. on Appearance* (Lulu Press, Edinburgh, UK, 2012), pp. 48–51.
- <sup>44</sup> A. Serrano, B. Chen, C. Wang, M. Piovarczy, H.-P. Seidel, P. Didyk, and K. Myszkowski, "The effect of shape and illumination on material perception: model and applications," *ACM Trans. Graphics (TOG)* **40**, 1–16 (2021).
- <sup>45</sup> B. Smits, "An RGB-to-spectrum conversion for reflectances," *J. Graphics Tools* **4**, 11–22 (1999).
- <sup>46</sup> I. D. Stephen, V. Coetzee, M. Law Smith, and D. I. Perrett, "Skin blood perfusion and oxygenation colour affect perceived human health," *PloS One* **4**, e5083 (2009).
- <sup>47</sup> K. R. Storrs, B. L. Anderson, and R. W. Fleming, "Unsupervised learning predicts human perception and misperception of gloss," *Nat. Human Behav.* **5**, 1402–1417 (2021).
- <sup>48</sup> Y. Tian, M. A. Abebe, and C. A. Thorstenson, "The influence of material roughness on perceived gloss and color appearance of graphical generated faces," *Color Imaging Conf.* **32**, 173–173 (2024) [Online]. Available: <https://library.imaging.org/cic/articles/32/1/30>.
- <sup>49</sup> J. T. Todd, J. F. Norman, and E. Mingolla, "Lightness constancy in the presence of specular highlights," *Psychol. Sci.* **15**, 33–39 (2004).
- <sup>50</sup> P. Vangorp, J. Laurijssen, and P. Dutré, "The influence of shape on the perception of material reflectance," *ACM Trans. Graphics* **26**, 77 (2007).
- <sup>51</sup> M. Vurro, Y. Ling, and A. C. Hurlbert, "Memory color of natural familiar objects: Effects of surface texture and 3-D shape," *J. Vision* **13**, 20–20 (2013).
- <sup>52</sup> C. V. Wa and H. I. Maibach, "Mapping the human face: biophysical properties," *Skin Res. Technol.* **16**, 38–54 (2010).



- <sup>53</sup> J. Wang, C. Kuesten, J. Mayne, G. Majmudar, and T. N. Pappas, "Human skin gloss perception based on texture statistics," *IEEE Trans. Image Process.* **30**, 3610–3622 (2021).
- <sup>54</sup> G. Wendt, F. Faul, V. Ekroll, and R. Mausfeld, "Disparity, motion, and color information improve gloss constancy performance," *J. Vision* **10**, 7–7 (2010).
- <sup>55</sup> H. Wickham, W. Chang, L. Henry, T. L. Pedersen, K. Takahashi, C. Wilke, K. Woo, H. Yutani, D. Dunnington, and T. van den Brand, 2024 Smoothed conditional means. Accessed: 2024-05-30. [Online]. Available: [https://ggplot2.tidyverse.org/reference/geom\\_smooth.html](https://ggplot2.tidyverse.org/reference/geom_smooth.html).
- <sup>56</sup> L. Xia, S. C. Pont, and I. Heynderickx, "Light diffuseness metric, part 2: Describing, measuring and visualising the light flow and diffuseness in three-dimensional spaces," *Light. Res. Technol.* **49**, 428–445 (2017).
- <sup>57</sup> B. Xiao and D. H. Brainard, "Surface gloss and color perception of 3D objects," *Vis. Neurosci.* **25**, 371–385 (2008).
- <sup>58</sup> B. Xiao, B. Hurst, L. MacIntyre, and D. H. Brainard, "The color constancy of three-dimensional objects," *J. Vision* **12**, 6–6 (2012).
- <sup>59</sup> K. Xiao, J. M. Yates, F. Zardawi, S. Sueeprasan, N. Liao, L. Gill, C. Li, and S. Wuerger, "Characterising the variations in ethnic skin colours: a new calibrated data base for human skin," *Skin Res. Technol.* **23**, 21–29 (2017).
- <sup>60</sup> L. A. Zebrowitz and J. M. Montepare, "Social psychological face perception: Why appearance matters," *Social Personality Psychol. Compass* **2**, 1497–1517 (2008).
- <sup>61</sup> F. Zhang, H. de Ridder, P. Barla, and S. Pont, "A systematic approach to testing and predicting light-material interactions," *J. Vision* **19**, 11–11 (2019).
- <sup>62</sup> F. Zhang, H. de Ridder, P. Barla, and S. Pont, "Effects of light map orientation and shape on the visual perception of canonical materials," *J. Vision* **20**, 13–13 (2020).