

SUPPLEMENTARY MATERIAL: The effect of shape and illumination on material perception: model and applications

ANA SERRANO, Max-Planck-Institut für Informatik, Germany and Centro Universitario de la Defensa, Spain

BIN CHEN, Max-Planck-Institut für Informatik, Germany

CHAO WANG, Max-Planck-Institut für Informatik, Germany

MICHAL PIOVARČI, Institute of Science and Technology Austria, Austria

HANS-PETER SEIDEL, Max-Planck-Institut für Informatik, Germany

PIOTR DIDYK, Università della Svizzera Italiana, Switzerland

KAROL MYSZKOWSKI, Max-Planck-Institut für Informatik, Germany

This document offers additional information and details on the following topics:

- (S1) Additional details on the experimental design (Sec. 3 in the main paper).
- (S2) Additional results from the statistical analysis (Sec. 4 in the main paper).
- (S3) Perceptual embeddings for different sets of geometries and illuminations, featuring different combinations of our attributes (Sec. 5 in the main paper).
- (S4) Additional figures and data related to the statistical analysis (Sec. 4 in the main paper).
- (S5) Additional BRDF editing results (Sec. 6 in the main paper).

S 1. ADDITIONAL DETAILS ON EXPERIMENTAL DESIGN

SIGN

Our goal is to characterize the perceived appearance of a wide variety of materials under different geometries and illuminations for which we have rendered 42,120 different images covering different materials, illuminations, and geometries. We describe in this section our stimuli and procedure for crowdsourcing perceptual data.

1.1 Stimuli

Materials We include 520 measured BRDFs, mostly focusing on isotropic appearances (389 BRDFs). We also include (131) anisotropic materials for further analysis. We initially include 100 materials from the MERL dataset [Matusik et al. 2003], 62 from the RGL dataset [Dupuy and Jakob 2018] (from which 11 are anisotropic), and 150 from the UTIA [Filip and Vávra 2014] dataset. From the UTIA dataset, we select the 50 materials with higher anisotropy effects and compute their isotropic equivalents [Filip 2015], in order to allow comparisons between perceived attributes with and without anisotropy. Then, in order to complete our dataset, we use the work of Sun et al. [2018] to perform edits over MERL BRDFs in a compact

Authors' addresses: Ana Serrano, anase@unizar.es, Max-Planck-Institut für Informatik, Saarbrücken, Germany, Centro Universitario de la Defensa, Zaragoza, Spain; Bin Chen, binchen@mpi-inf.mpg.de, Max-Planck-Institut für Informatik, Saarbrücken, Germany; Chao Wang, chaowang@mpi-inf.mpg.de, Max-Planck-Institut für Informatik, Saarbrücken, Germany; Michal Piovarči, michael.piovarci@ist.ac.at, Institute of Science and Technology Austria, Klosterneuburg, Austria; Hans-Peter Seidel, hpsseidel@mpibsb.mpg.de, Max-Planck-Institut für Informatik, Saarbrücken, Germany; Piotr Didyk, piotr.didyk@usi.ch, Università della Svizzera Italiana, Lugano, Switzerland; Karol Myszkowski, karol@mpi-inf.mpg.de, Max-Planck-Institut für Informatik, Saarbrücken, Germany.

model, including mixing diffuse and specular components of different BRDFs, and random edits of diffuse hue, diffuse saturation, and specular hue. Finally, in order to balance the presence of different appearances, we use the work of Lagunas et al. [2019] to perform clustering of all these materials by similarity. We detected a very large amount of diffuse brown-black materials present in the UTIA dataset, therefore, we decided to keep only 120 materials from UTIA, and complement the dataset with more specular materials coming from our MERL edits. This results in 520 materials (100 MERL, 62 RGL, 120 UTIA, 50 isotropic UTIA, and 188 MERL edits).

Rendering We use Mitsuba1 for rendering our stimuli. Details are shown in Tab. S1.

Table S1. Rendering details of the dataset

Item	Information
Software	Mitsuba1 v0.6.0
Resolution	1024 × 1024
Integrator	path
Sensor film	hdrfilm
Sensor sample count	1024
Integrator maximum depth	32
Saved channels	10 = color (3) + depth (1) + albedo (3) + normal (3)
Storage	815 G for all channels in .exr format 32 G for color channels in .jpg format

1.2 Methods

Procedure We implemented a web interface for performing our perceptual survey. Fig. S1 shows screenshots of this interface, displaying several stages of the survey (introduction, attribute explanation, training, and an example trial). Before launching our large-scale study, we ran several iterations of a pilot with a subset of our dataset in order to refine the interface and explanations given to the participants, as well as to check that the chosen attributes were well understood and meaningful to them. After this procedure, the agreement between users for our selected attributes (computed as the percentage of agreement to the majority) was very high (between 0.85 and 0.92), and manual inspection of their answers showed that they did understand the attributes they were being asked to rate, therefore, we moved to the large-scale study. Before starting the survey, the task and different attributes were explained to the participants, and to allow them to familiarize with the interface and

minimize worker unreliability [Welinder et al. 2010] a short training session with obvious examples (not part of our tested stimuli) was presented. We also collected demographic information: gender (male - female - other - prefer not to answer), age, display used for the survey (computer - tablet - smartphone), knowledge of computer graphics, experience with design or modeling software, and artistic knowledge or experience (none - basic - intermediate - professional). Similar to previous works [Serrano et al. 2016; Zell et al. 2015; Zhang et al. 2020] we chose a Likert rating task, which offers a good trade-off between the amount of needed trials and difficulty of the task. We chose to use a seven-point Likert scale since they offer a good balance between granularity and complexity [Nunnally 1994]. The extremes of the scales were labeled as "not at all" and "a lot". Each perceptual survey consisted of 20 images selected randomly from our pool and 3 control images containing one very diffuse material and two very specular materials displayed at a random order within the survey. We use these images to detect inattentive or malicious participants: data from a participant is discarded if the response for the diffuse material is very glossy (glossiness > 3) or the responses for the specular ones are not glossy at all (glossiness < 5). We rejected 11% of the participants. We obtained 215,680 valid responses, from which we ensure that each image has at least 5 views.

S 2. ADDITIONAL RESULTS FROM THE STATISTICAL ANALYSIS

2.1 Image statistics

In addition to geometry and shape statistics reported in the main paper, we also compute low-level statistics and mid-level features of the final image and try to use these features as predictors for changes in the attribute ratings. Based on previous literature we compute the following low-level statistics: high-frequency content, wavelet coefficient sparsity, mean, standard deviation, skewness, kurtosis, and maximum and minimum [Toscani et al. 2017] (computed as the 90 and 10 percentile, respectively), and mid-level features: sharpness, contrast and coverage of highlights [Marlow and Anderson 2013]. Note that we compute these statistics only in the surfaces of interest, and we do not take into account the background. We then use these statistics as predictors for changes in the attribute ratings via elastic net regression [Kuhn et al. 2008]. The coefficients of this regression (proportional odds ratios) correspond to the expected change in the ratings for changes in each of the corresponding predictor variables while holding the other predictors constant, therefore we use them as indicators of the importance of each predictor.

Glossiness These predictors explain 46% of the variance in glossiness, with the strongest predictors being the minimum and the standard deviation. Although we have found skewness to also be strongly related to changes in glossiness, recent work has suggested that the standard deviation (as a measure of contrast) may correlate better with perceived gloss than skewness for realistic images [Wiebel et al. 2015]. Our images depict real-world illuminations with complex shapes: we found that in our case, standard deviation also has a stronger contribution in the prediction of gloss. The specular lowlights have been shown to strongly contribute to the appearance of gloss [Kim et al. 2012]. The gradients (indicative of

sharpness) also play a strong role in predicting glossiness. This is in agreement with previous research suggesting that mid-level perceptual features such as the coverage, contrast, and sharpness of specular highlights (which are related to intensity gradients) as good predictors for perceived gloss [2013].

Lightness The tested predictors explain 25% of the variance in perceived lightness, with the strongest predictors being the minimum, maximum, mean, and skewness. Although less research has been devoted to lightness perception, several works have shown that it may be also correlated with image skewness [Motoyoshi et al. 2007; Sharan et al. 2008]; our data also shows this correlation. Maximum, minimum, and mean luminance also play an important role in the perception of lightness, as expected.

Metallicness In this case, predictors explain 38% of the variance of metallicness, with the strongest ones being the same as for glossiness: minimum, and standard deviation. We hypothesize that in our study, observers associated metallicness to color, and therefore in this case, the percentage of variance explained with the tested predictors is lower than for glossiness.

Although the experiments of Motoyoshi et al. [2007] showed that skewness alone already explained a high variance in glossiness perception, Anderson et al. [2009] showed that when a larger range of different materials and geometry are presented, the correlation between skewness and perceived gloss heavily drops. Marlow et al. [2013] show a high correlation between glossiness judgments and perceived sharpness, contrast, and coverage of highlights, although as Nishida [Nishida 2019] points out, these are also outputs of visual processing estimated from human judgments; instead, we estimate these features directly from images. Simple models for prediction based on simple statistics yield valuable insights but our analysis supports that they do not explain all the variability in the data, specially when attempting to deal with large variability in illumination, geometry, and reflectance. This is in accordance with previous works [Nishida 2019; Olkkonen and Brainard 2010], and it has been recently suggested that more complex encodings of the visual input may better explain the underlying processes of material perception [Delanoy et al. 2020; Fleming and Storrs 2019; Lagunas et al. 2021]. In the main paper, we rely on deep features in order to predict perceived attributes.

2.2 Attribute correlation

We analyze Spearman rank correlations between the users' answers to our five reflectance attributes, shown in Figure 2. We also include the distribution of all the raw collected data (Figures 37 and 38), as well as correlations between our attributes (Figures 39 and 39) by geometry and illumination.

Our perceptual data shows very high correlations among gloss-related attributes, as well as metallicness, which is to be expected, since they all contribute to the overall perception of gloss. These correlations are in line with those observed by Serrano et al. [2016]. We can also observe that lightness is not completely orthogonal to glossiness, which is in agreement with previous work [Chadwick and Kentridge 2015; Pellacini et al. 2000; Sèvè 1993] since lightness also plays a role in the perception of gloss.

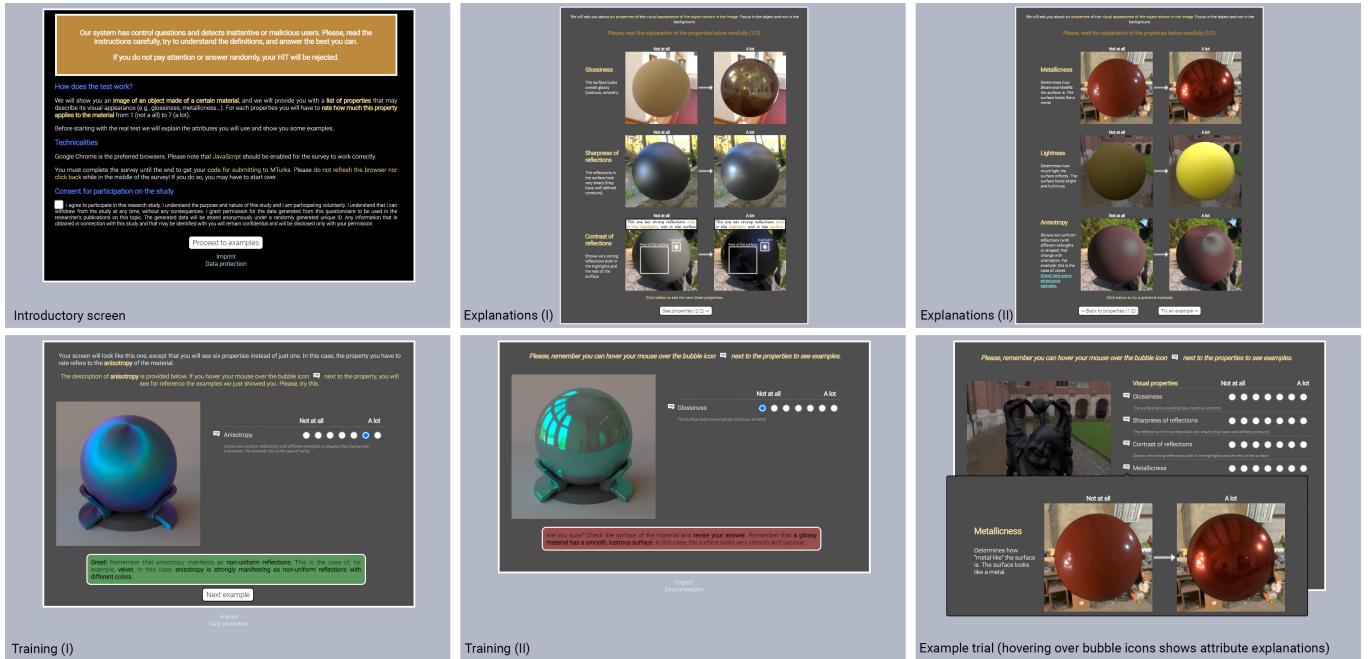


Fig. S1. Interface of our web-based online experiment for collecting perceptual ratings. In order: introduction to the experiment, explanation of the attributes, training examples, and an example trial.

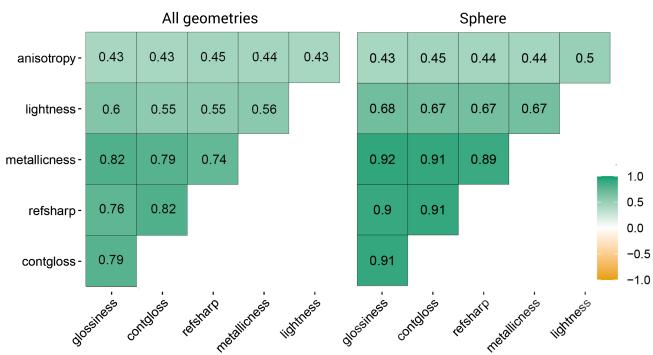


Fig. S2. Spearman rank correlations between our six attributes for all geometries pooled (left), and for the *sphere* only (right). In both cases pooled for all environment maps.

When looking at the different geometries, we can observe that the *sphere* shows higher correlations for gloss-related attributes than the rest of the geometries. This could indicate that, as already suggested by Vangorp et al. [2007], very simple smooth surfaces do not provide sufficient cues for discriminating more sophisticated gloss behaviors such as sharpness of reflections, contrast gloss, or metallicness.

2.3 Anisotropy

Due to the limited amount of measured anisotropic materials in existing datasets, our dataset is biased towards isotropic materials.

Therefore, in order to analyze anisotropy, we select 50 anisotropic materials from the UTIA dataset, and their isotropic counterparts. First, we want to analyze the effect of illumination and geometry in the perception of anisotropy. Second, we want to analyze if the presence of anisotropy affects the perception of gloss, lightness, and metallicness. Similarly to the previous section, we make use of cumulative link mixed models and in this case, we additionally include as a fixed factor a binary variable that encodes whether each image contains an anisotropic material, and the interactions of this variable with illumination and geometry.

For the materials tested, we found that anisotropy ratings are only significantly different for isotropic-anisotropic pairs for the *sphere* ($p < 0.001$), and for the illuminations (H) ($p = 0.0012$) and (I) ($p = 0.0039$). In Fig. S3 we show an example of one of the tested anisotropic materials for all our illuminations and geometries. Filip et al. [2015] analyzed the perception of anisotropy for three different geometries (sphere, blob, and tablecloth) under point-light illumination, and for one geometry (sphere) under a natural environment map (Grace). They found that for the sphere and blob sensitivity to anisotropy was similar under point-light illumination, while for the more complex shape, it dropped significantly. When comparing the sphere under point-light illumination and the complex environment map, perception of anisotropy heavily decreased for the latter. Although our experimental design is different (they used pairwise comparisons and we use single rating), we found the same trends in our data, however in our case, only the sphere shows significant differences: this may be due to the differences in experimental design, or simply due to the higher complexity of

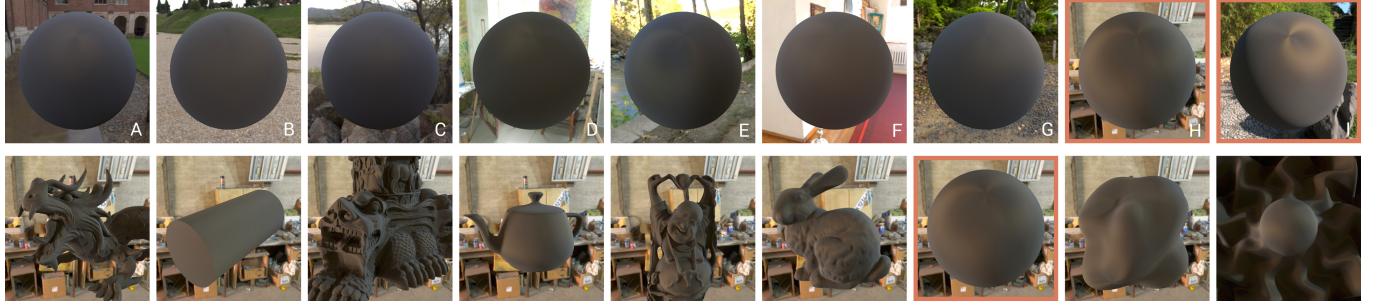


Fig. S3. Examples of our nine illuminations (top) and geometries (bottom) with the material *fabric002* from the UTIA dataset. In our experiment, we included 50 materials from the UTIA dataset, and their isotropic counterparts: we found that users could only consistently identify anisotropy (significant differences between anisotropic-isotropic pairs) for the sphere, and two of our illuminations (marked in orange).

our environment maps. Interestingly, we found that some environment maps (H and I), even though complex, allow for consistently identifying anisotropy. These two environment maps, even though very different at first sight, correspond to the lowest I_{area} , and the highest I_{range} , i.e., they contain point-like light sources. We show that these two statistics are highly correlated to the perception of anisotropy.

We did not find significant differences in glossiness, metallicity, and lightness under the presence of anisotropy. However, note that of these 50 BRDFs only some feature strong anisotropy (as shown in Fig. S3), while others have more subtle anisotropy: In this analysis, we focus on deriving general trends, but further analysis could be performed in our data, e.g., the analysis could be restricted to individual anisotropic materials, and the relationship between strength of anisotropy and whether it is visible or not for different combinations of envmap and geometry could be further studied.

After these results, we limit the anisotropy attribute only to this analysis, and we do not include it for prediction. Unfortunately, current datasets of measured materials do not contain enough variety of anisotropic materials, and therefore, our data is biased towards isotropic materials. Additionally, most of these anisotropic materials show very subtle effects, and are hard to see for most geometries even for human observers, as we have shown in this section. This produces unreliable results for anisotropy estimation as shown in the main paper.

S 3. PREDICTOR

In this section we include additional results for the predictor and perceptual embeddings for different sets of geometries and illuminations, featuring different combinations of our attributes.

3.1 Predictor Results

We provide more predictor results with all attributes to demonstrate the effectiveness of our predictor.

3.1.1 Predictions for Different Points of View

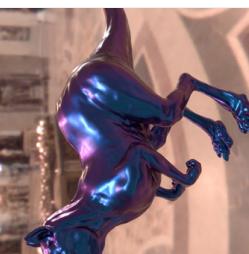
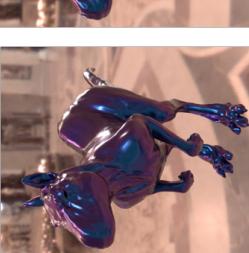
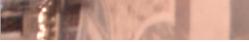
	
	
	
	
	
	
Glossiness 0.380943	0.369327 0.311193 0.2744 0.349502
Refsharp 0.242475	0.185116 0.141114 0.090896 0.151476
Contgloss 0.143992	0.150962 0.118105 0.055661 0.093019
Metallicness 0.303627	0.25526 0.179073 0.171755 0.247553
Lightness 0.433607	0.422175 0.432612 0.40314 0.427466
Glossiness 0.676348	0.79405 0.841923 0.751458 0.742788
Refsharp 0.322322	0.490562 0.613431 0.511079 0.313417
Contgloss 0.226098	0.364963 0.540821 0.348757 0.312349
Metallicness 0.403685	0.739041 0.776426 0.698096 0.551455
Lightness 0.358642	0.407588 0.505088 0.5104 0.404786

Fig. S4. Predictions for different points of view.

3.1.2 Predictions for increasing surface relief

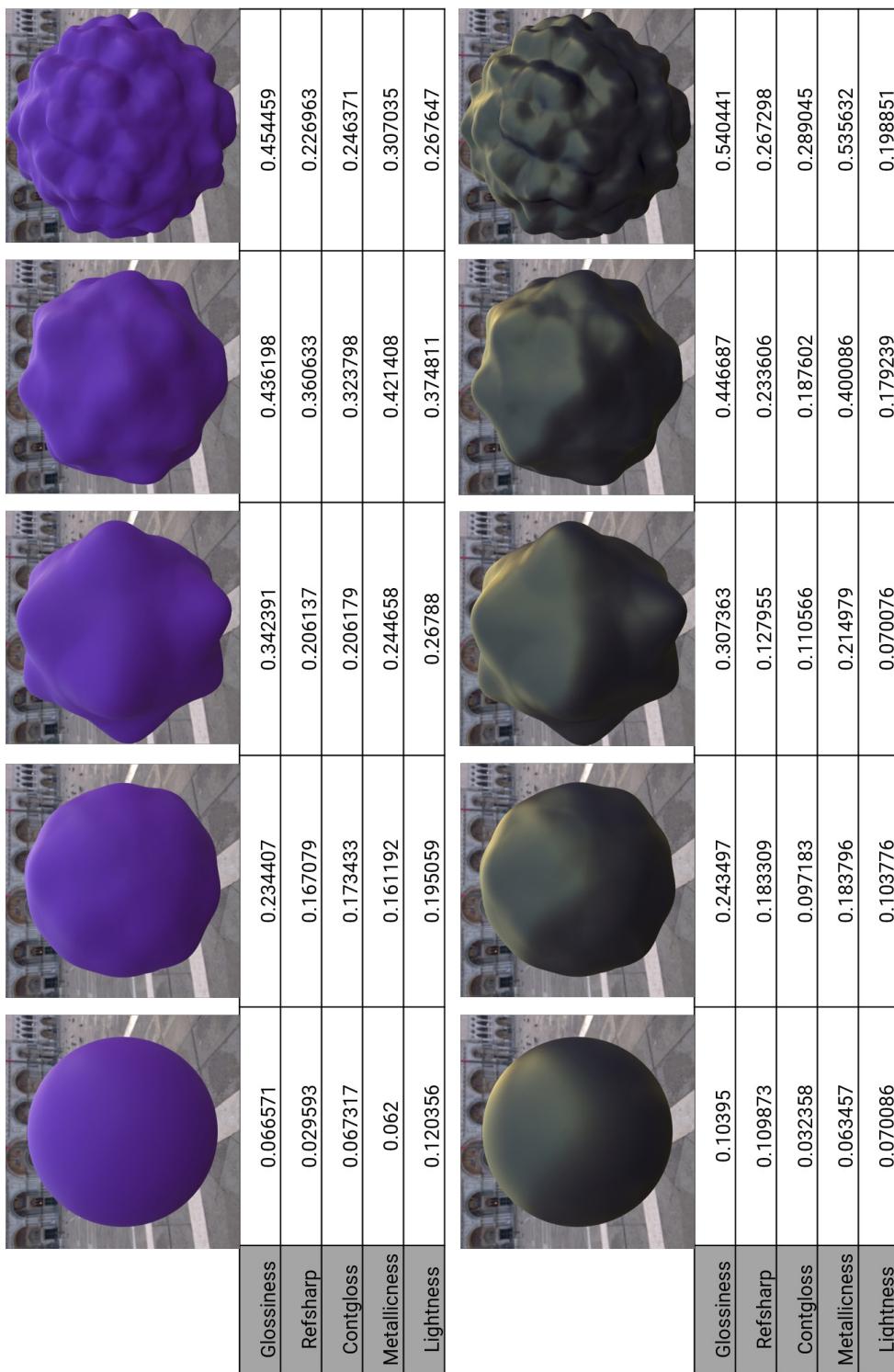


Fig. S5. Predictions for increasing surface relief.

3.1.3 Predictions for Increasing Blurry Environment Map

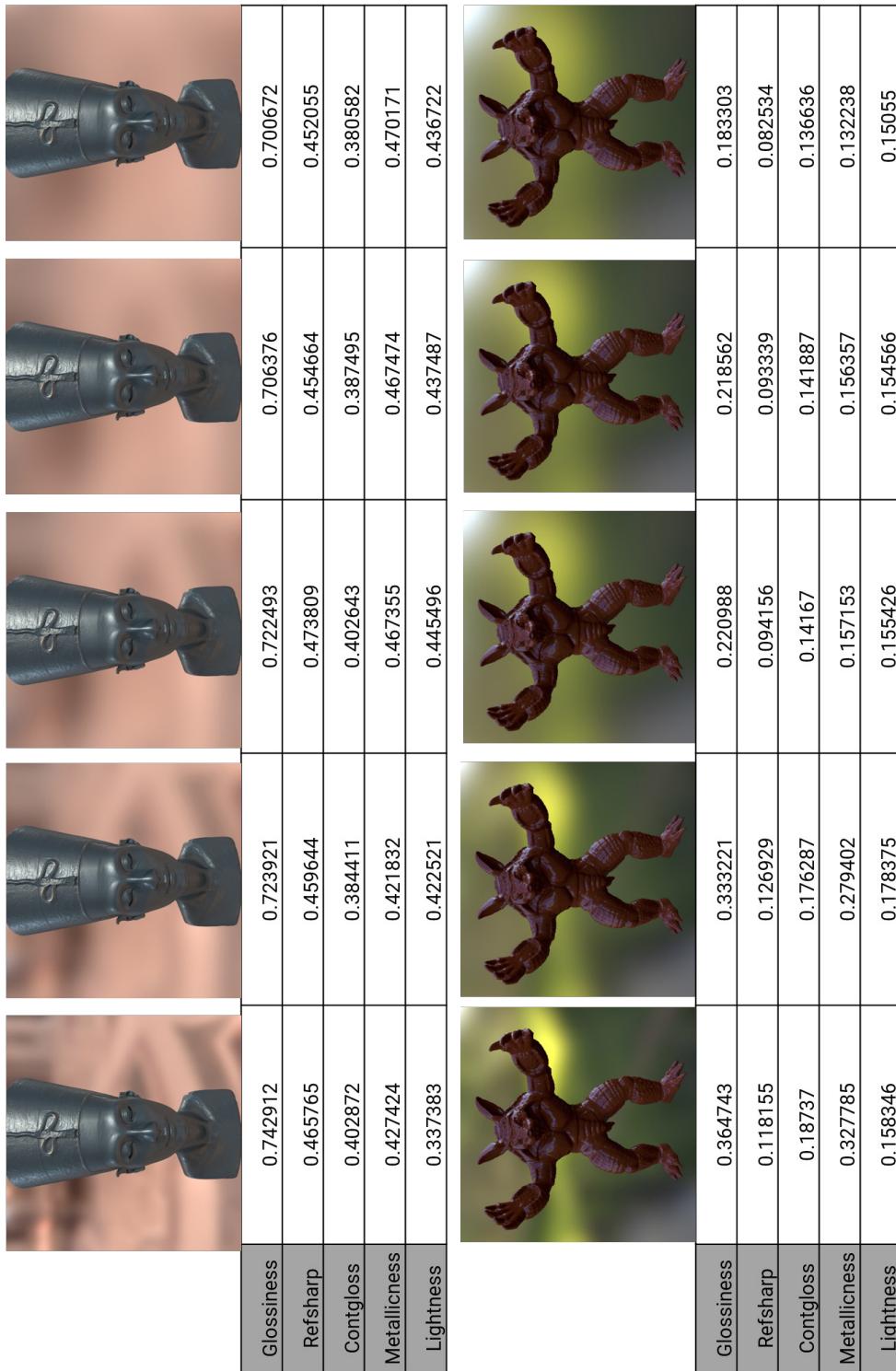


Fig. S6. Predictions for different blurriness of environment map.

3.2 Embedding Results

In this section we show additional embeddings for various datasets including: Lagunas et al., our validation dataset, and a few additional datasets with geometries and illuminations never seen during training. The embeddings show the relationship between attributes: lightness-glossiness, glossiness-sharpness of reflections, glossiness-metallicness, and metallicness-contrast of reflections.

3.2.1 Embeddings for the Lagunas et al. dataset

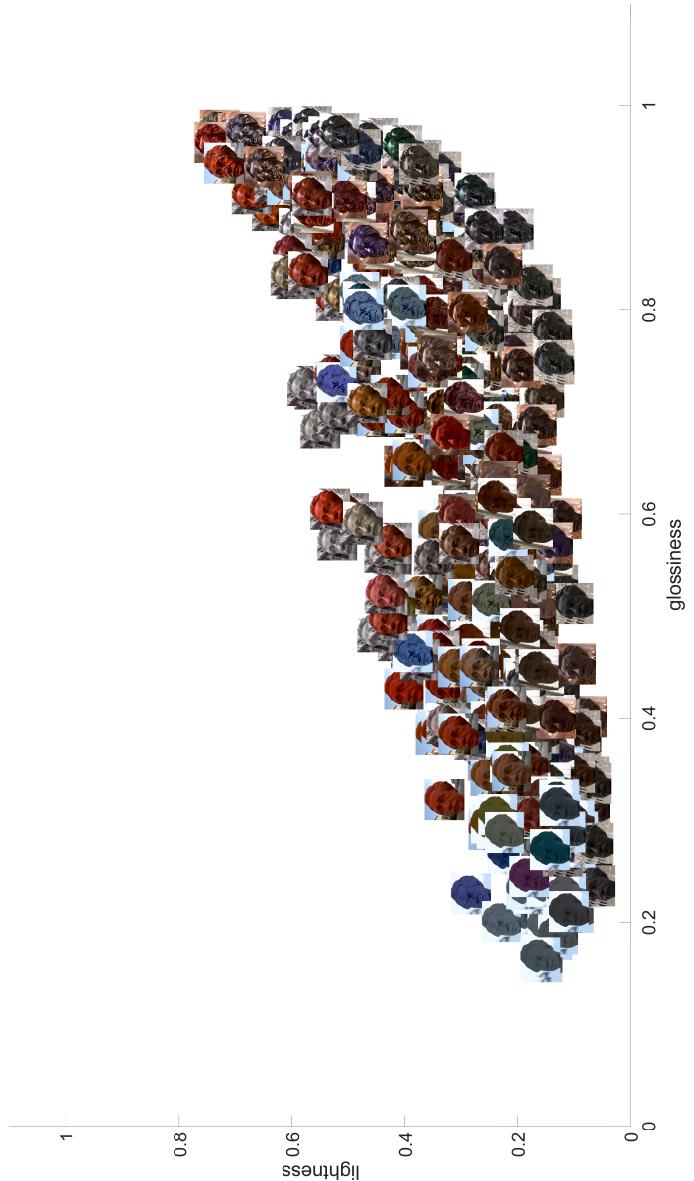


Fig. S7. Lightness-glossiness embedding generated from our predictor with geometry Einstein.

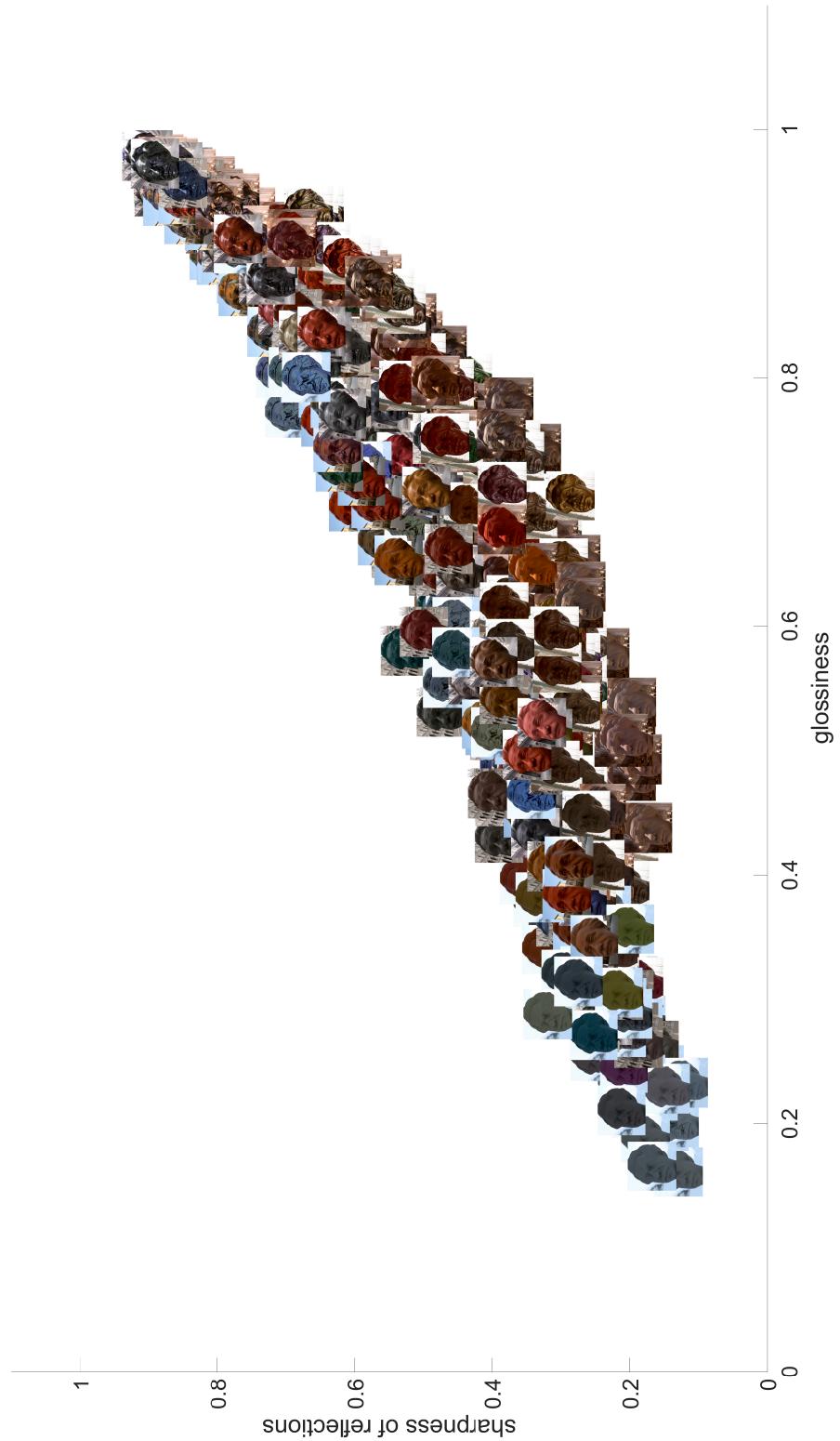


Fig. S8. Glossiness-sharpness of reflections embedding generated from our predictor with geometry Einstein.

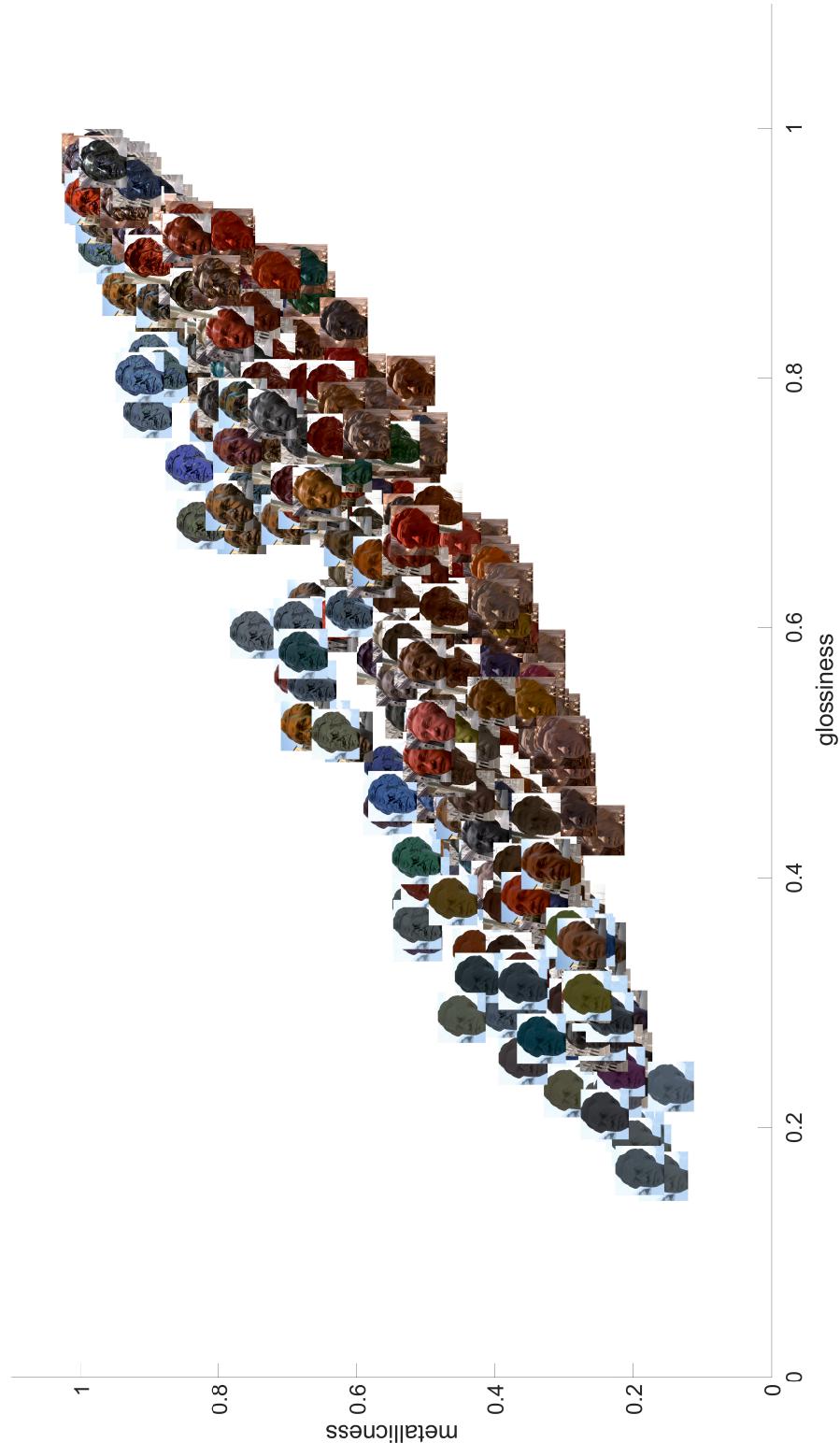


Fig. S9. Glossiness–metallicness embedding generated from our predictor with geometry Einstein.

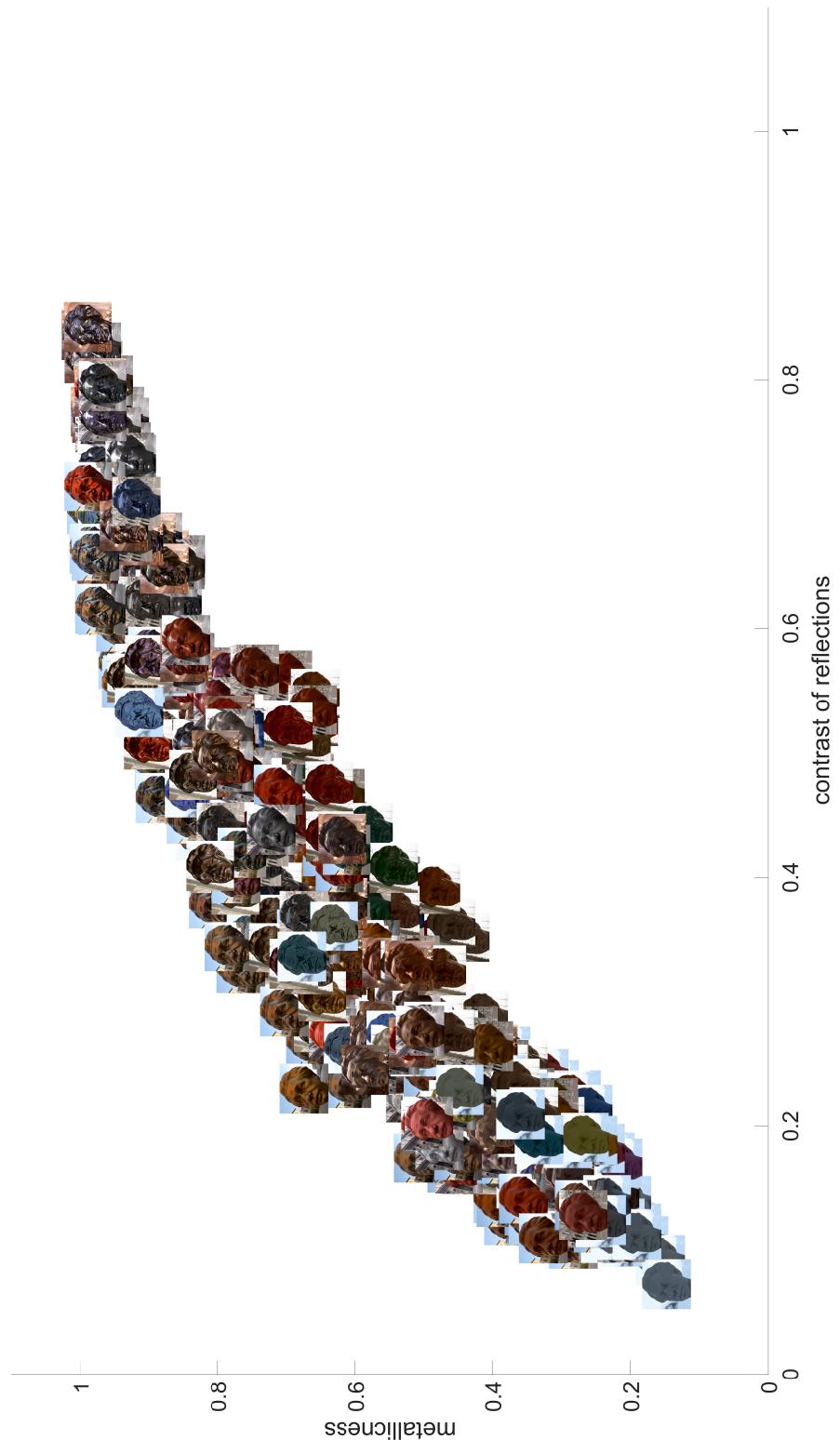


Fig. S10. Metallicness-contrast of reflections embedding generated from our predictor with geometry Einstein.

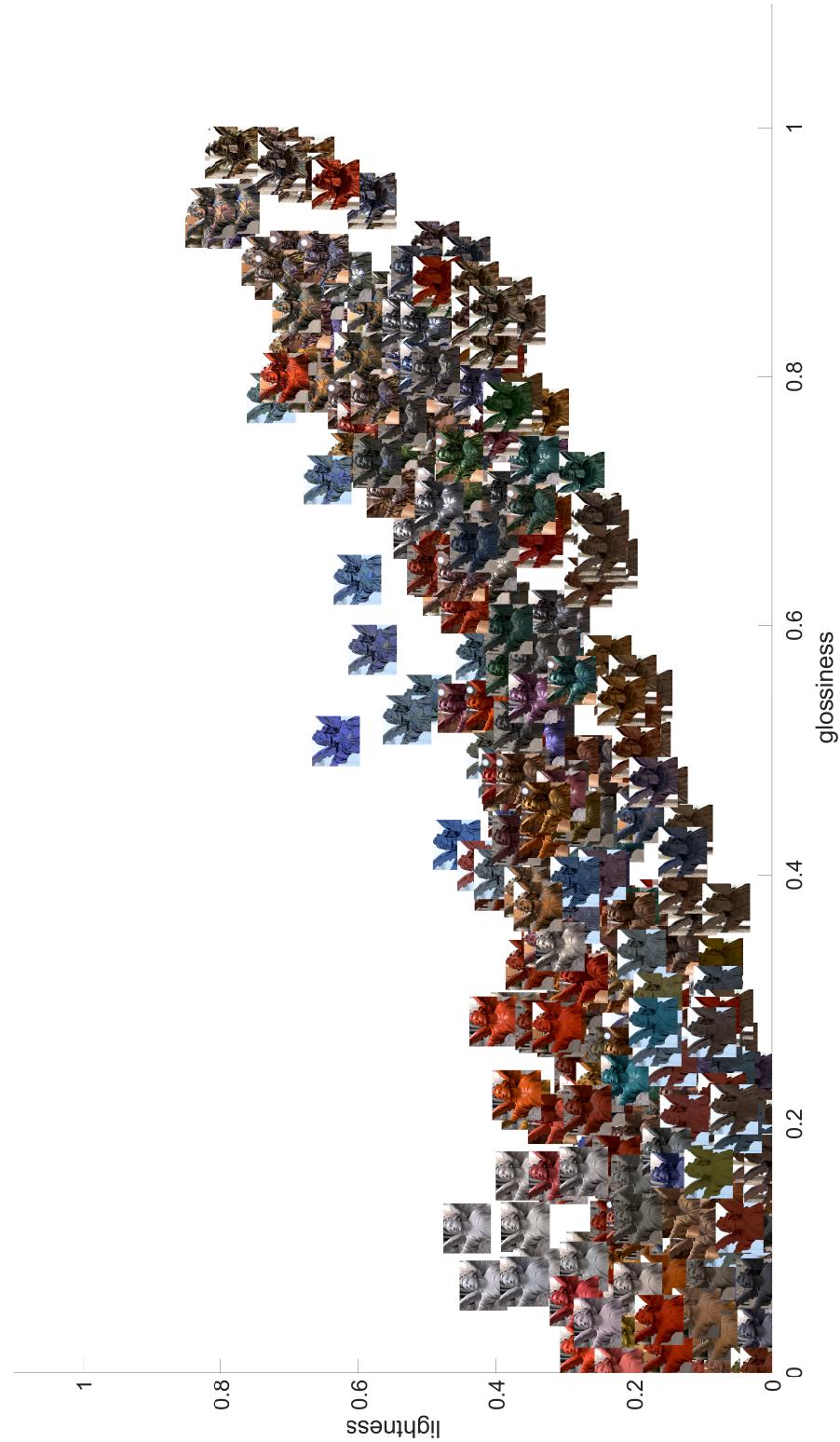


Fig. S11. Lightness-glossiness embedding generated from our predictor with geometry Lucy).

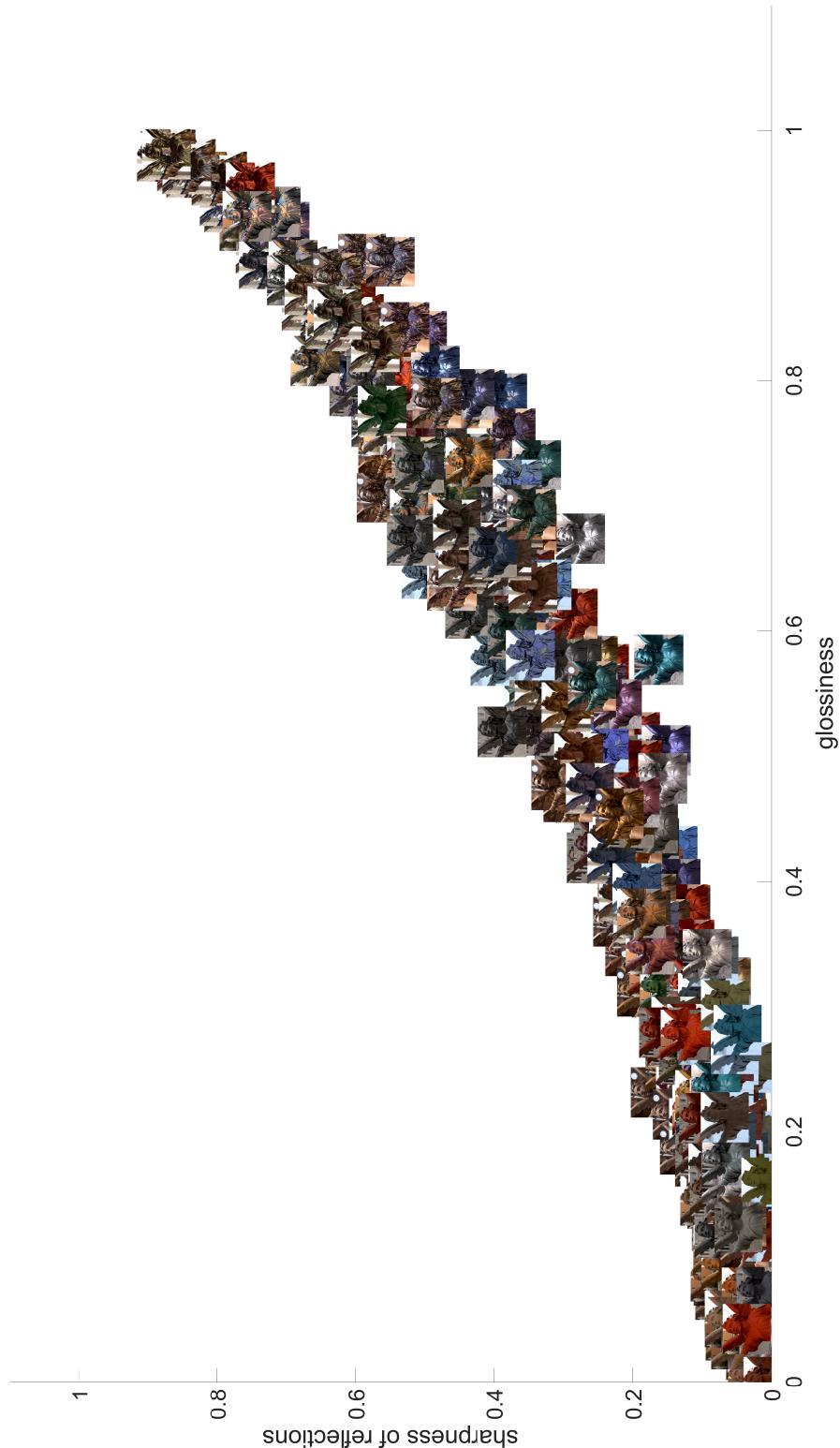


Fig. S12. Glossiness-sharpness of reflections embedding generated from our predictor with geometry Lucy.

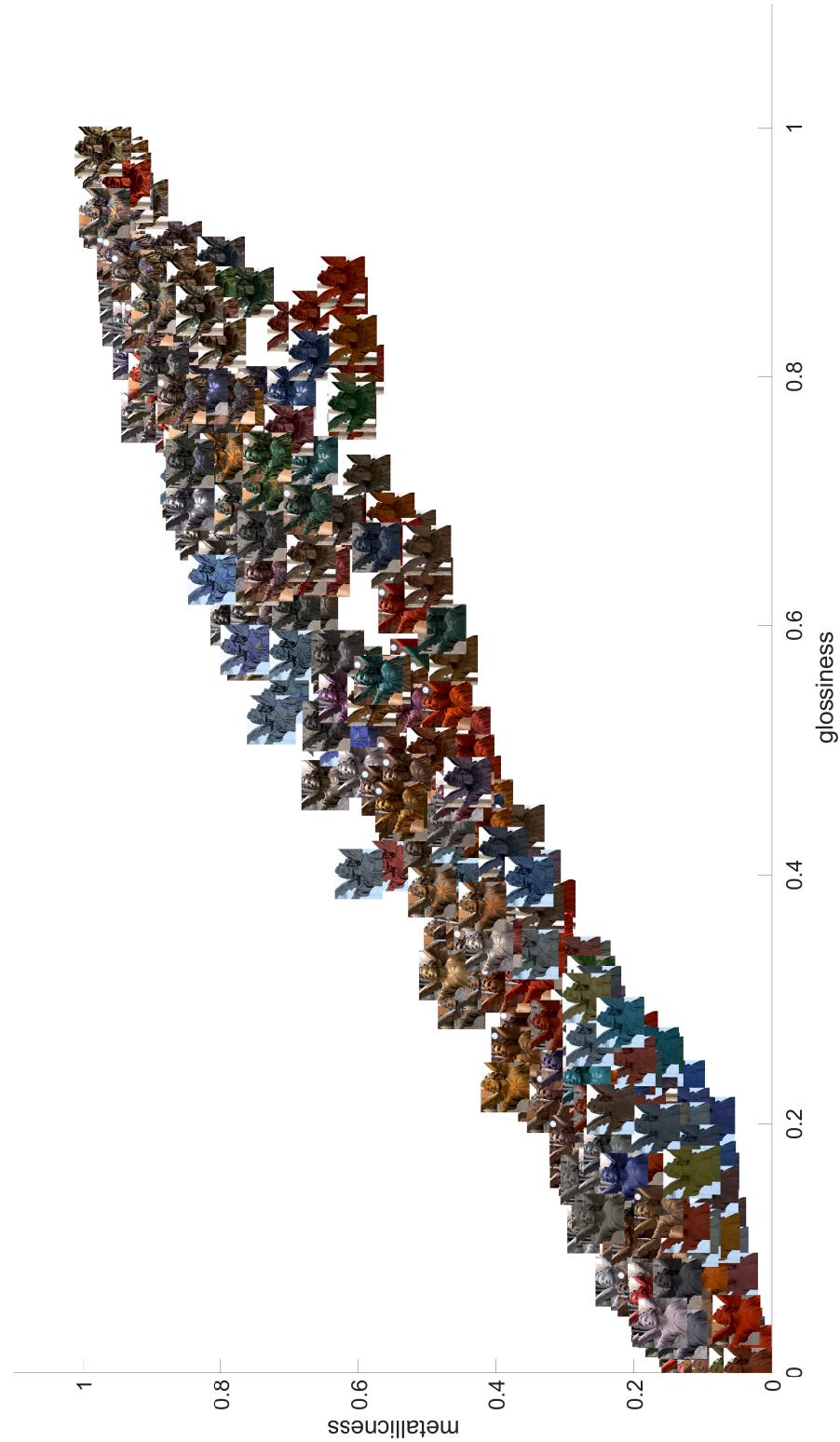


Fig. S13. Glossiness–metallicness embedding generated from our predictor with geometry Lucy.

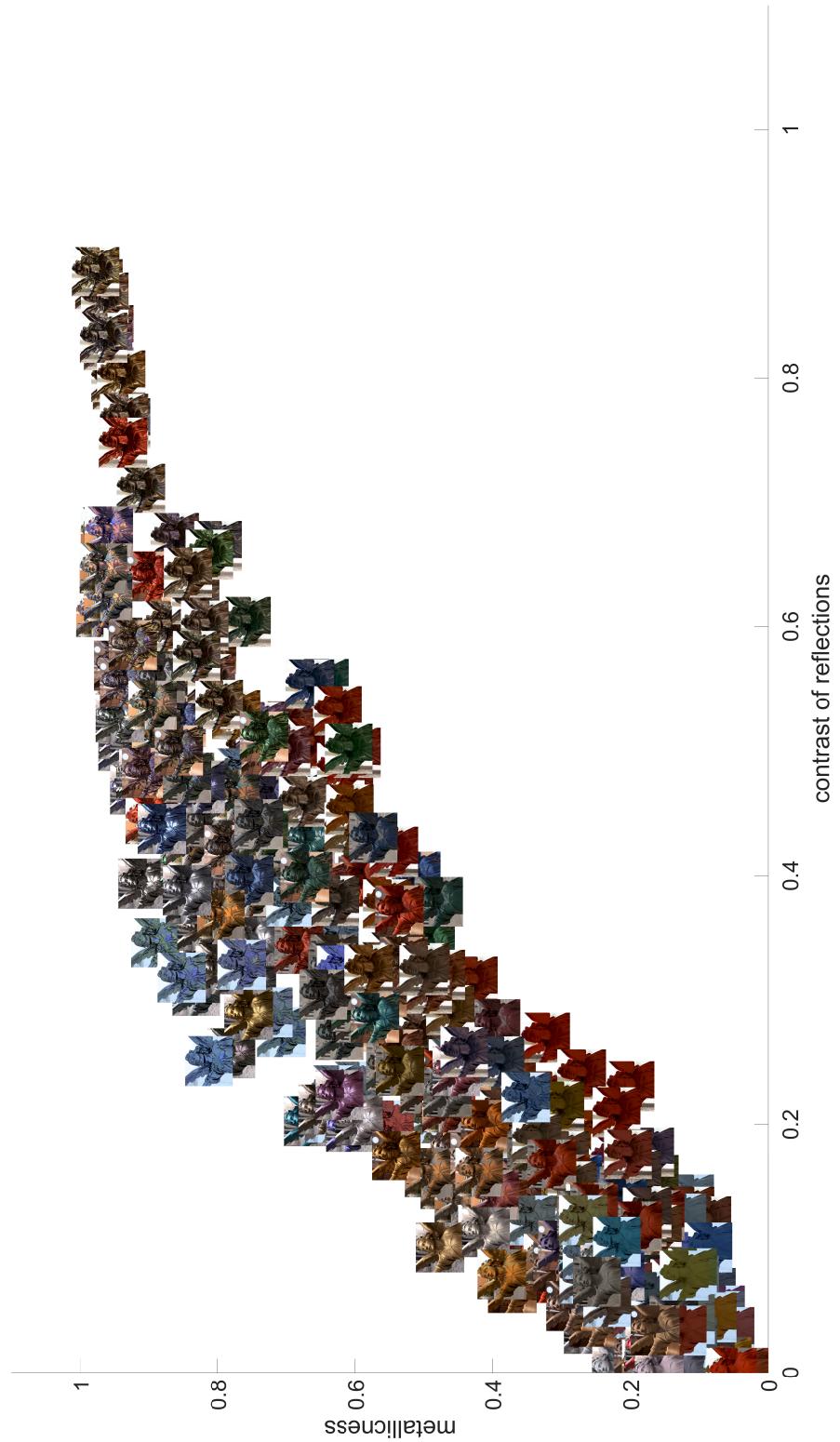


Fig. S14. Metallicness-contrast of reflections embedding generated from our predictor with geometry Lucy.

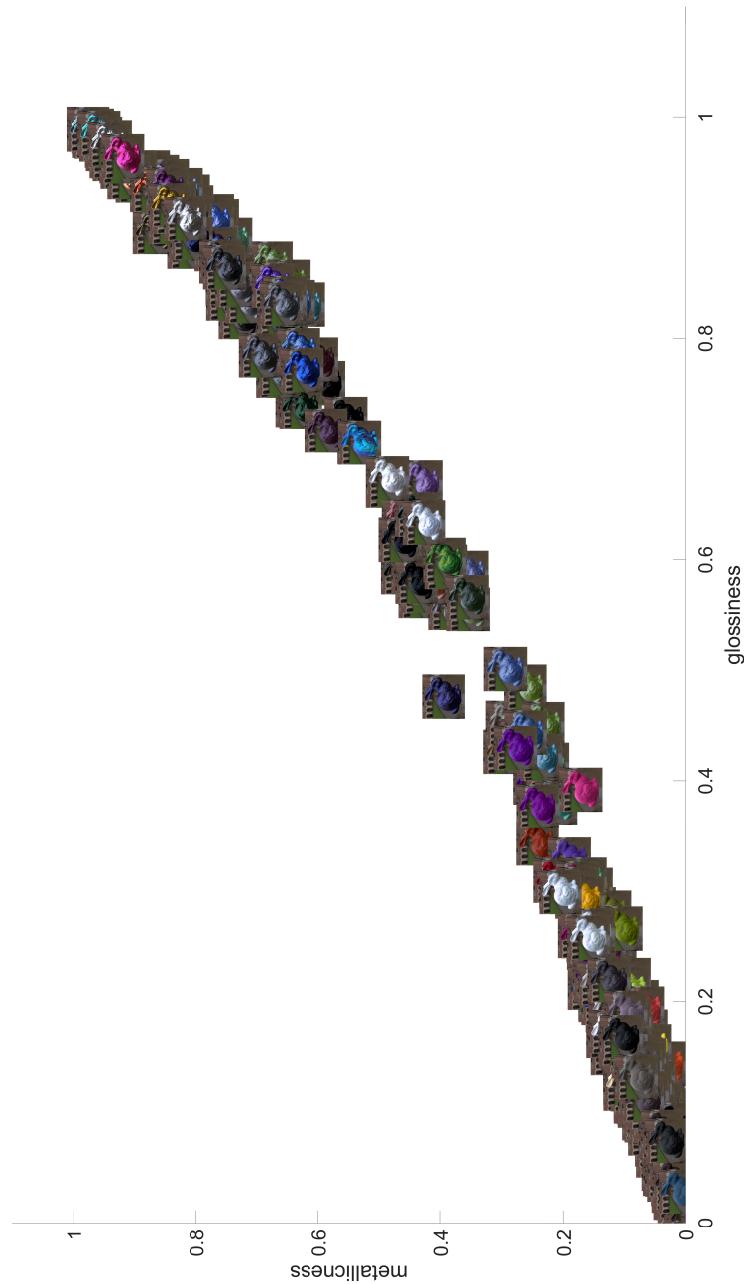
3.2.2 *Embeddings for our validation dataset*

Fig. S15. Glossiness-metallicity embedding generated from our predictor.

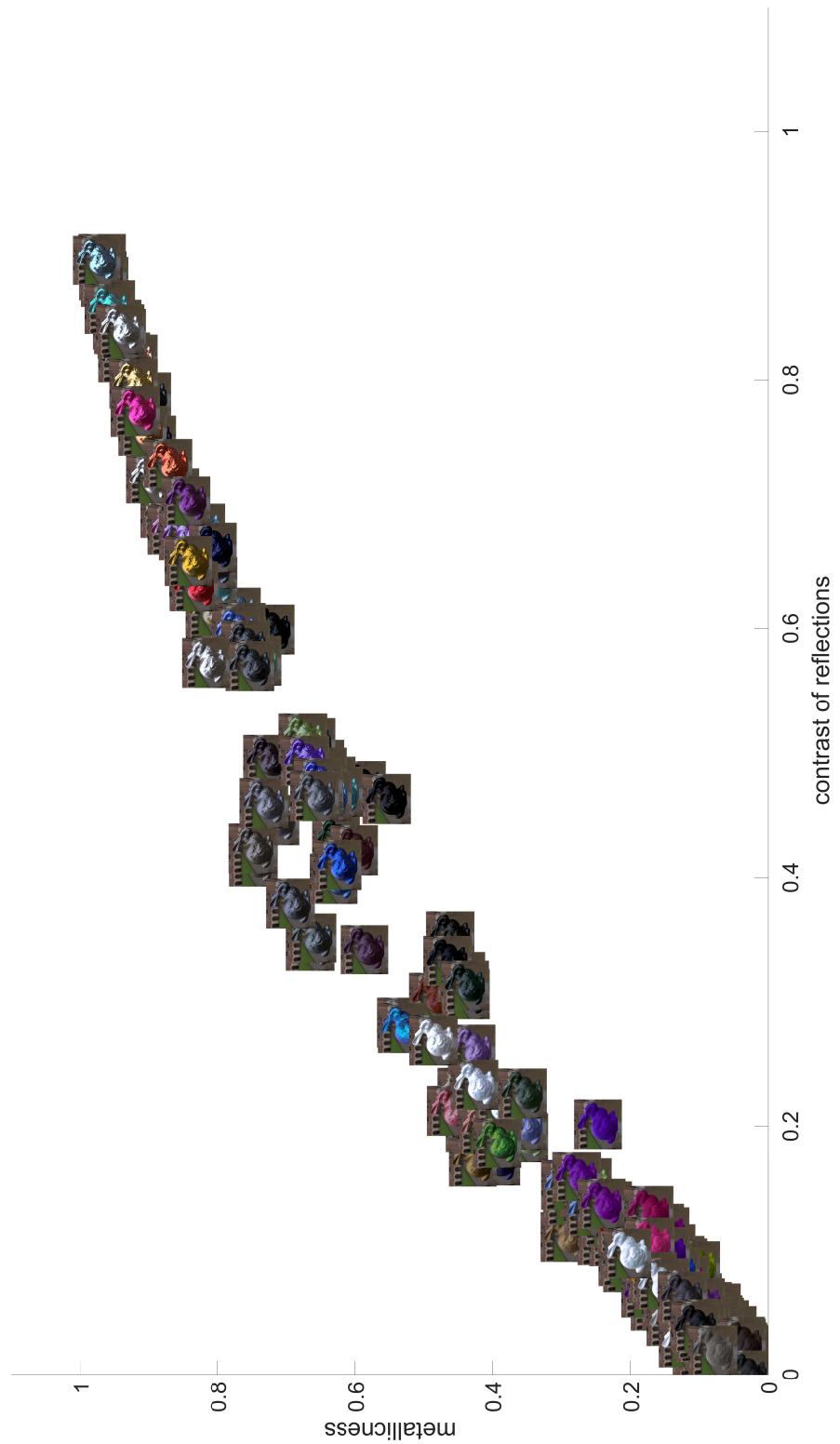


Fig. S16. Metallicity-contrast of reflections embedding generated from our predictor.

3.2.3 Embeddings for additional datasets We test our predictor with additional images rendered from more environment maps and geometries. We name them after their illumination and geometries: ennis-centaur, ennis-ganesha, pisa-centaur, pisa-ganesha, and pisa-sphere.

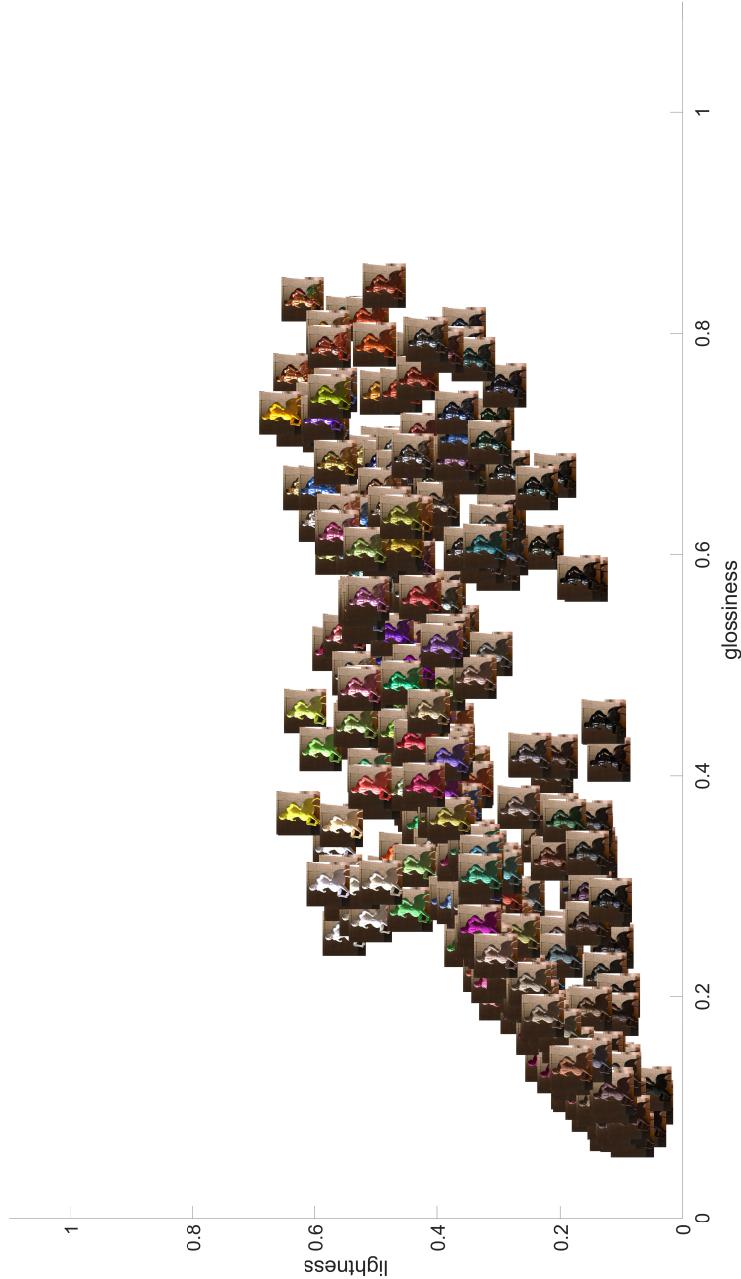


Fig. S17. Lightness-glossiness embedding generated from our predictor (ennis-centaur).

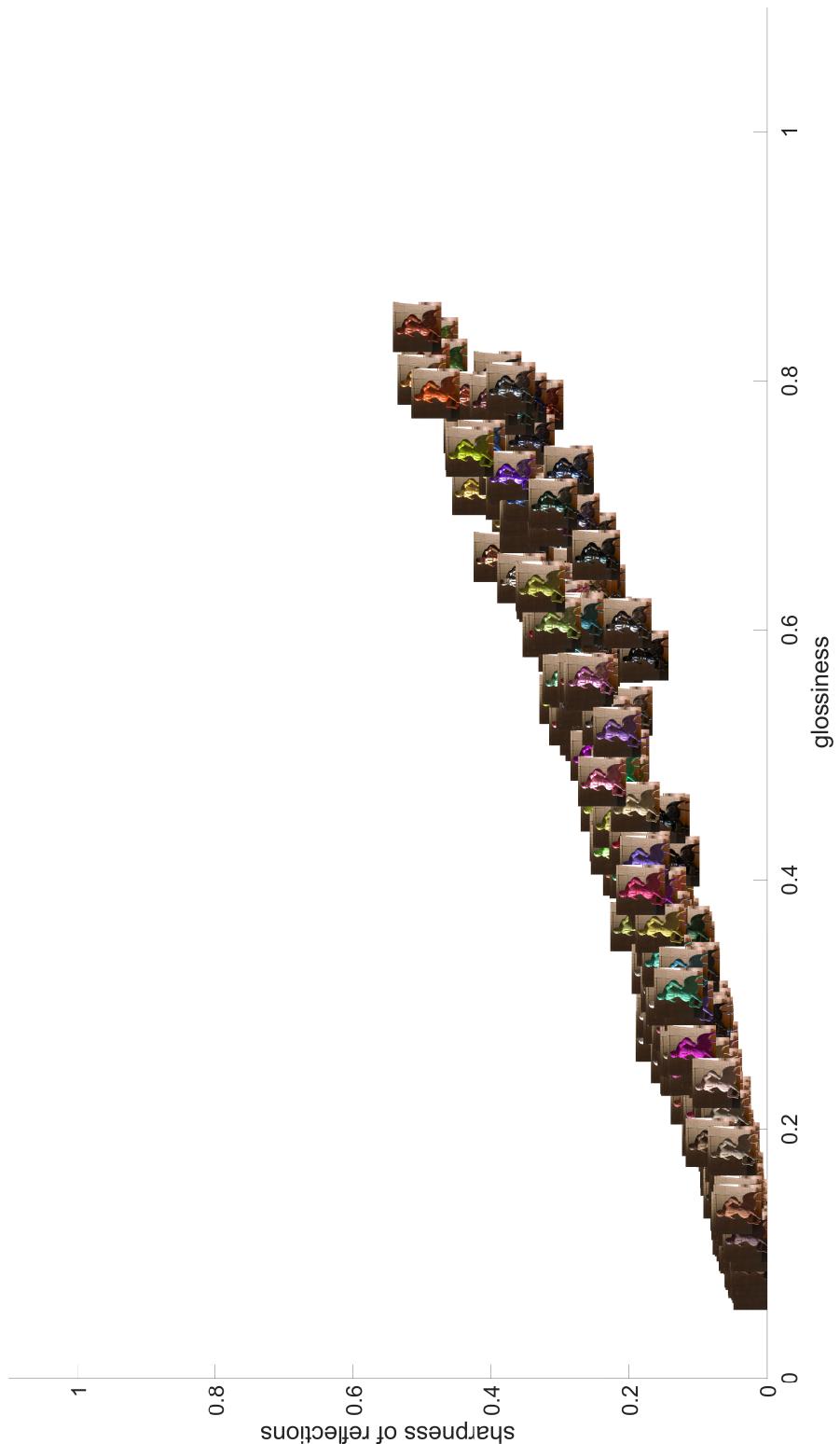


Fig. S18. Glossiness-sharpness of reflections embedding generated from our predictor (ennis-centaur).

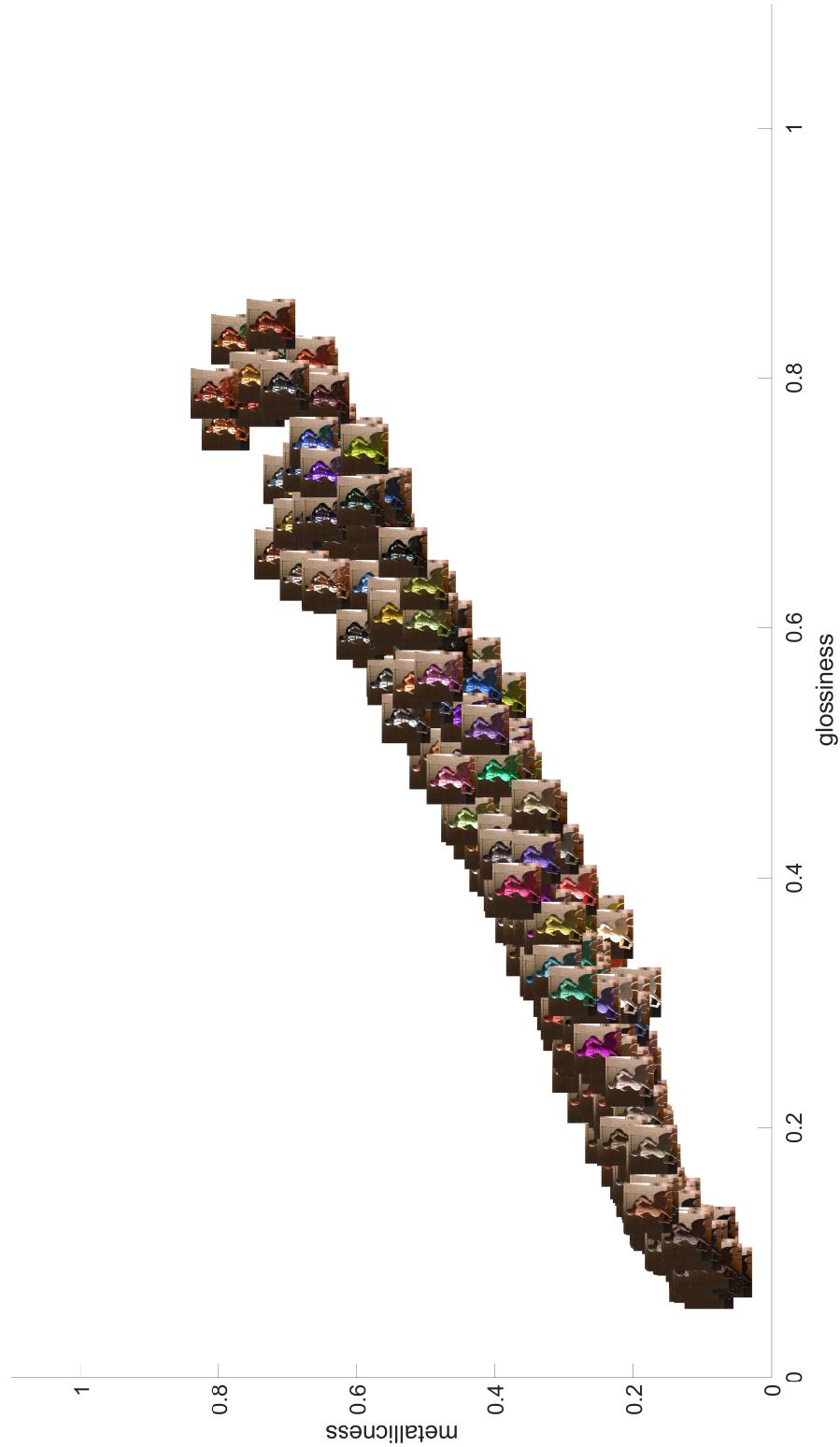


Fig. S19. Glossiness-metallicness embedding generated from our predictor (ennis-centaur).

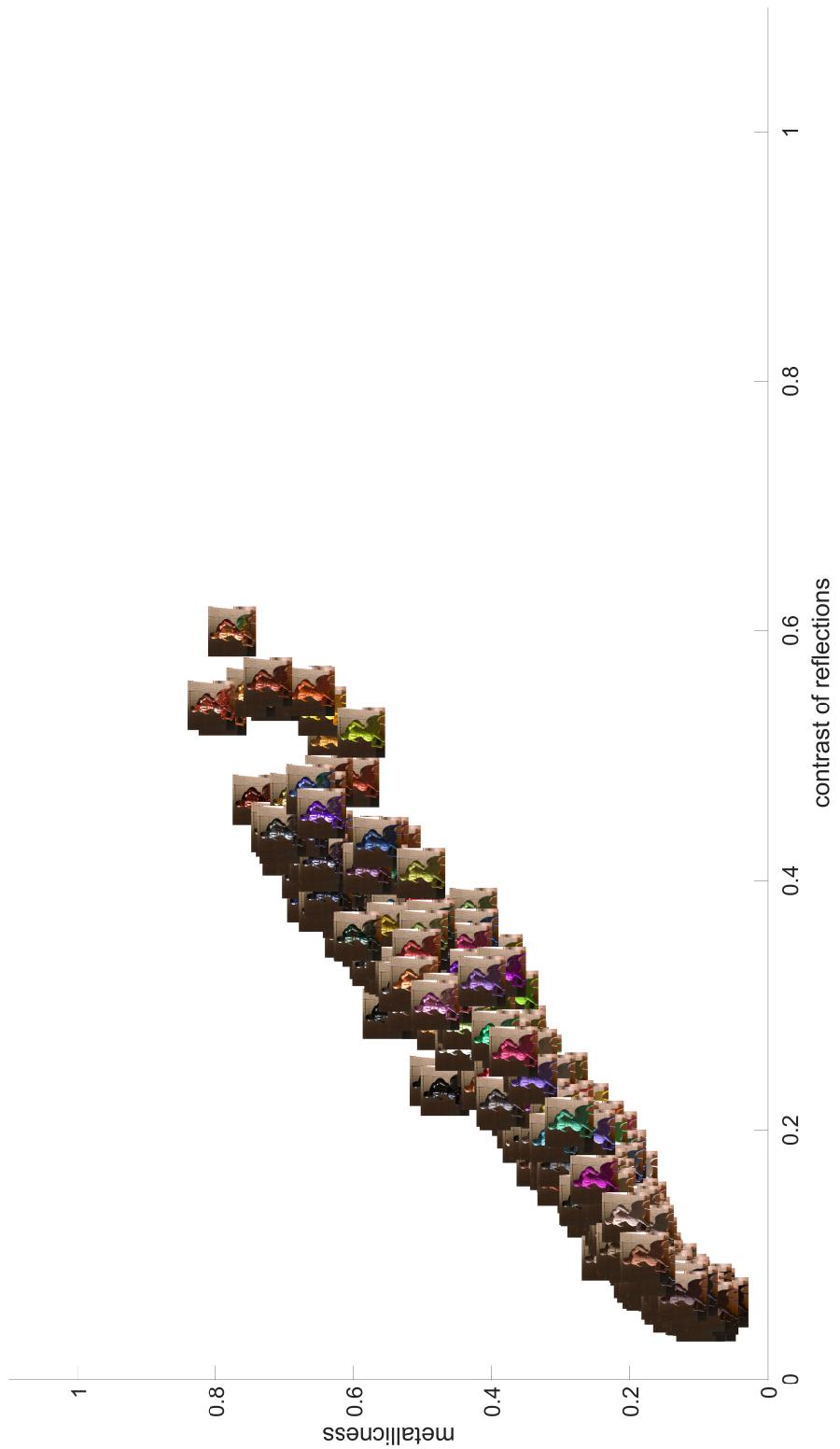


Fig. S20. Metallicity-contrast embedding generated from our predictor (ennis-centaur).

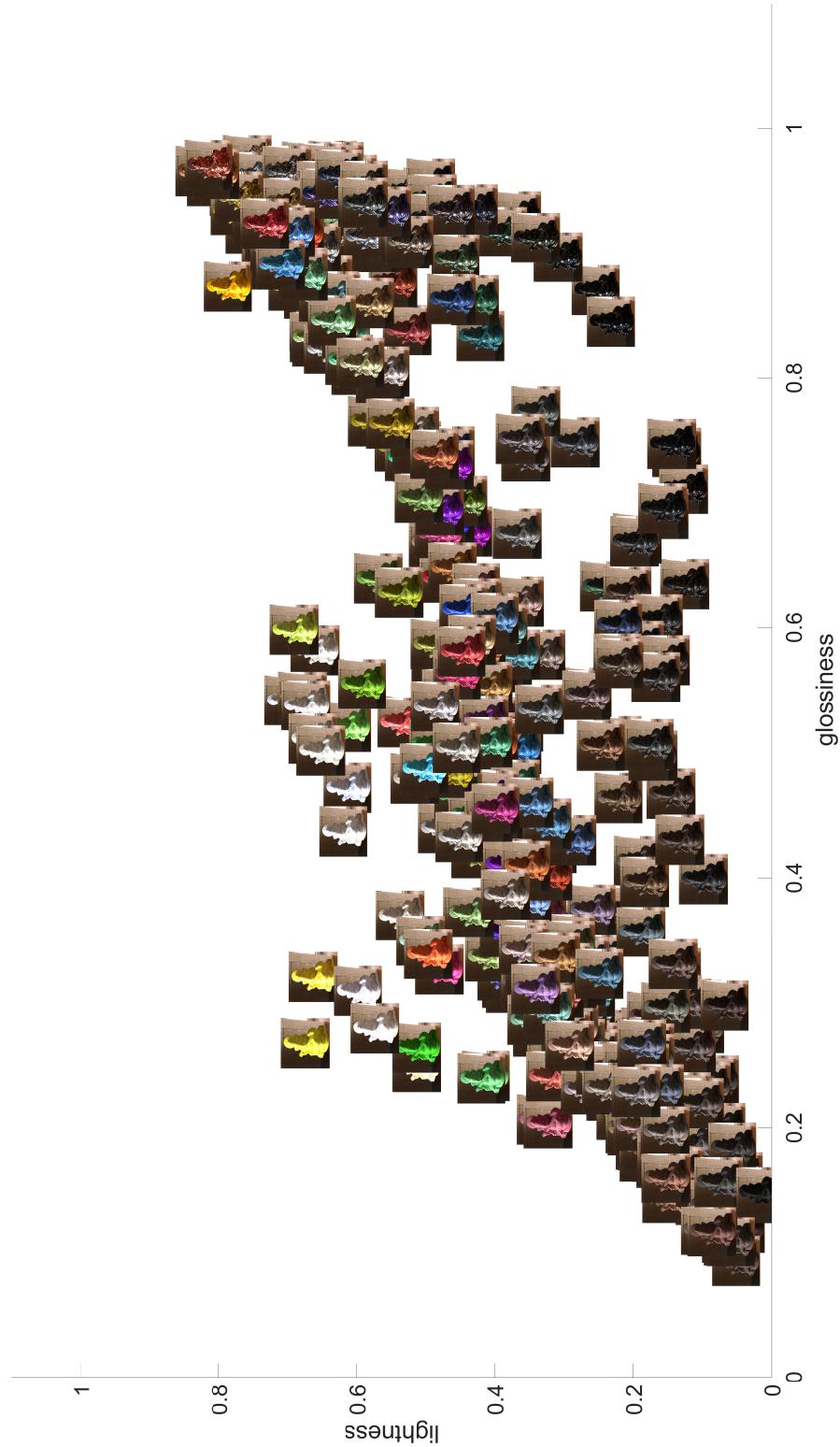


Fig. S21. Lightness-glossiness embedding generated from our predictor (ennis-ganesha).

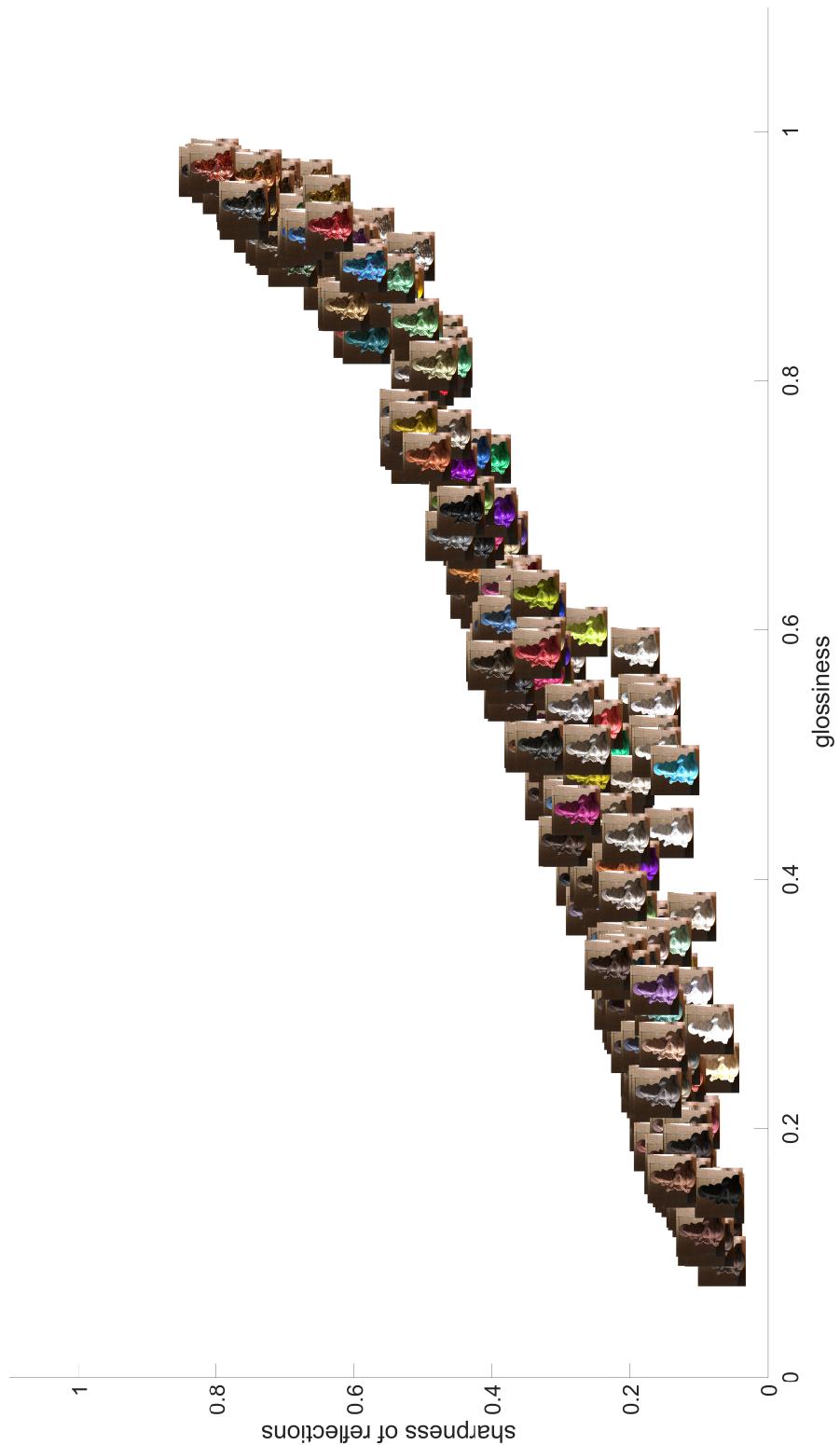


Fig. S22. Glossiness-sharpness of reflections embedding generated from our predictor (ennis-ganesha).

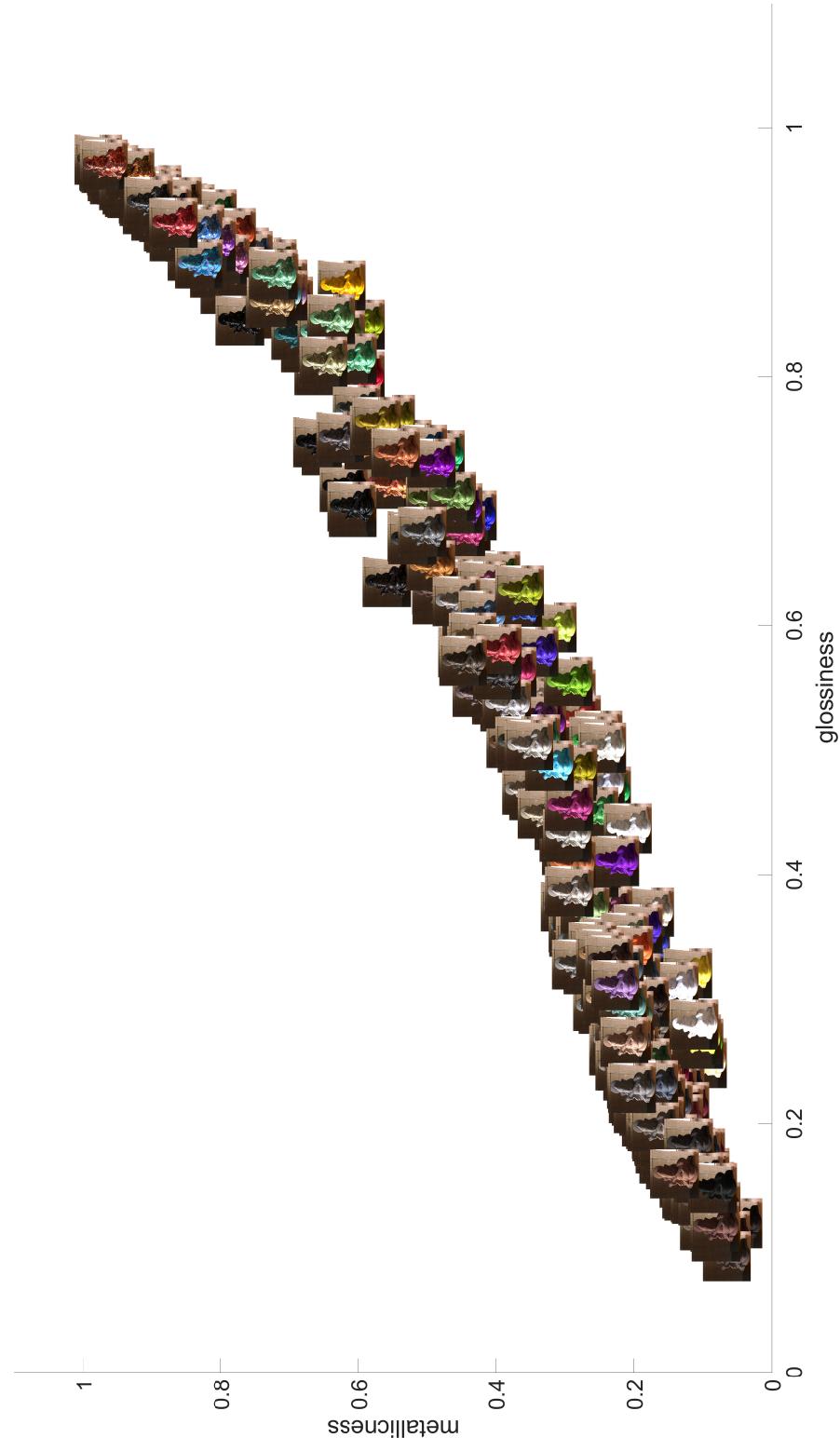


Fig. S23. Glossiness-metallicness embedding generated from our predictor (ennis-ganesha).

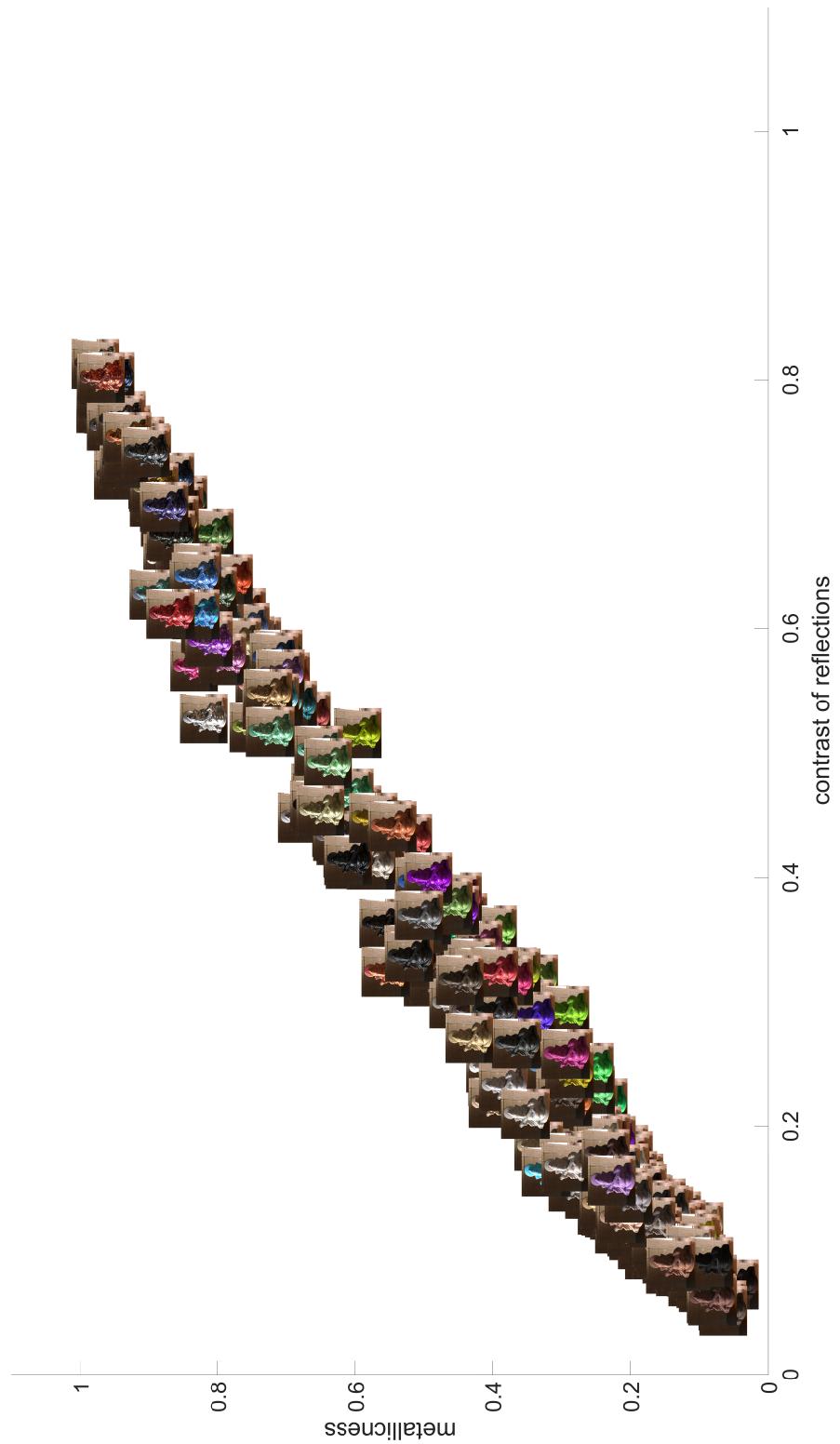


Fig. S24. Metallicness-contrast of reflections embedding generated from our predictor (ennis-ganesha).

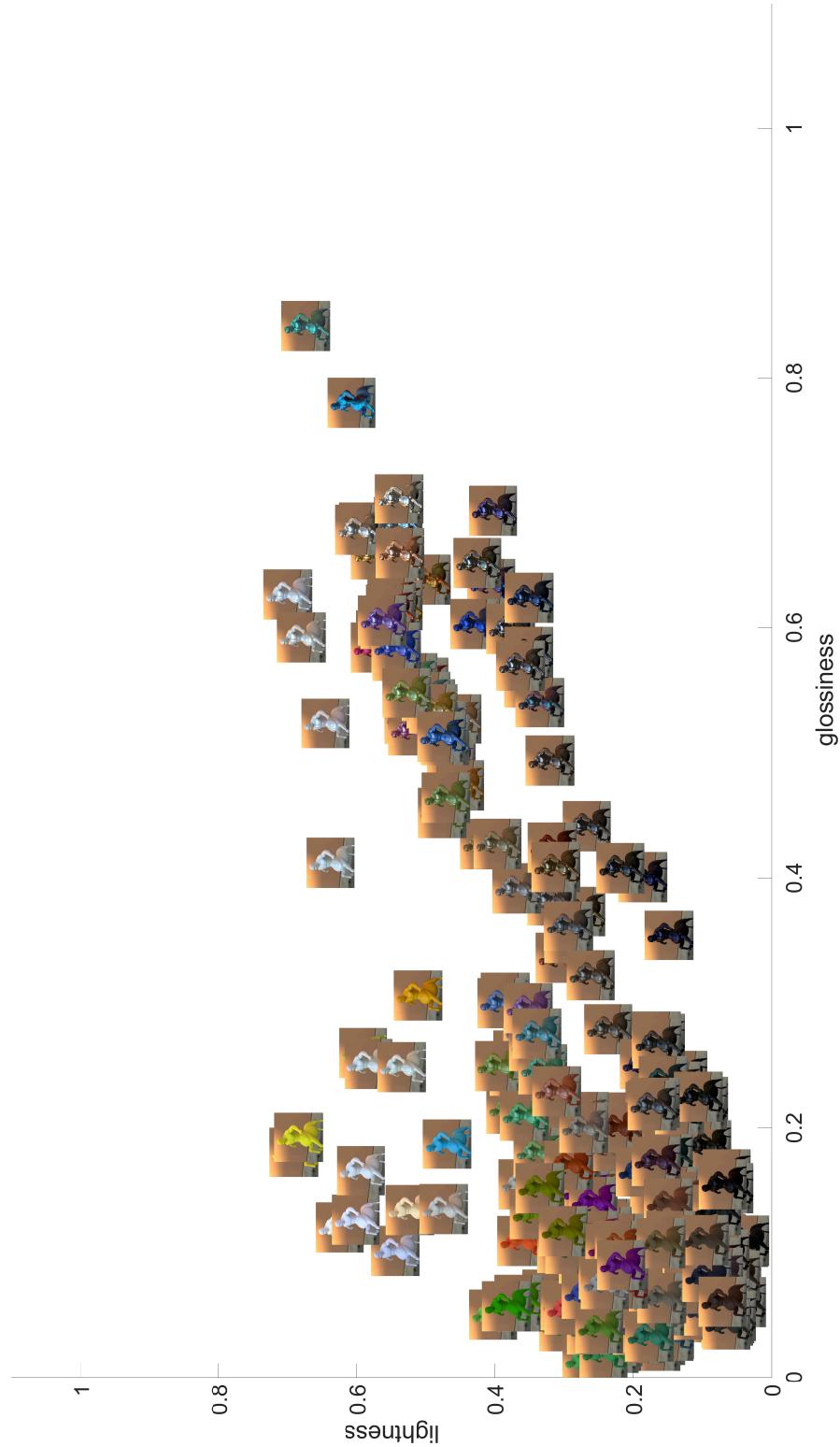


Fig. S25. Lightness-glossiness embedding generated from our predictor (pisa-centaur).

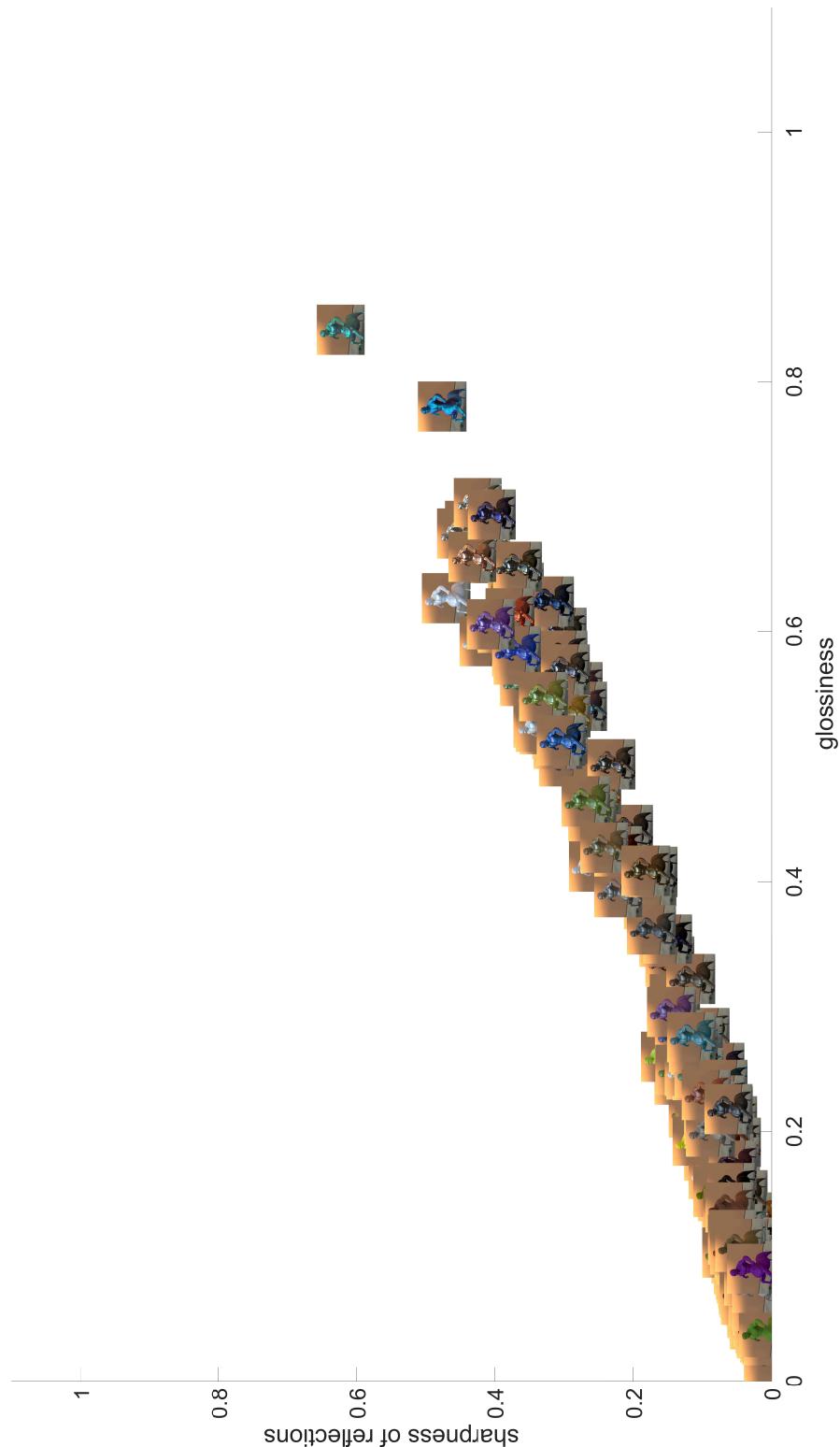


Fig. S26. Glossiness-sharpness of reflections embedding generated from our predictor (pisa-centaur).

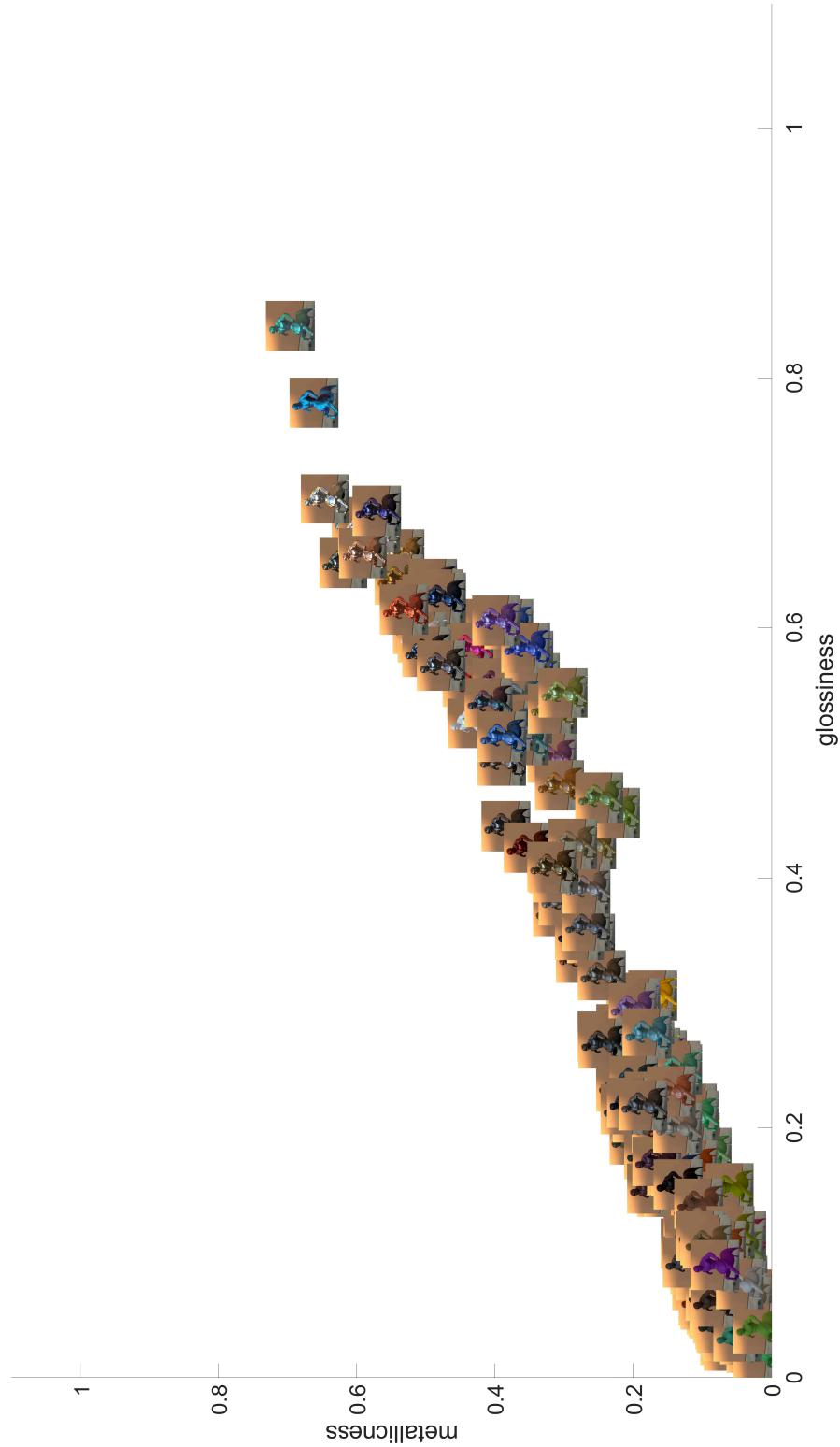


Fig. S27. Glossiness-metallicness embedding generated from our predictor (pisa-centaur).

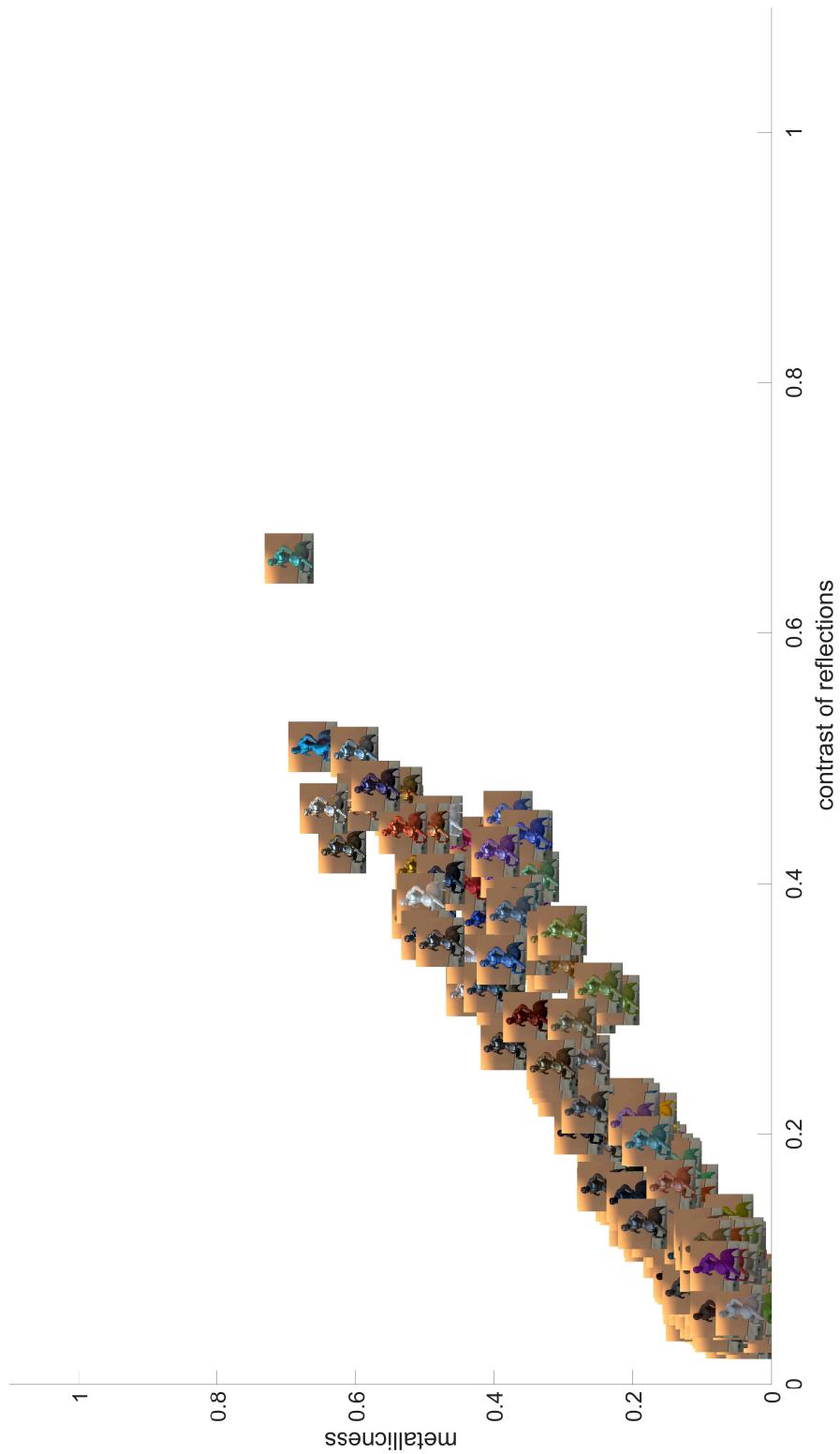


Fig. S28. Metallicness-contrast of reflections embedding generated from our predictor (pisa-centaur).

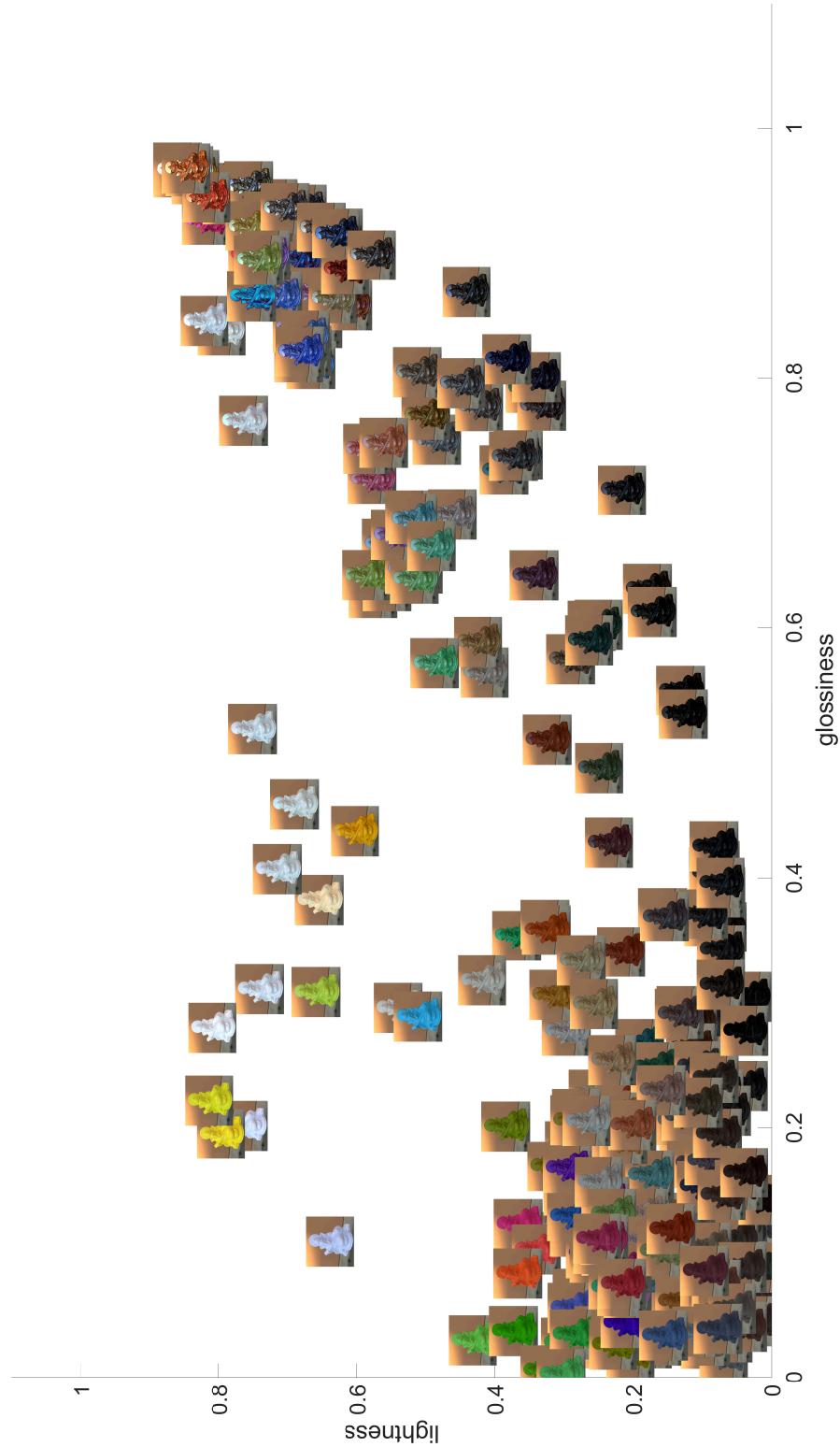


Fig. S29. Lightness-glossiness embedding generated from our predictor (pisa-ganesha).

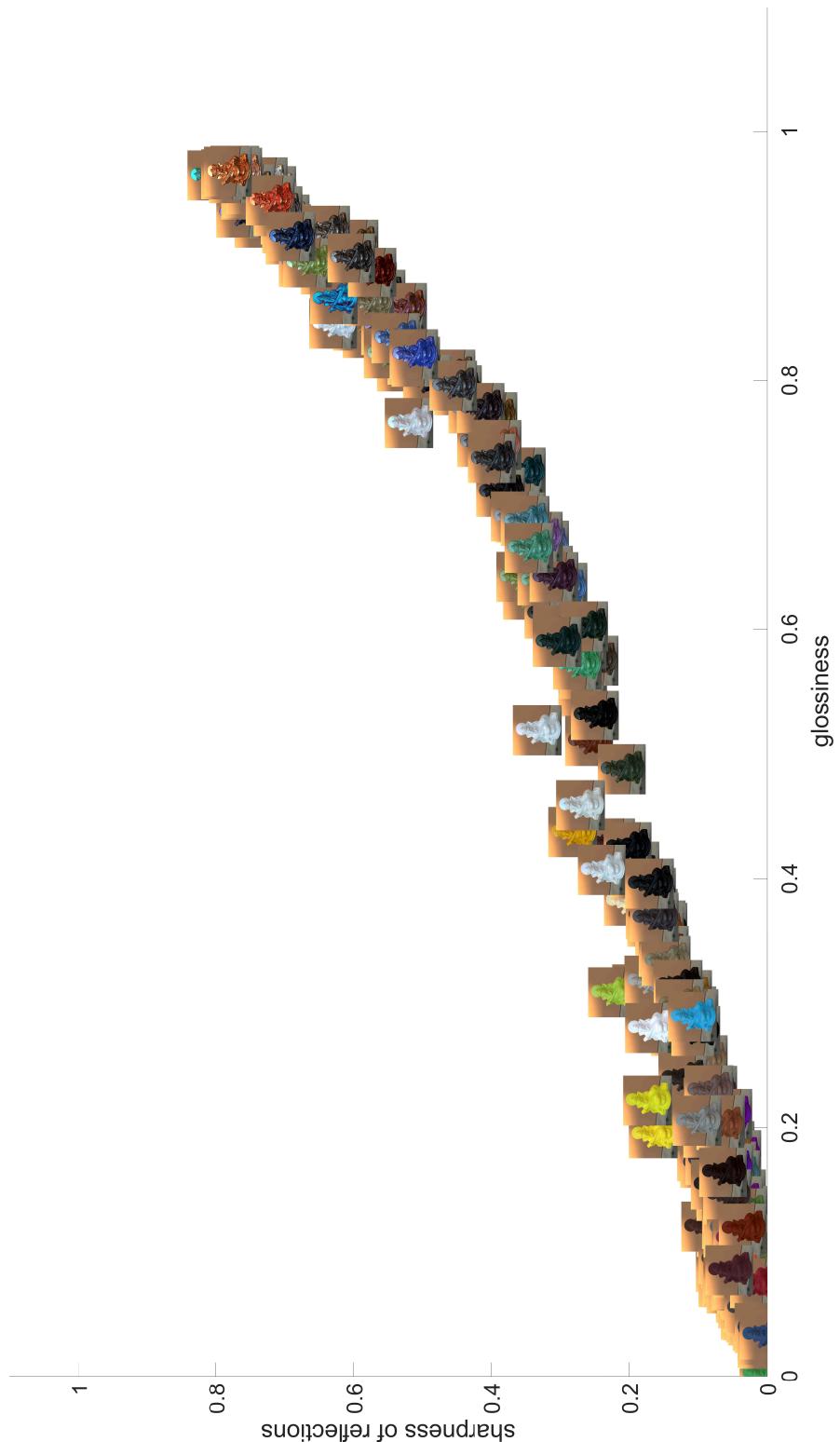


Fig. S30. Glossiness-sharpness of reflections embedding generated from our predictor (pisa-ganesha).

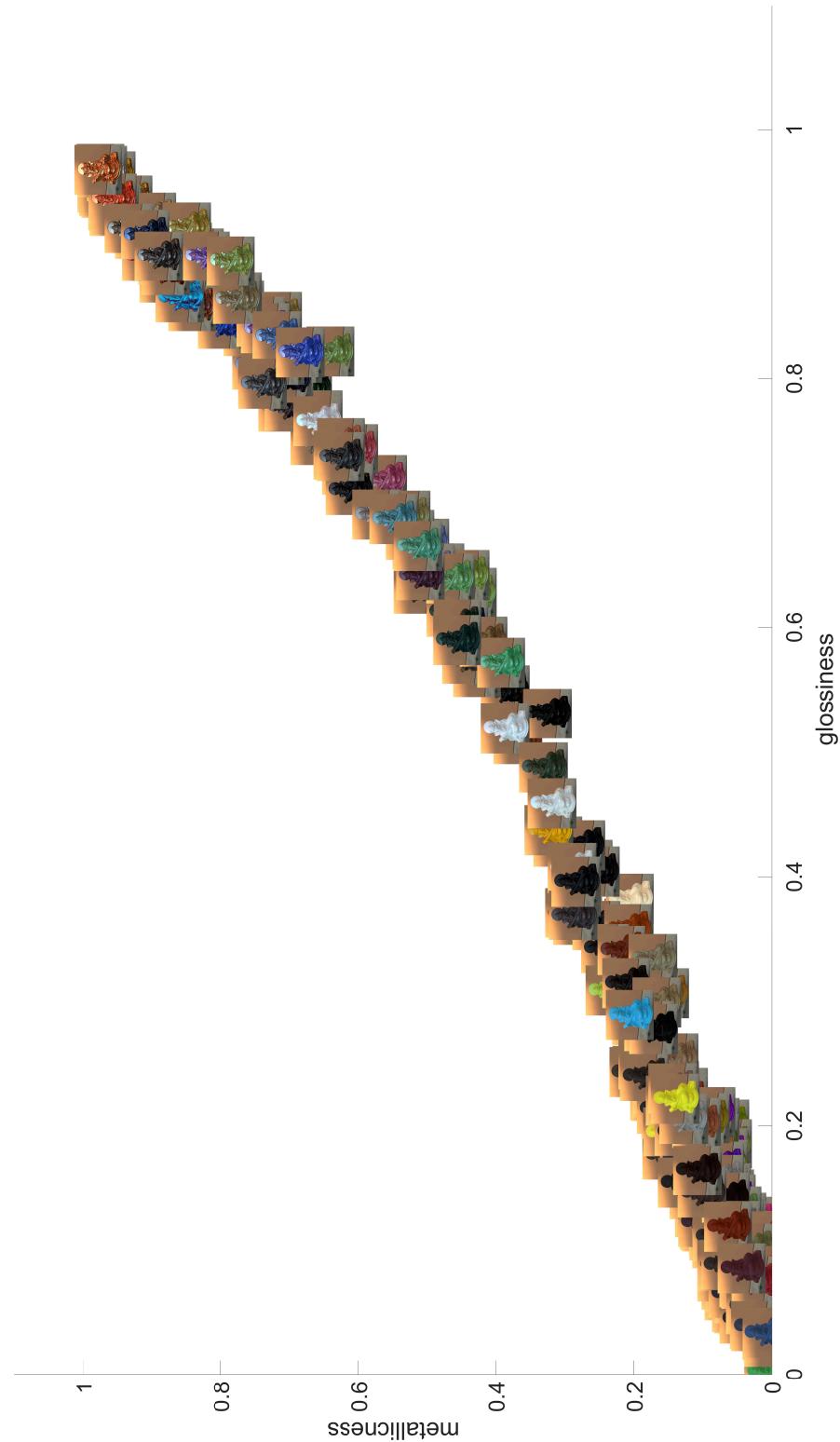


Fig. S31. Glossiness-metallicness embedding generated from our predictor (pisa-ganesha).

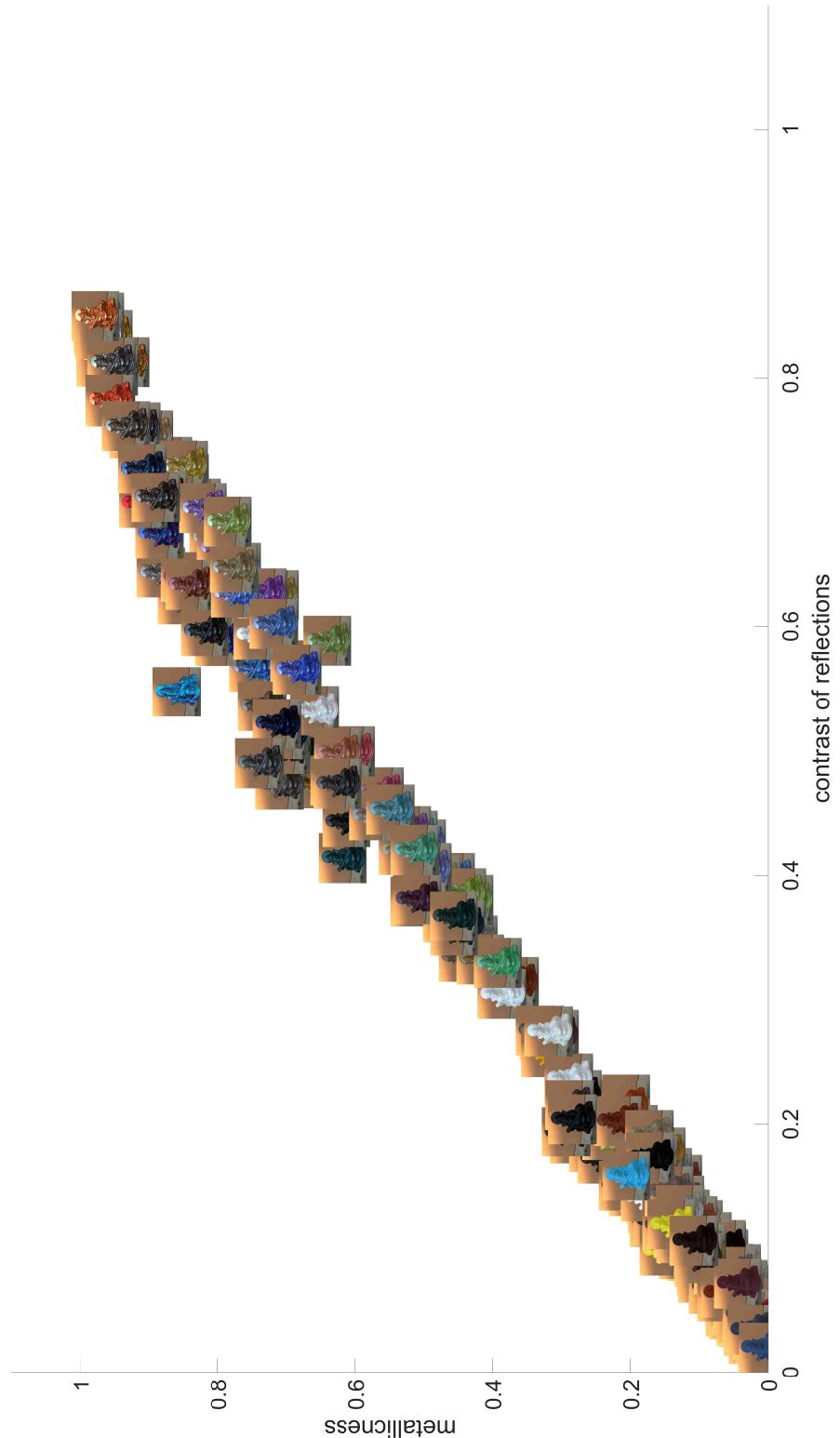


Fig. S32. Metallicness-contrast of reflections embedding generated from our predictor (pisa-ganesh).

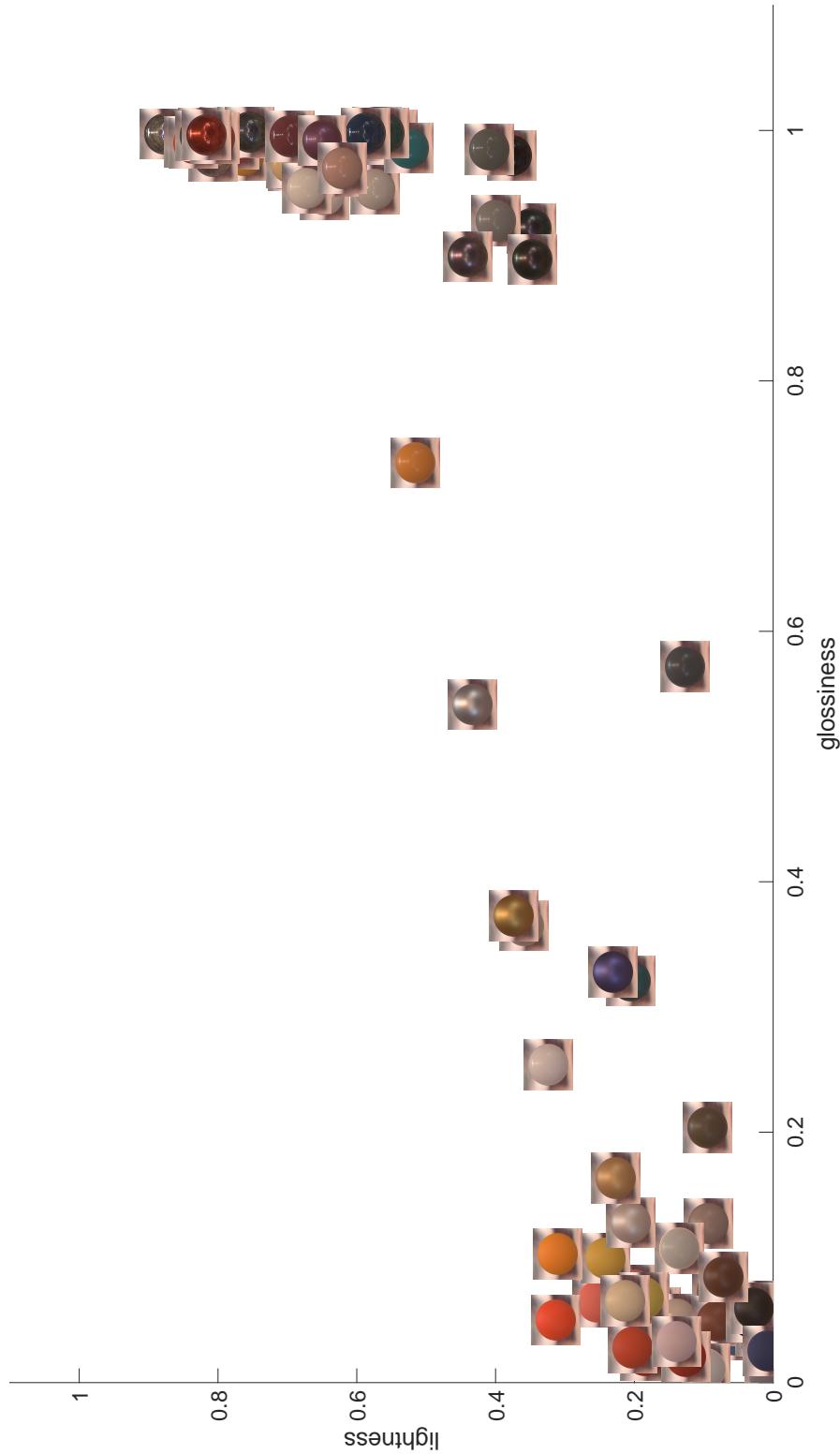


Fig. S33. Lightness-glossiness embedding generated from our predictor (pisa-sphere).

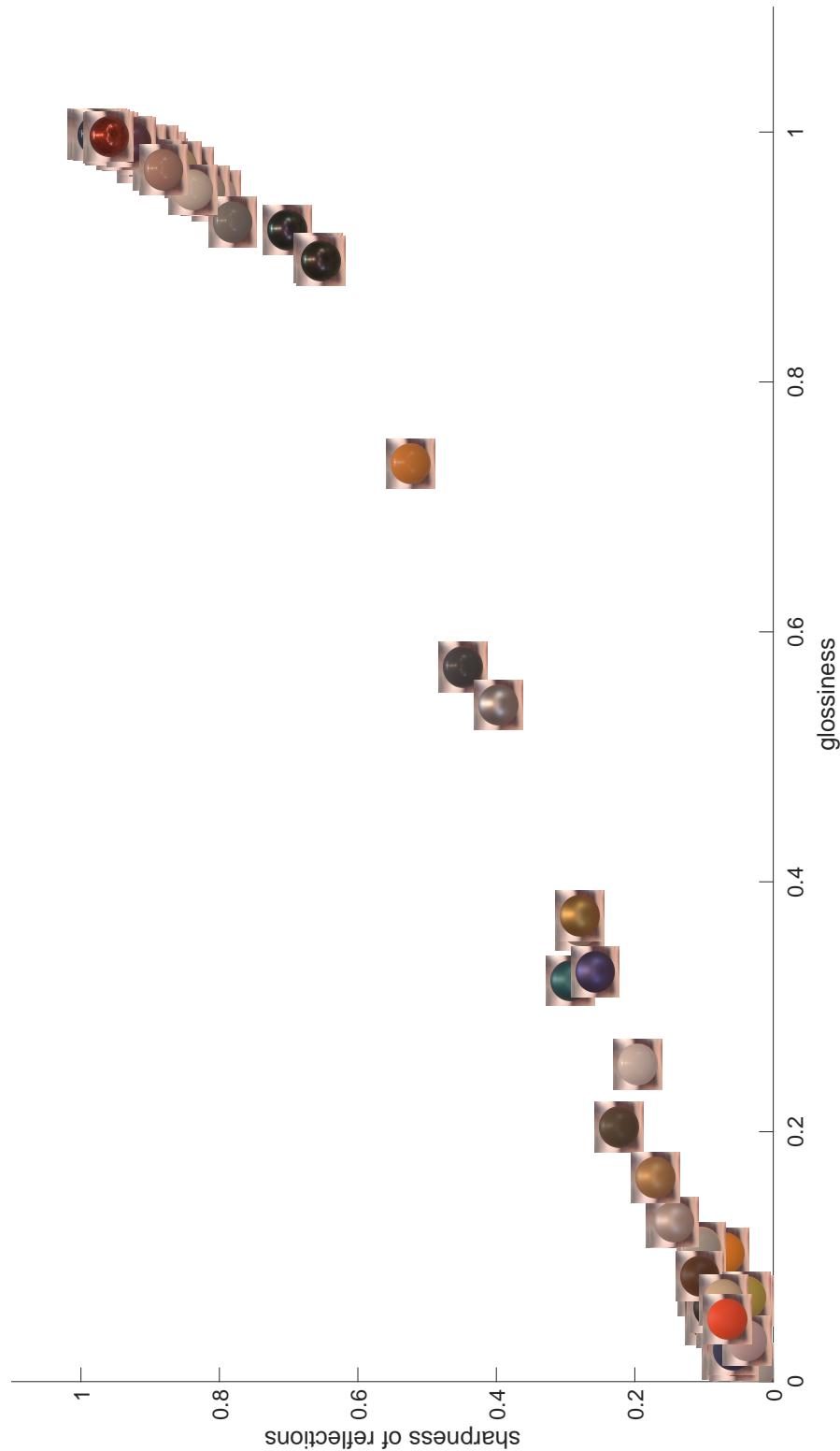


Fig. S34. Glossiness-sharpness of reflections embedding generated from our predictor (pisa-sphere).

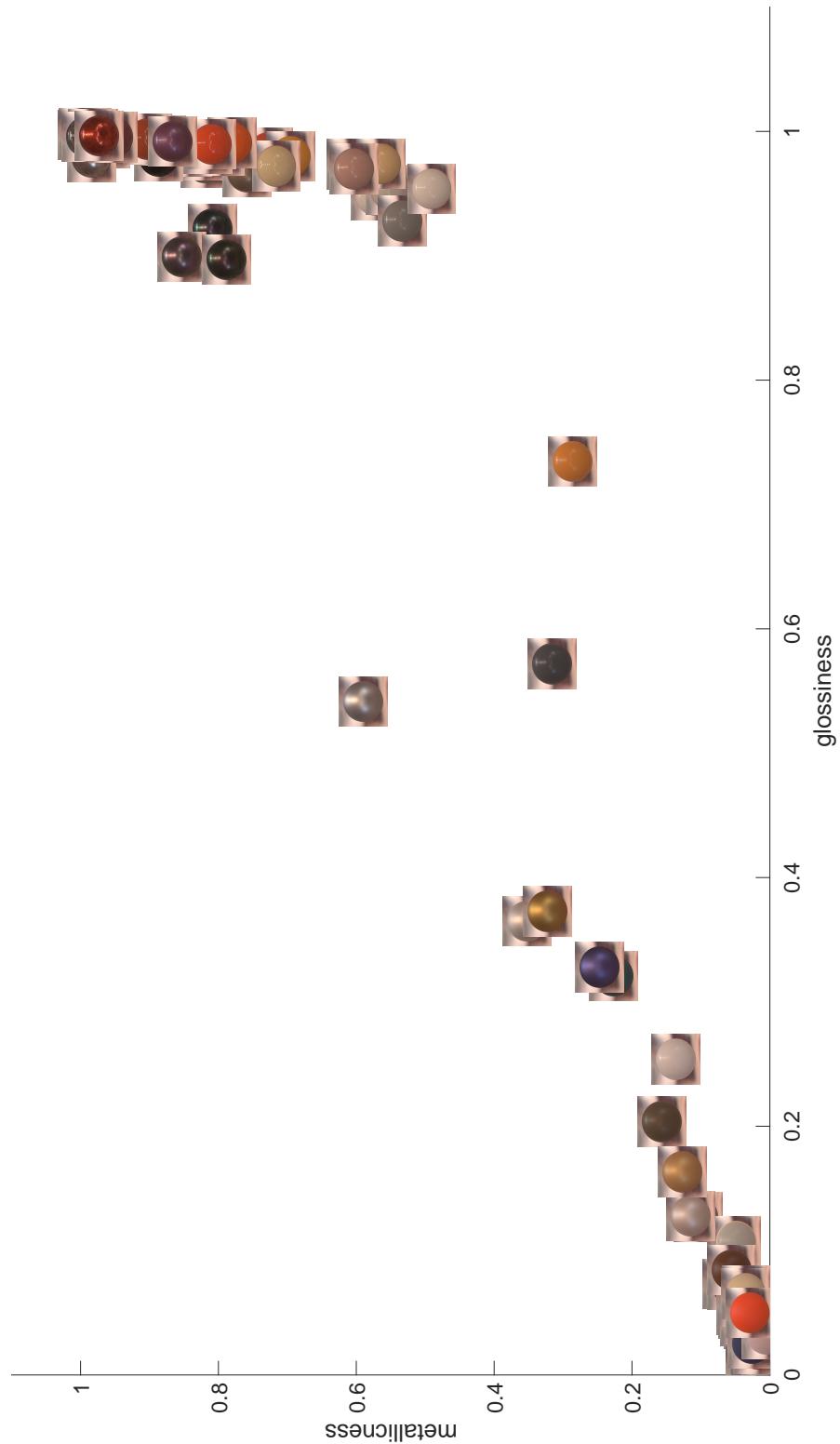


Fig. S35. Glossiness-metallicness embedding generated from our predictor (pisa-sphere).

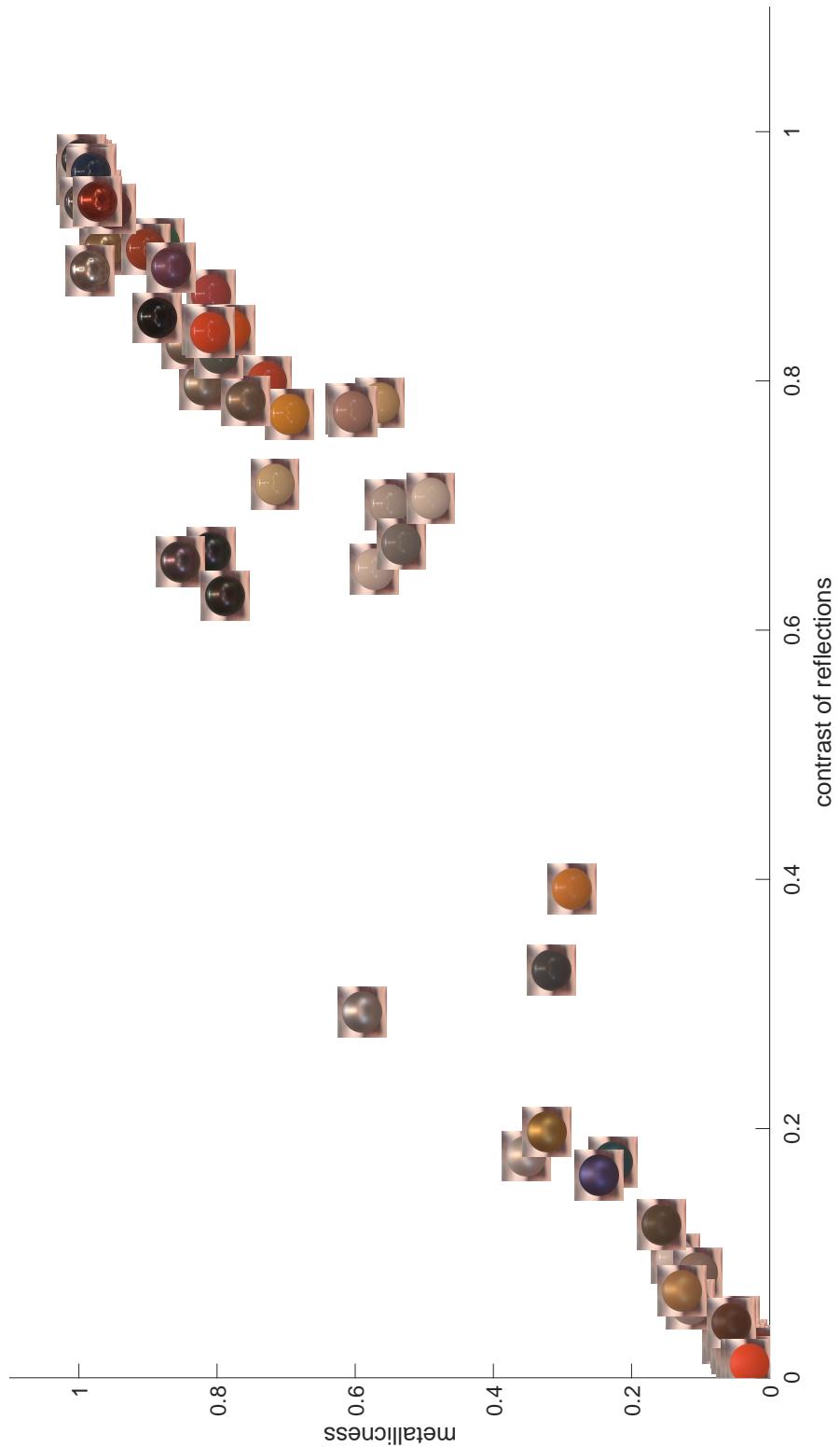


Fig. S36. Metallicness-contrast of reflections embedding generated from our predictor (pisa-sphere).

S 4. STATISTICAL ANALYSIS

In this section we include additional information about the data and fitted models for the statistical analysis described in Section 4 in the main paper. In particular, we include the collected raw data, correlations, fitted cumulative mixed models, predicted probabilities for each rating, and the estimated mixed effects of user and material.

4.1 Collected data

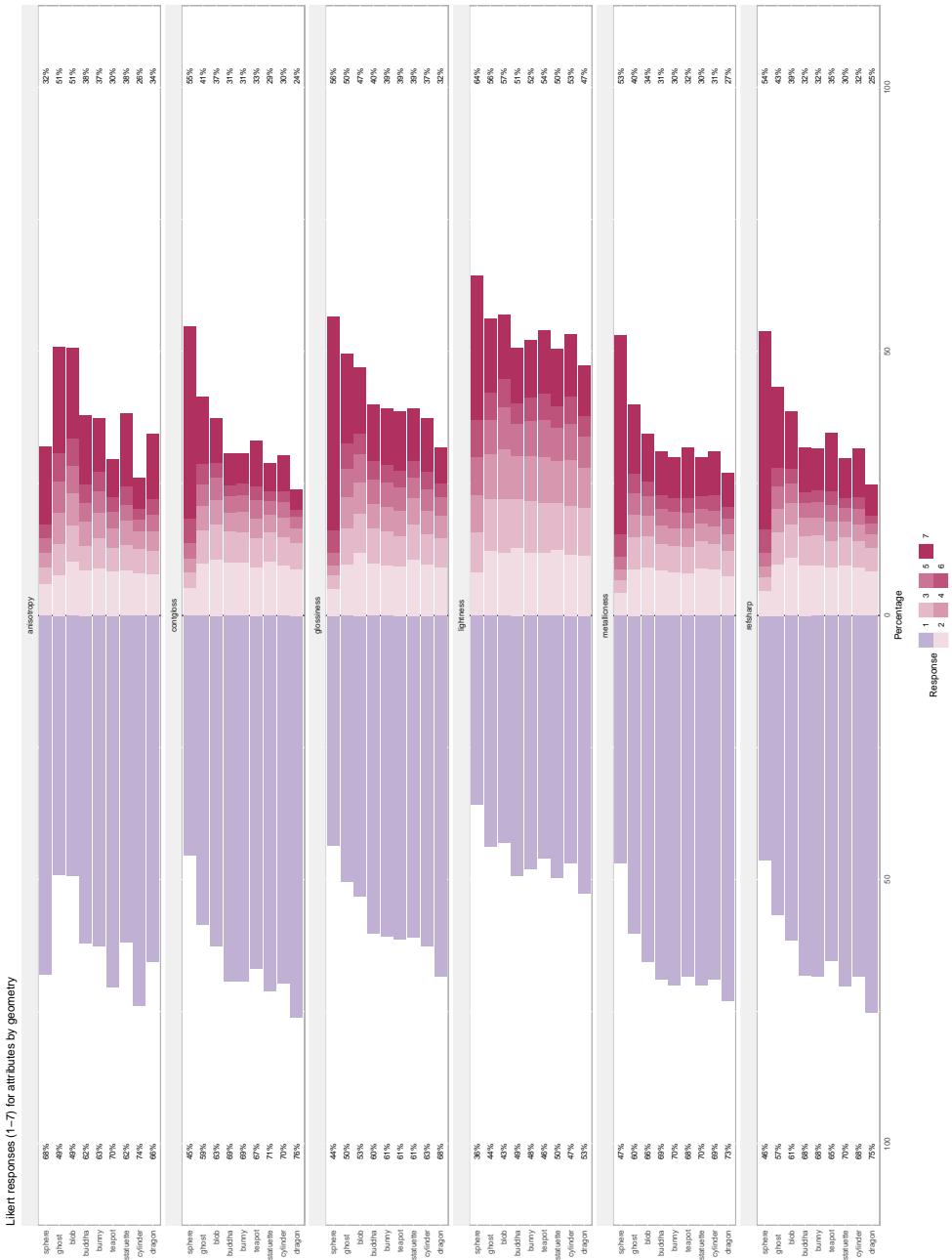


Fig. S37. Answers distribution by geometry for each of our six attributes.

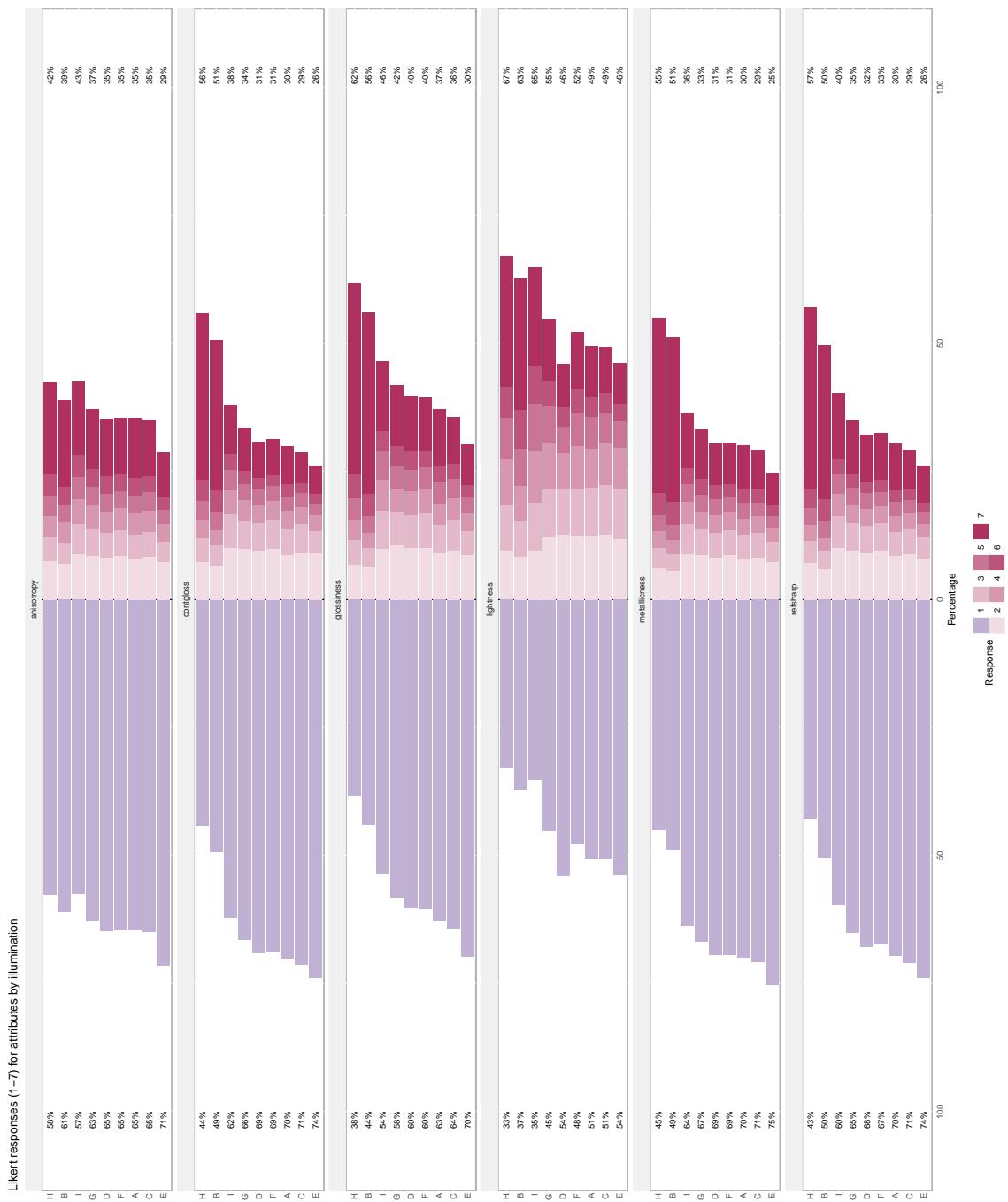


Fig. S38. Answers distribution by illumination for each of our six attributes.

Relationship between attributes across illuminations

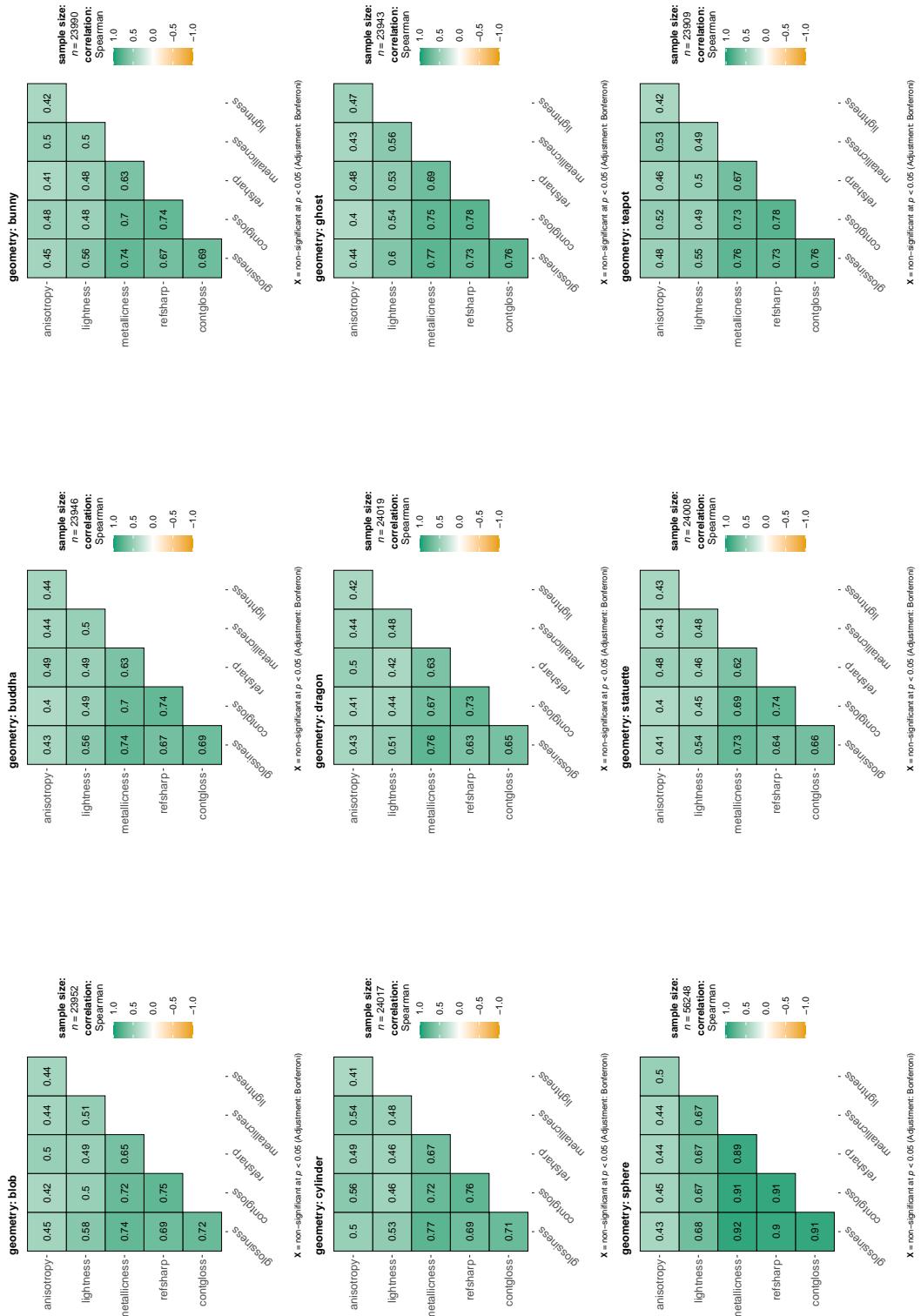


Fig. S39. Correlations between our attributes by geometry.

Relationship between attributes across illuminations

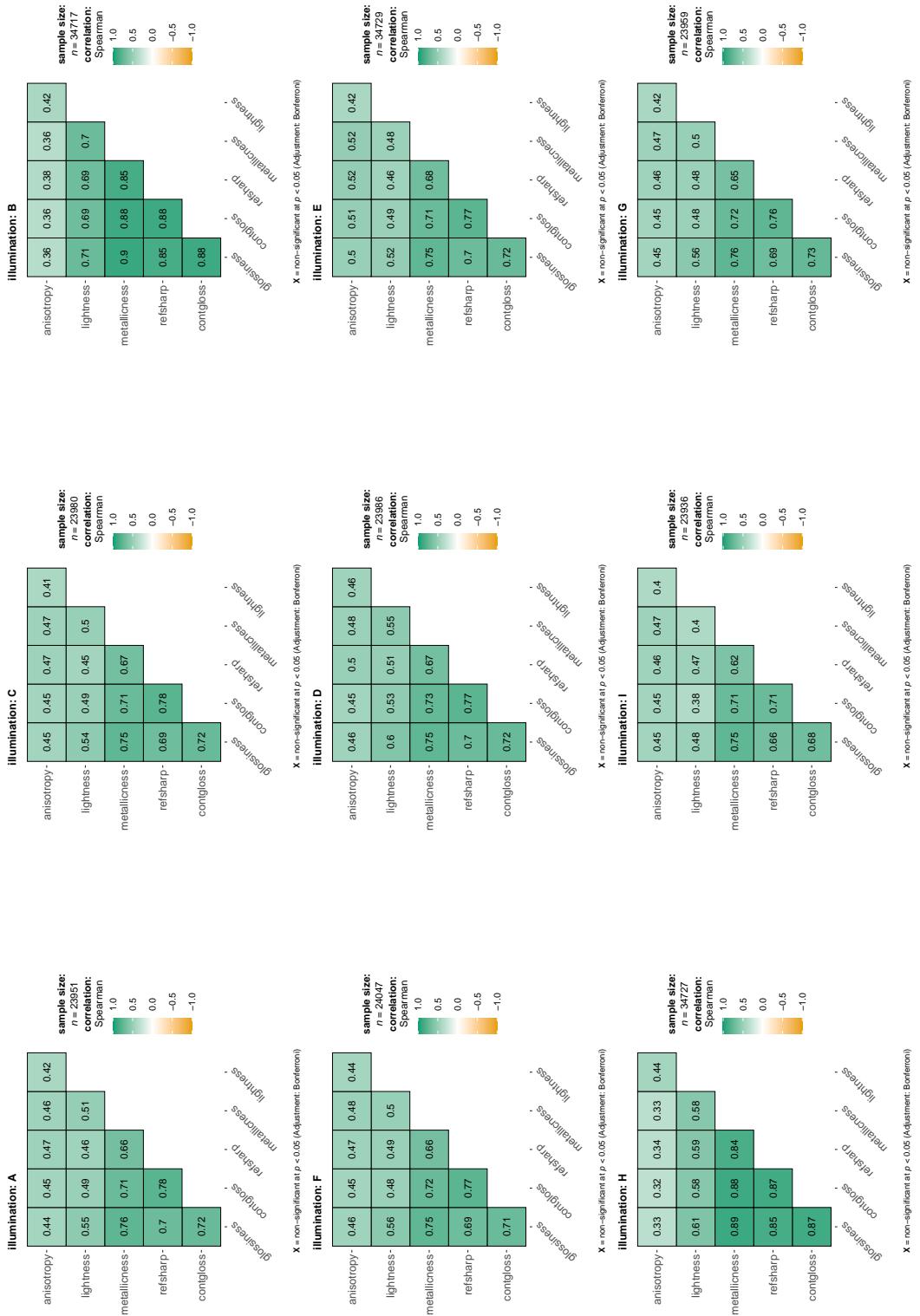


Fig. S40. Correlations between our attributes by illumination.

4.2 Models for each attribute

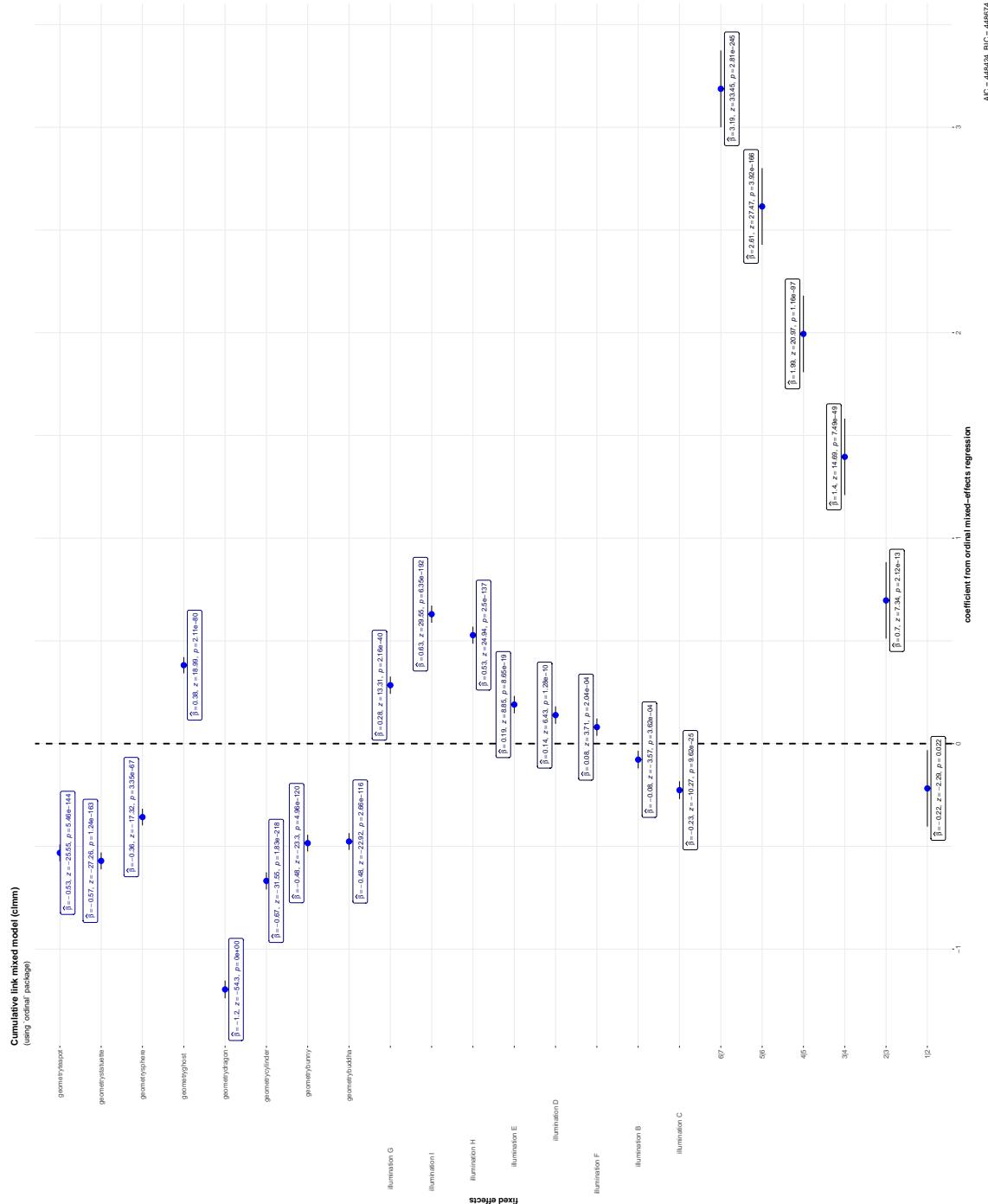


Fig. S41. Model for the glossiness attribute.

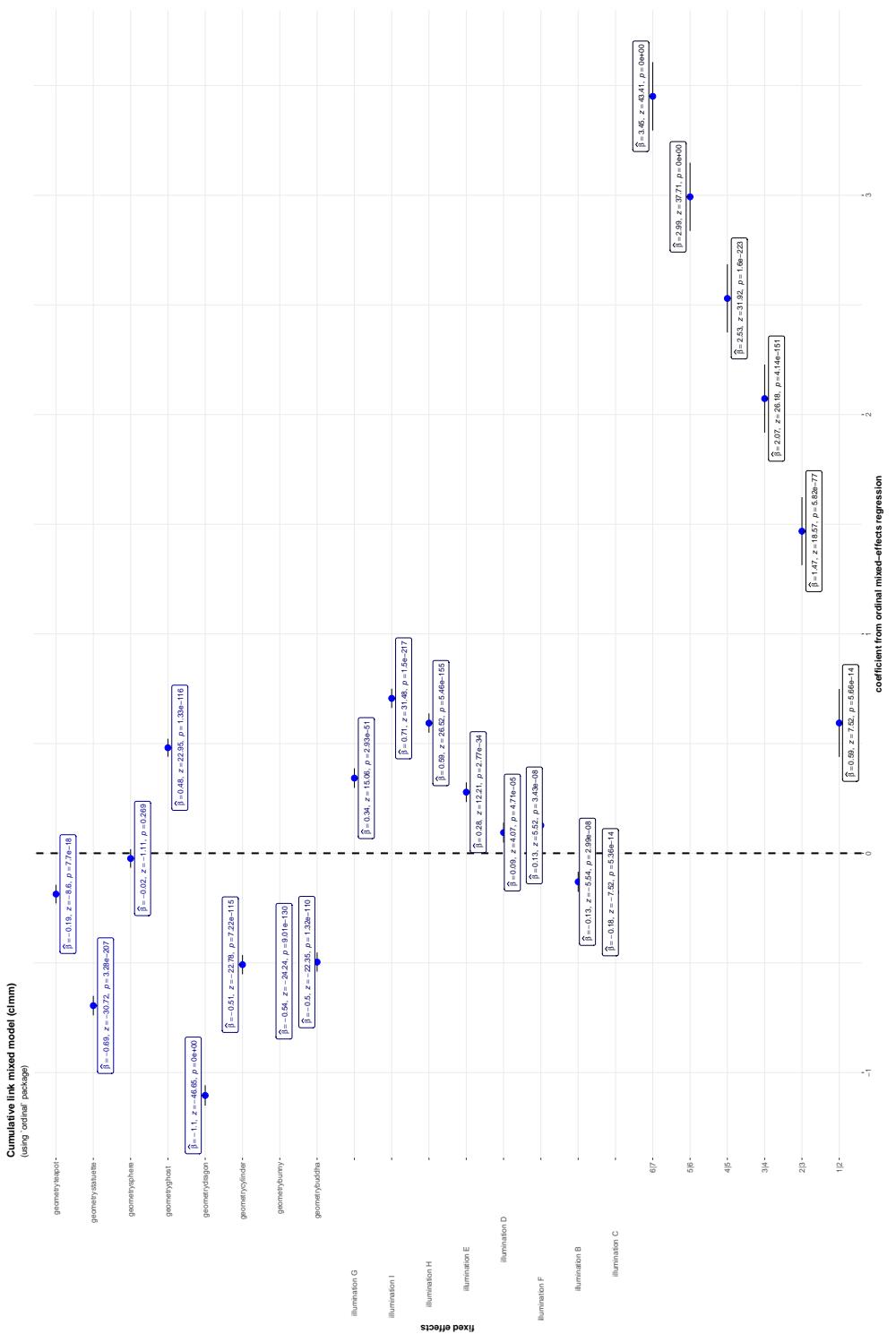


Fig. S42. Model for the sharpness of reflections attribute.

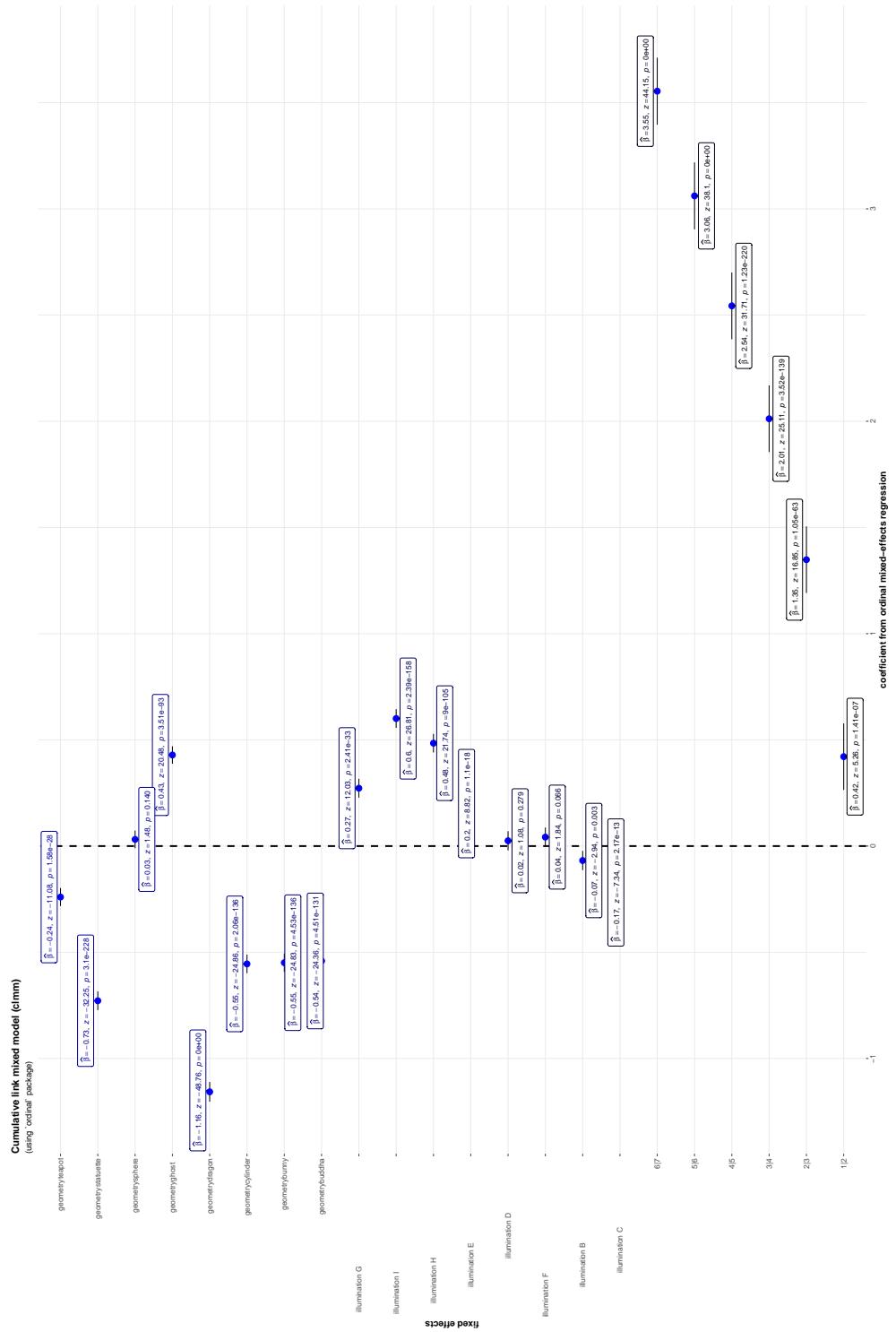


Fig. S43. Model for the contrast gloss attribute.

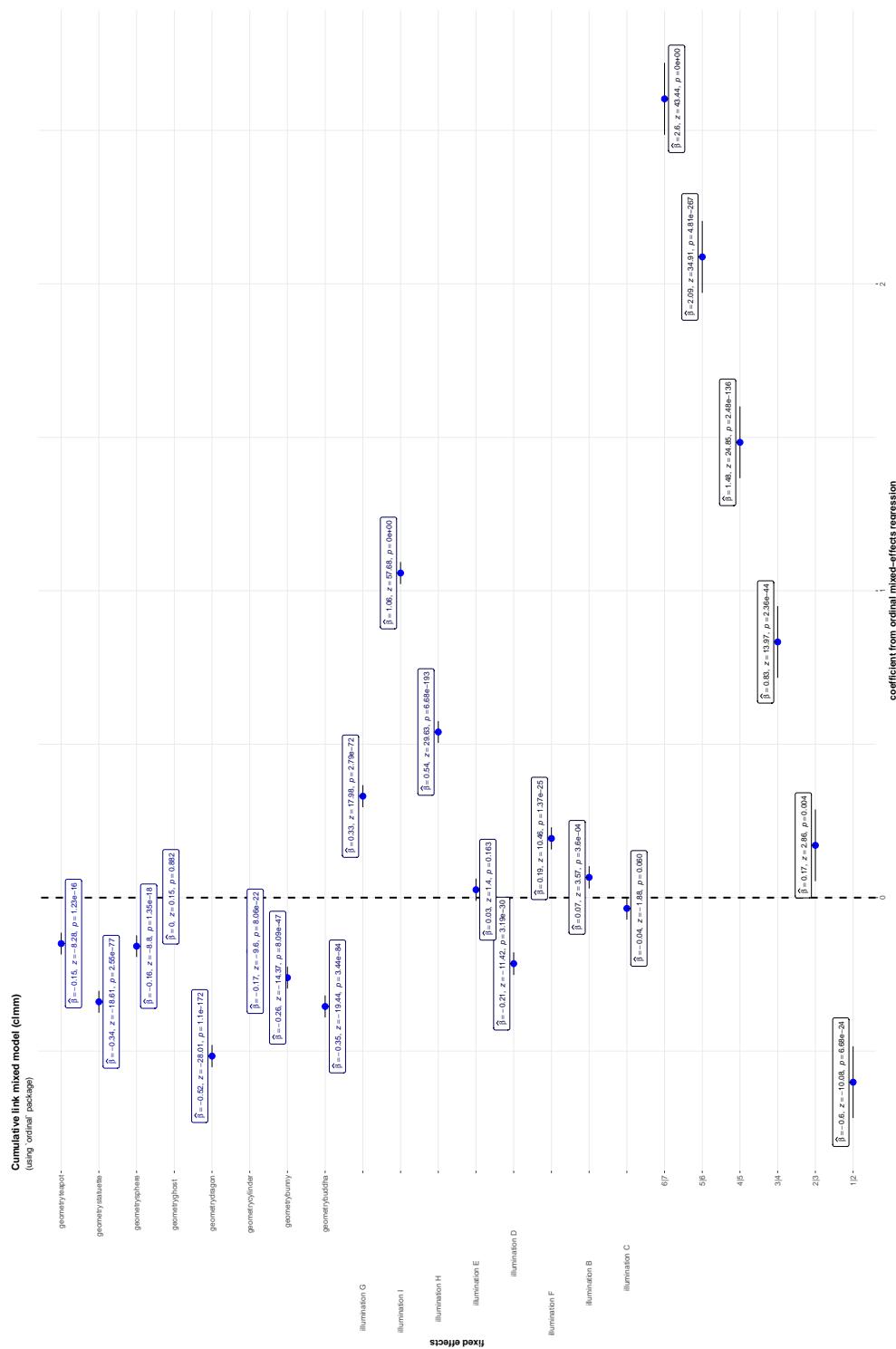


Fig. S44. Model for the light attribute.

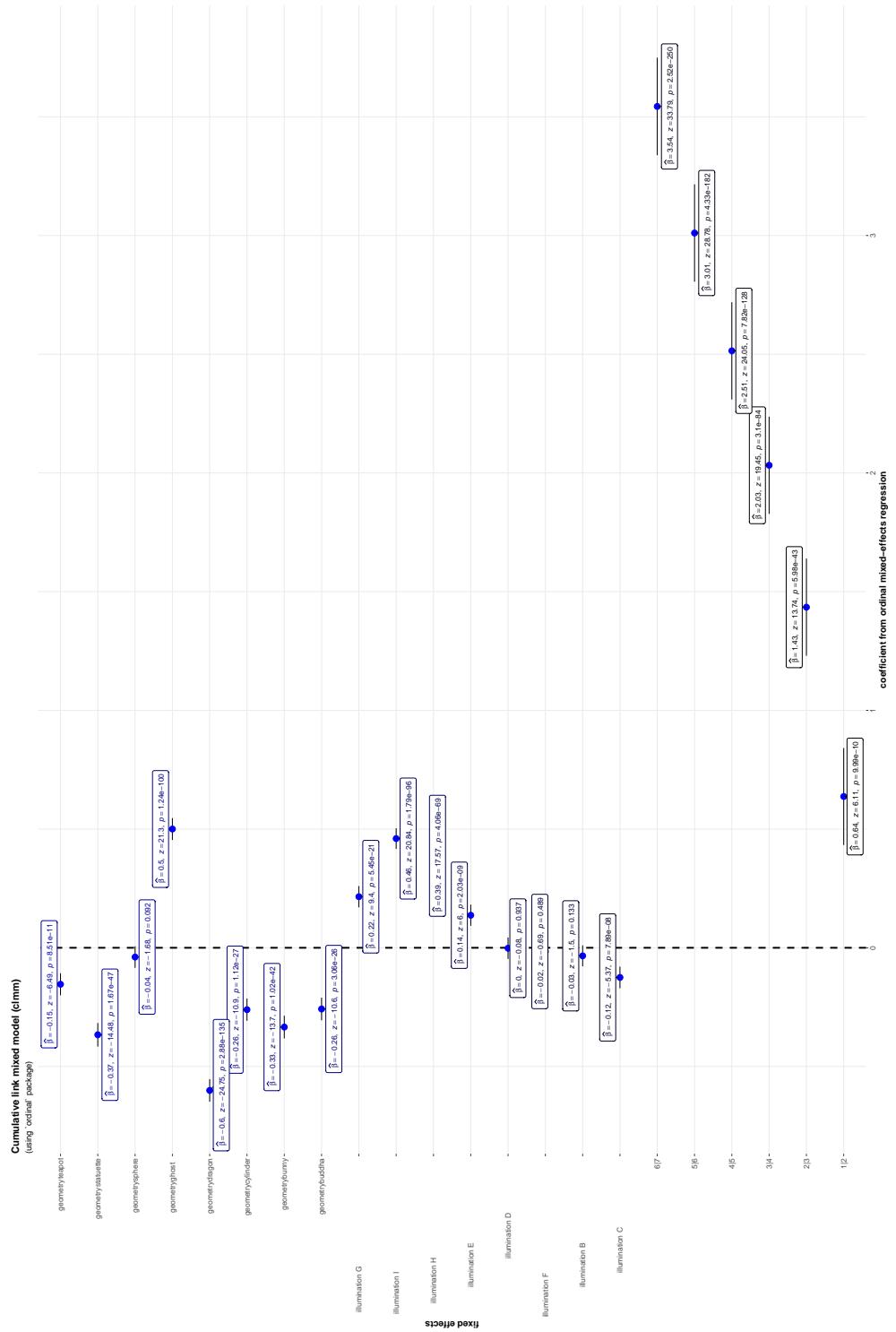


Fig. S45. Model for the metallic attribute.

4.3 Analysis

4.3.1 Glossiness

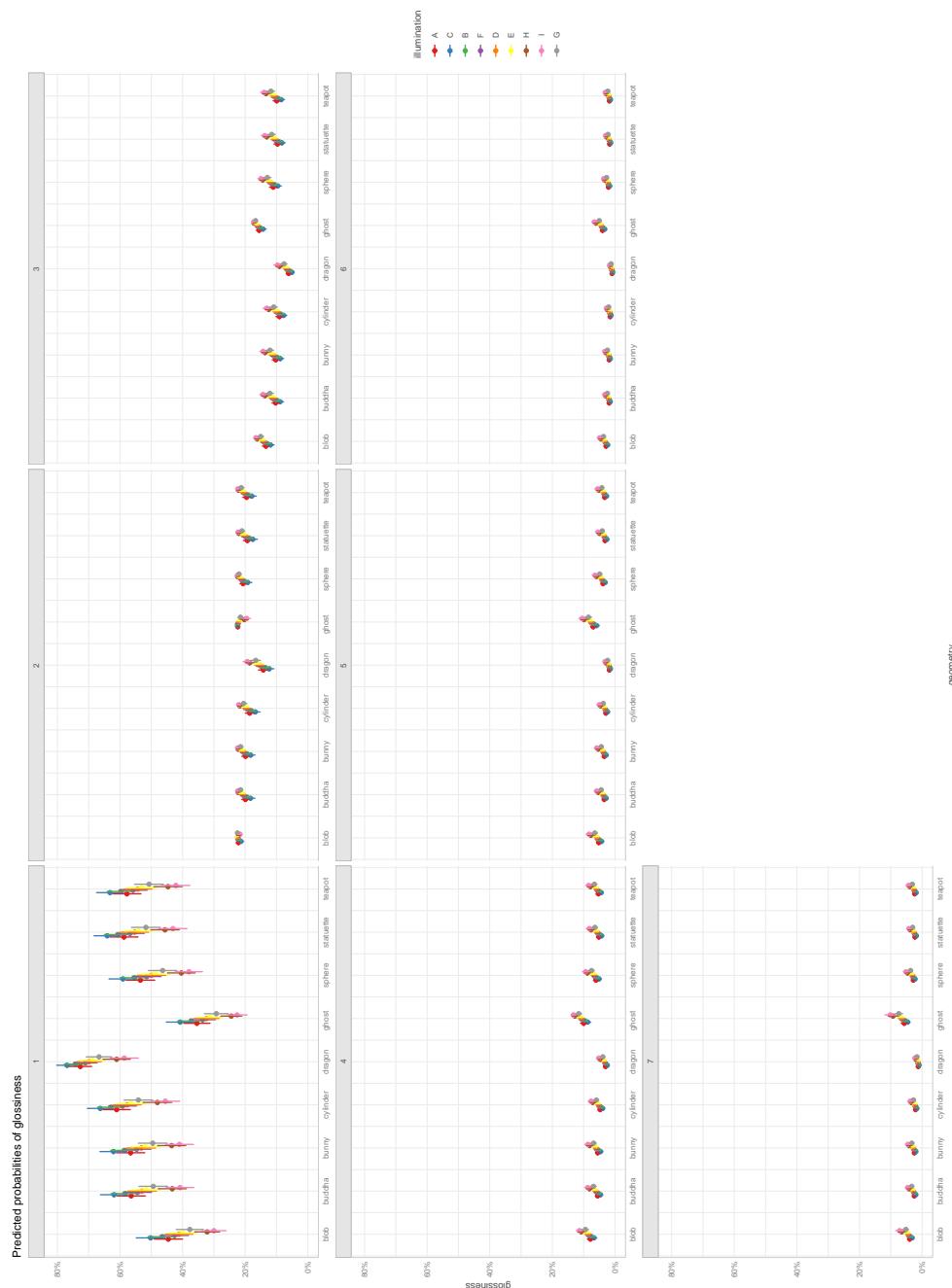
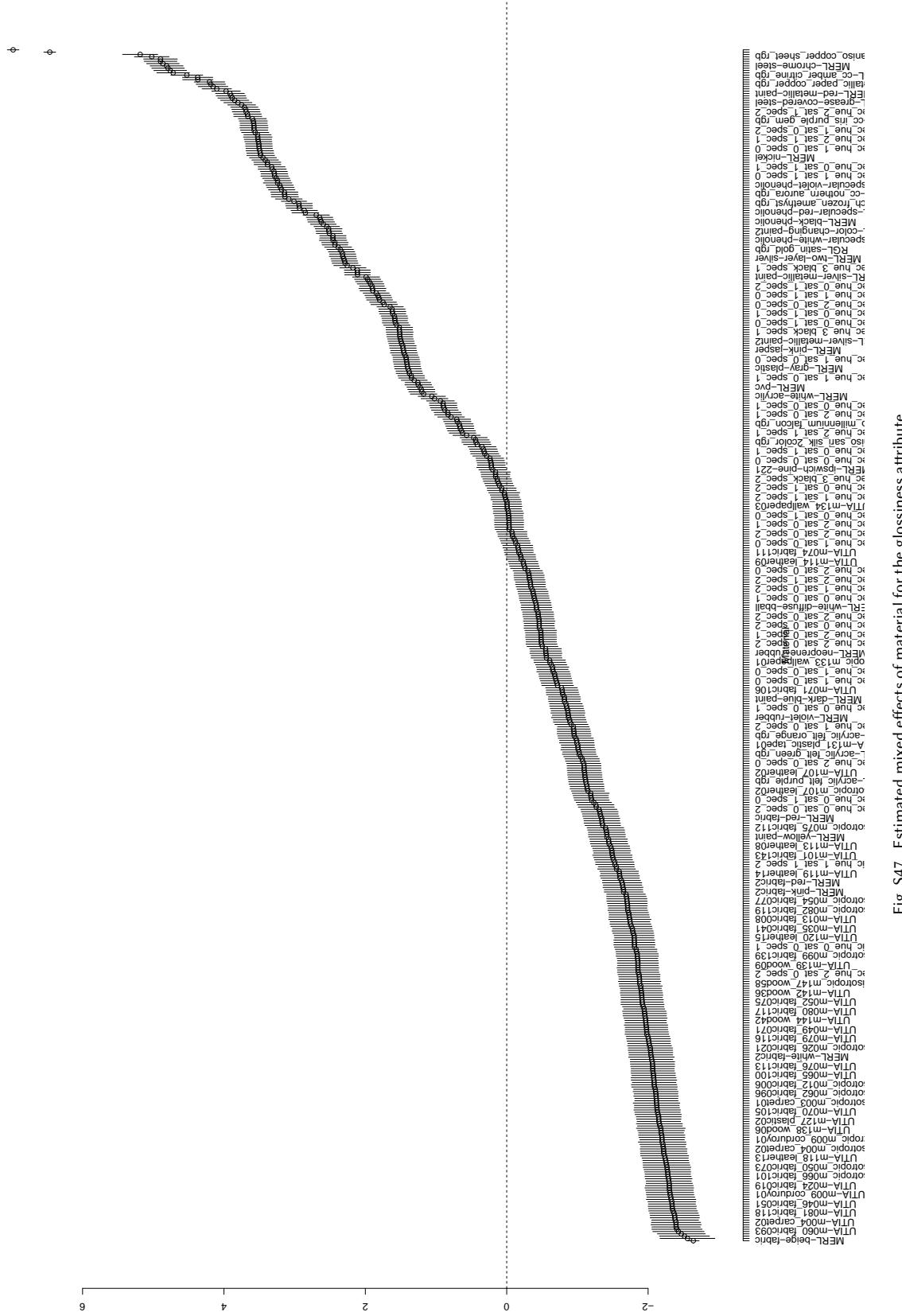


Fig. S46. Predicted probabilities for glossiness as a function of geometry and illumination.



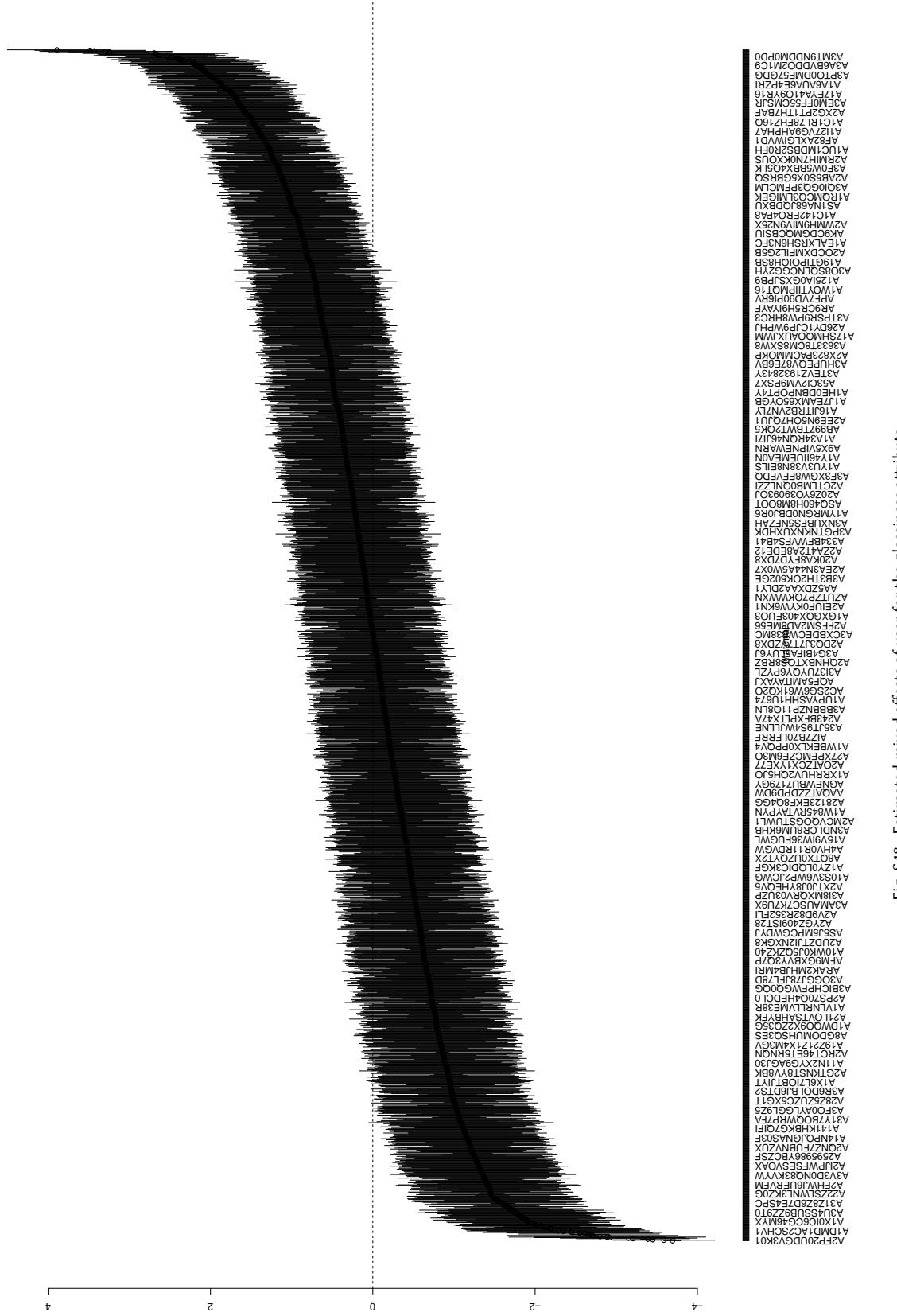


Fig. S48. Estimated mixed effects of user for the glossiness attribute.

4.3.2 Sharpness of reflections



Fig. S49. Predicted probabilities for sharpness of reflections as a function of geometry and illumination.

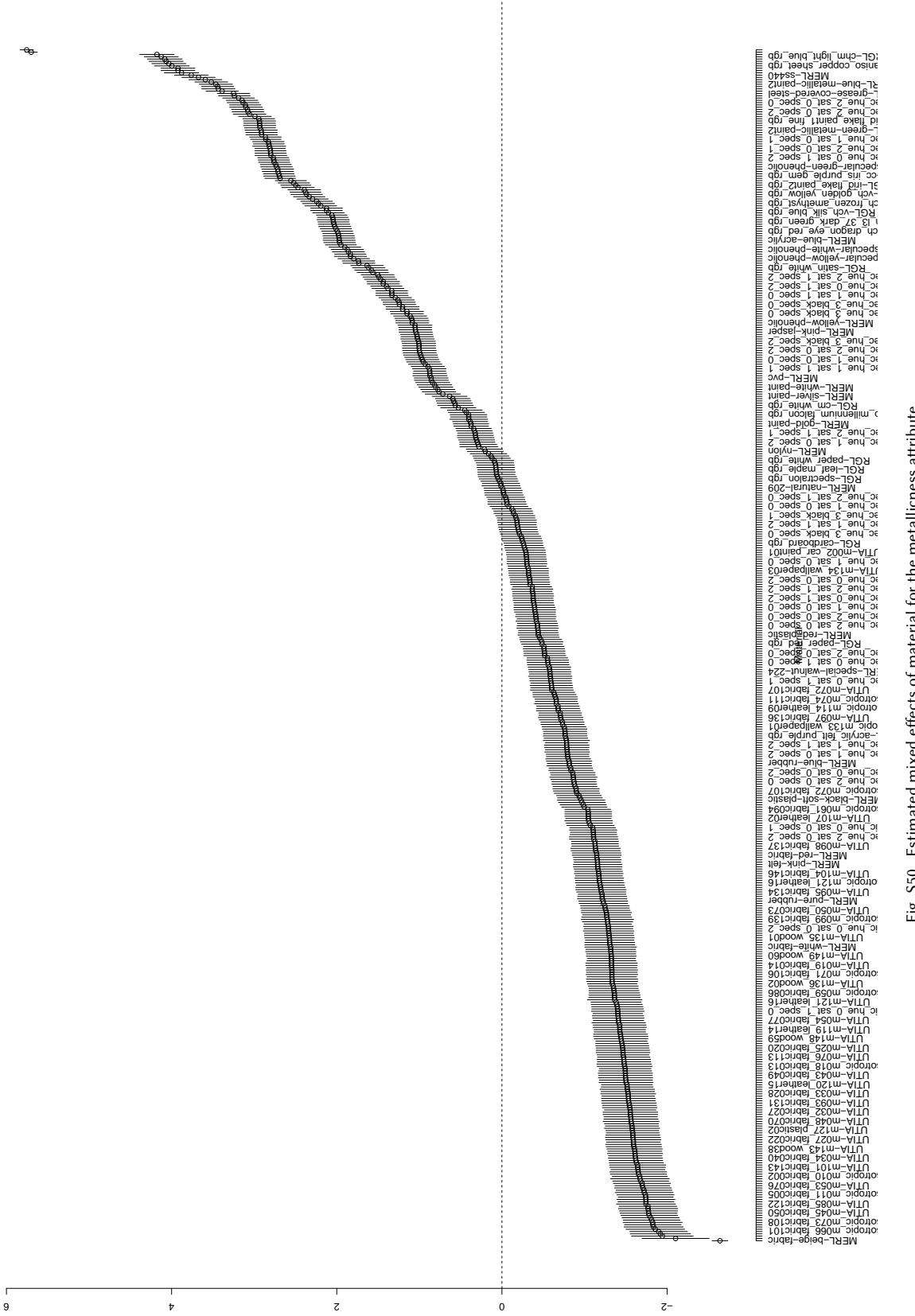


Fig. S50. Estimated mixed effects of material for the metallineness attribute.

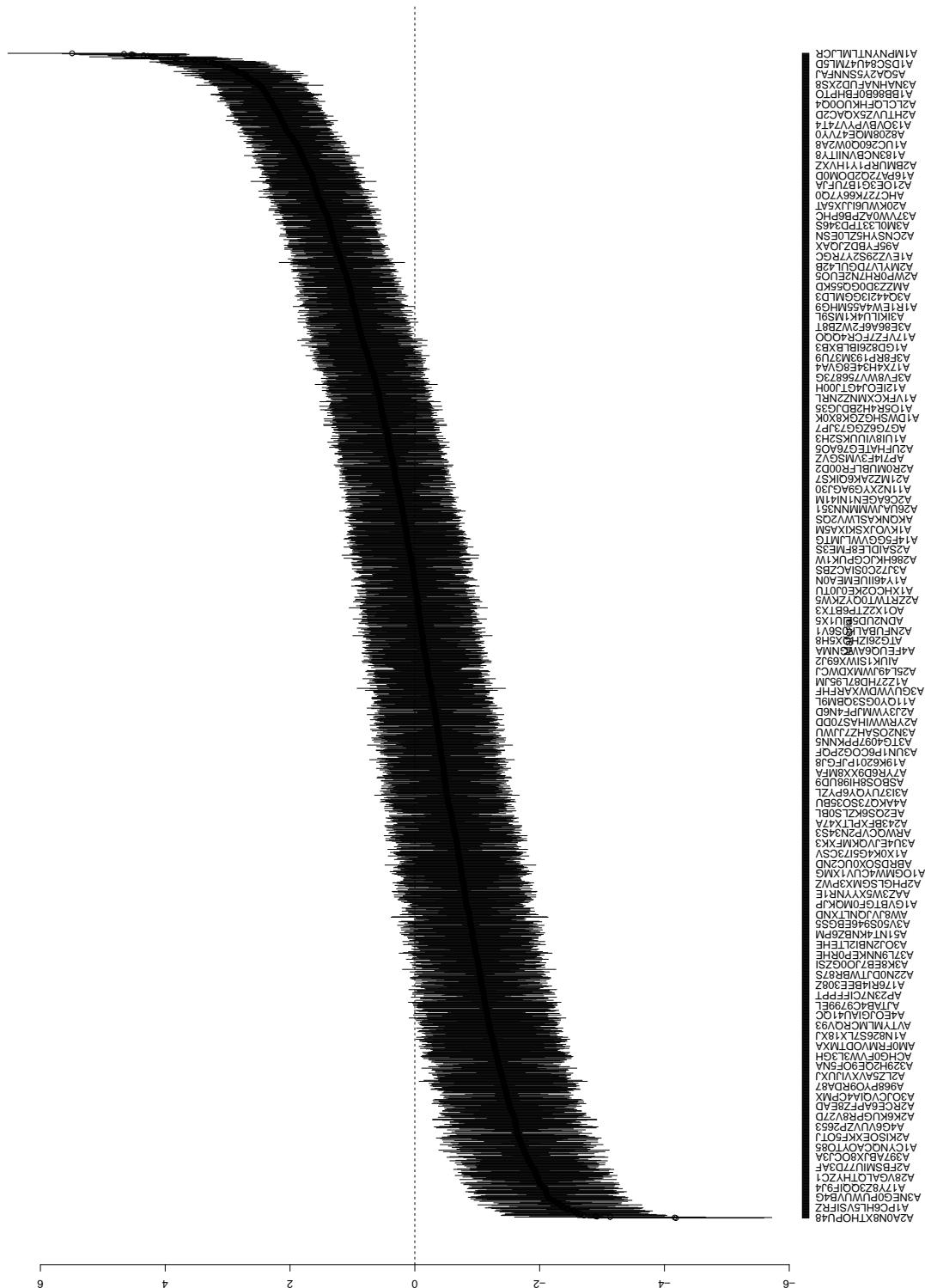


Fig. S51. Estimated mixed effects of user for the sharpness of reflections attribute.

4.3.3 Contrast of reflections



Fig. S52. Predicted probabilities for contrast of reflections as a function of geometry and illumination.

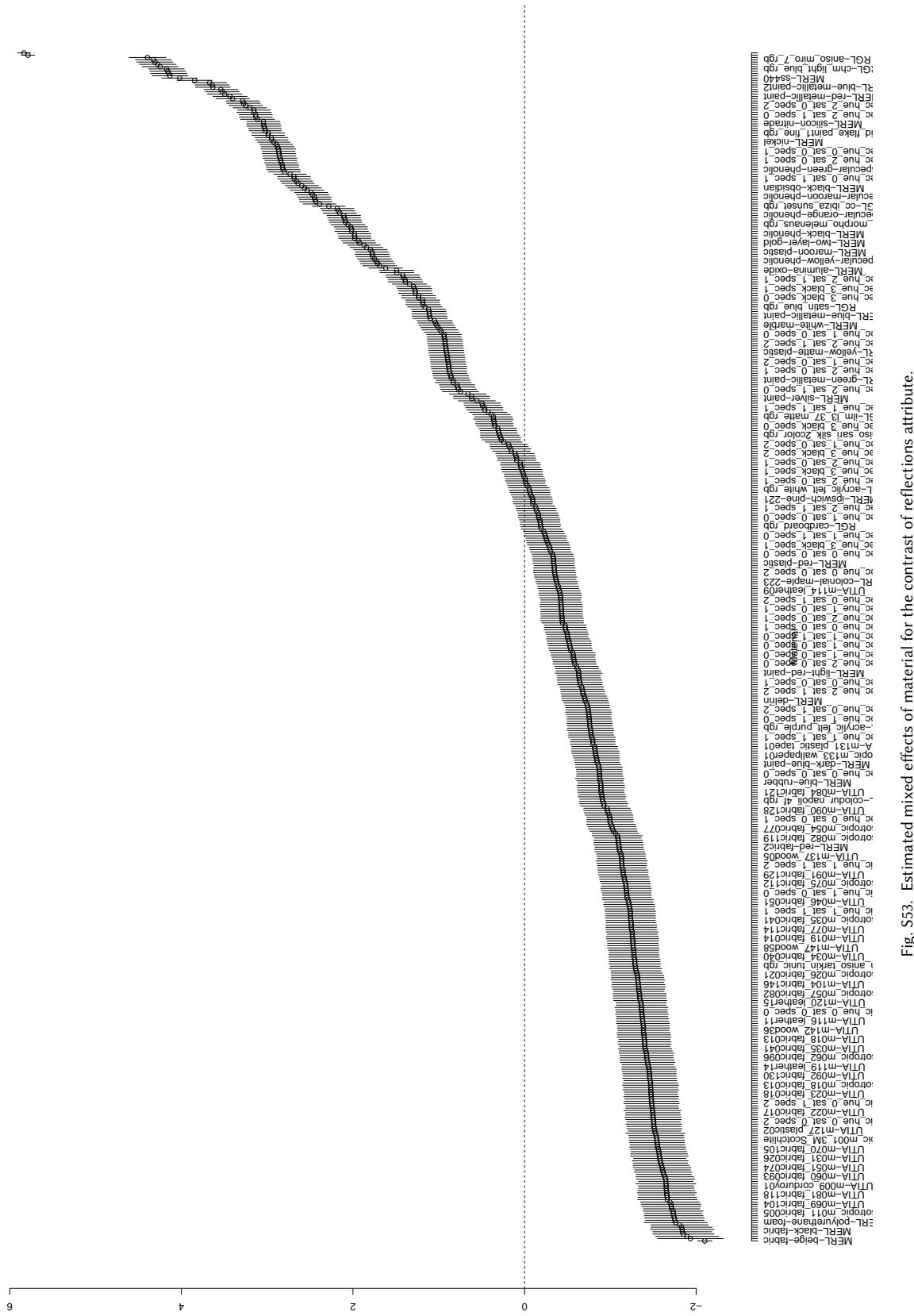


Fig. S53. Estimated mixed effects of material for the contrast of reflections attribute.



Fig. S54. Estimated mixed effects of user for the contrast of reflections attribute.

4.3.4 Lightness



Fig. S55. Predicted probabilities for lightness as a function of geometry and illumination.

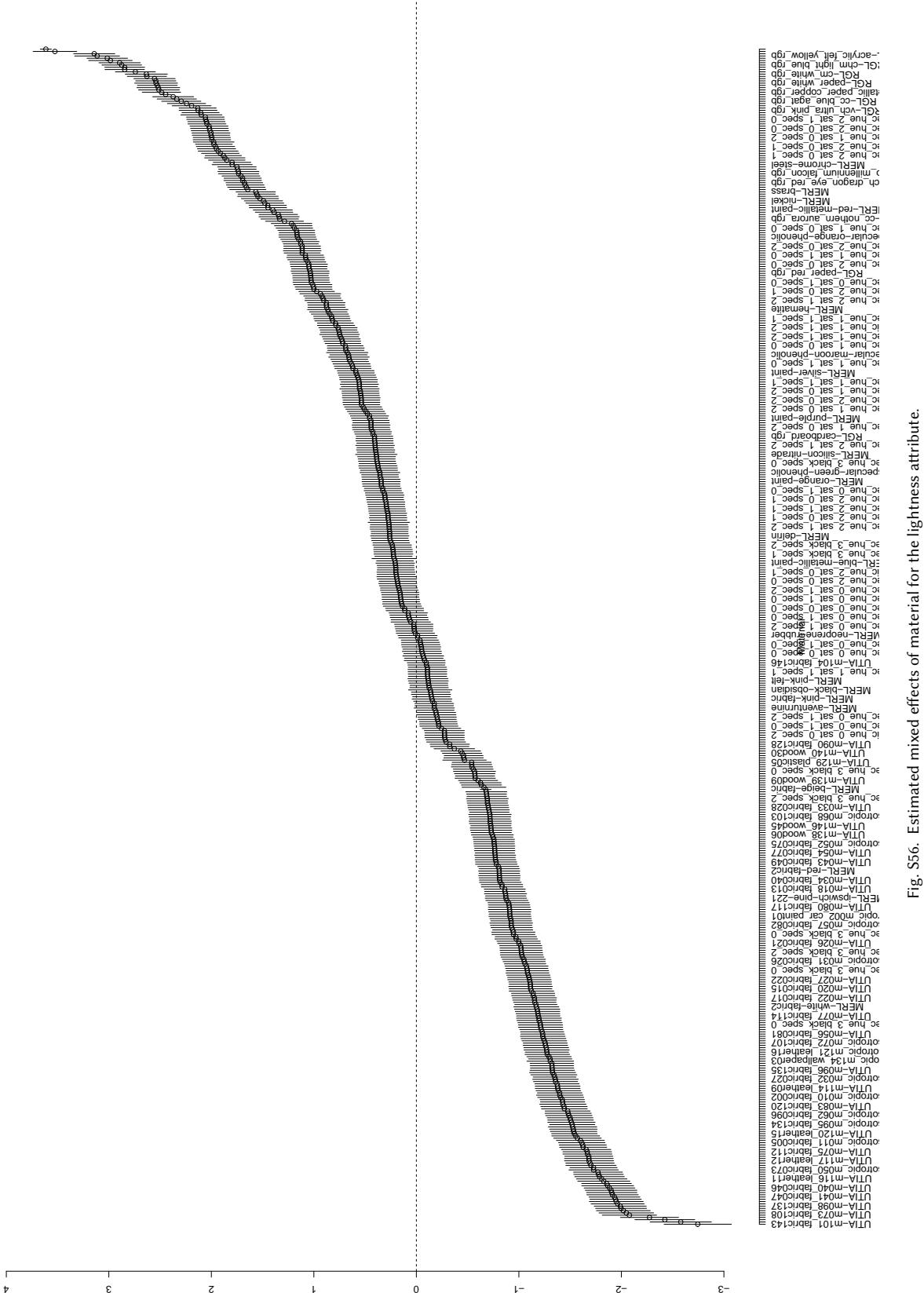


Fig. S56. Estimated mixed effects of material for the lightness attribute.

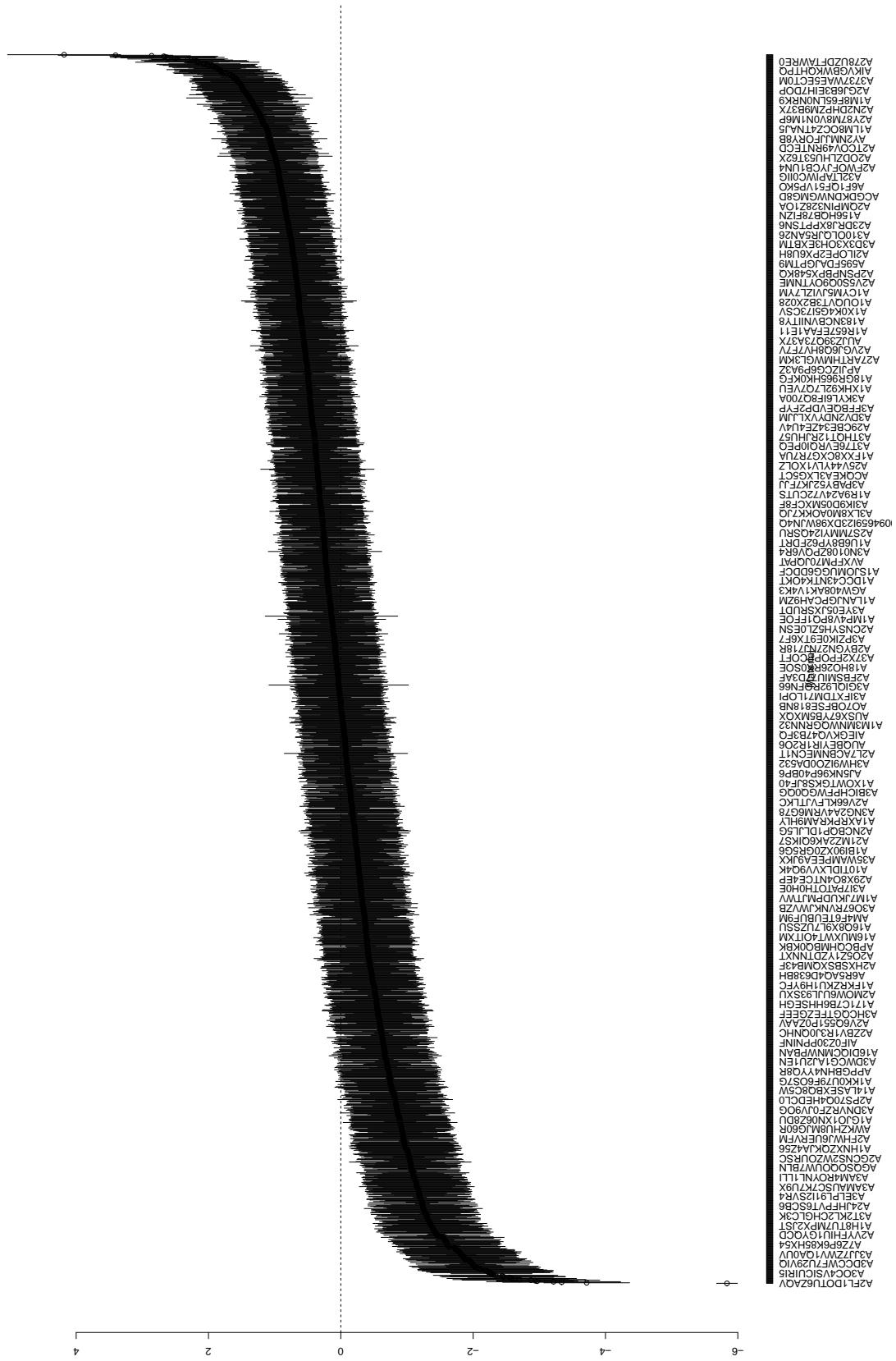


Fig. S57. Estimated mixed effects of user for the lightness attribute.

4.3.5 Metallicness

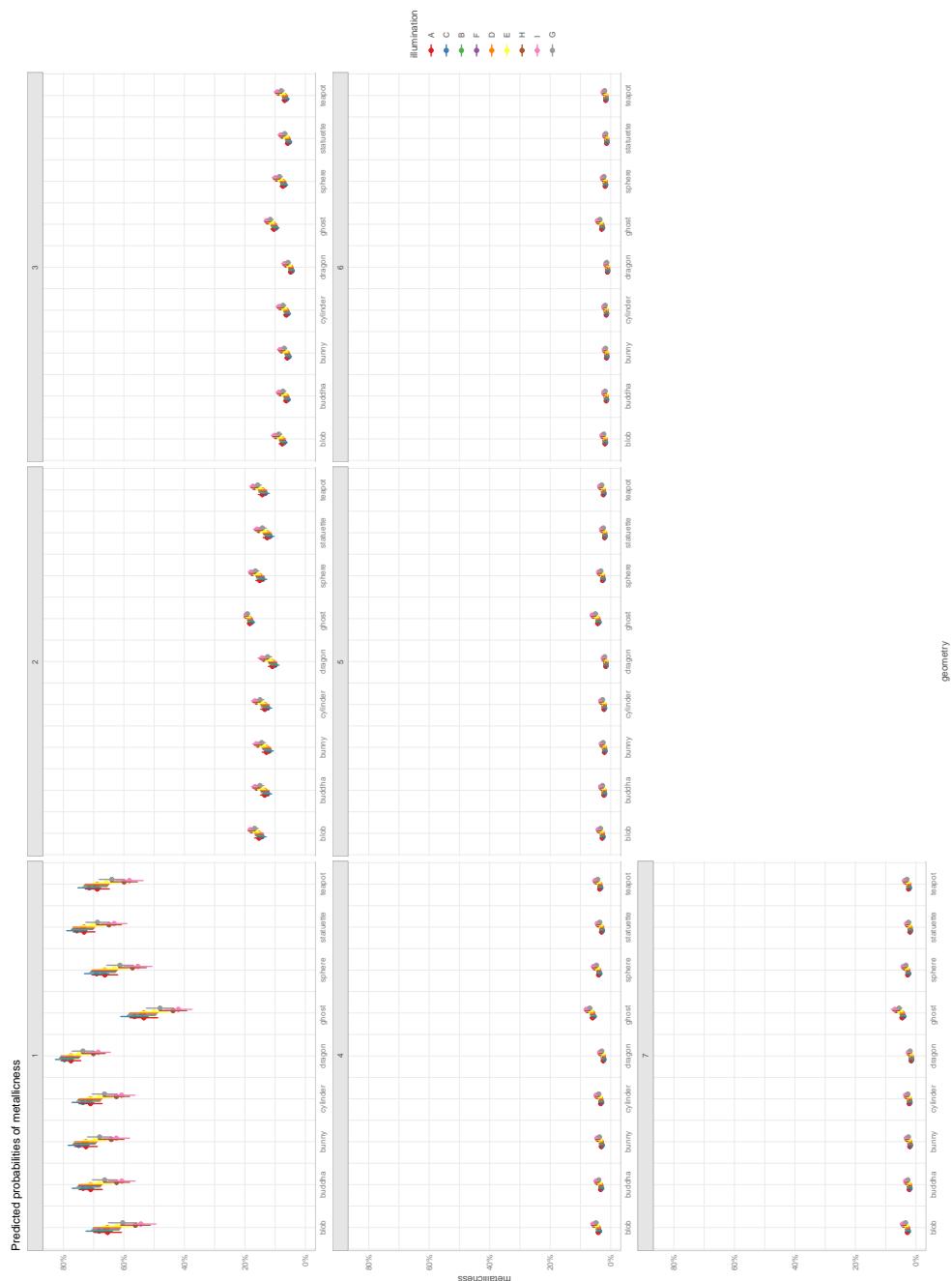


Fig. S53. Predicted probabilities for metallicness as a function of geometry and illumination.

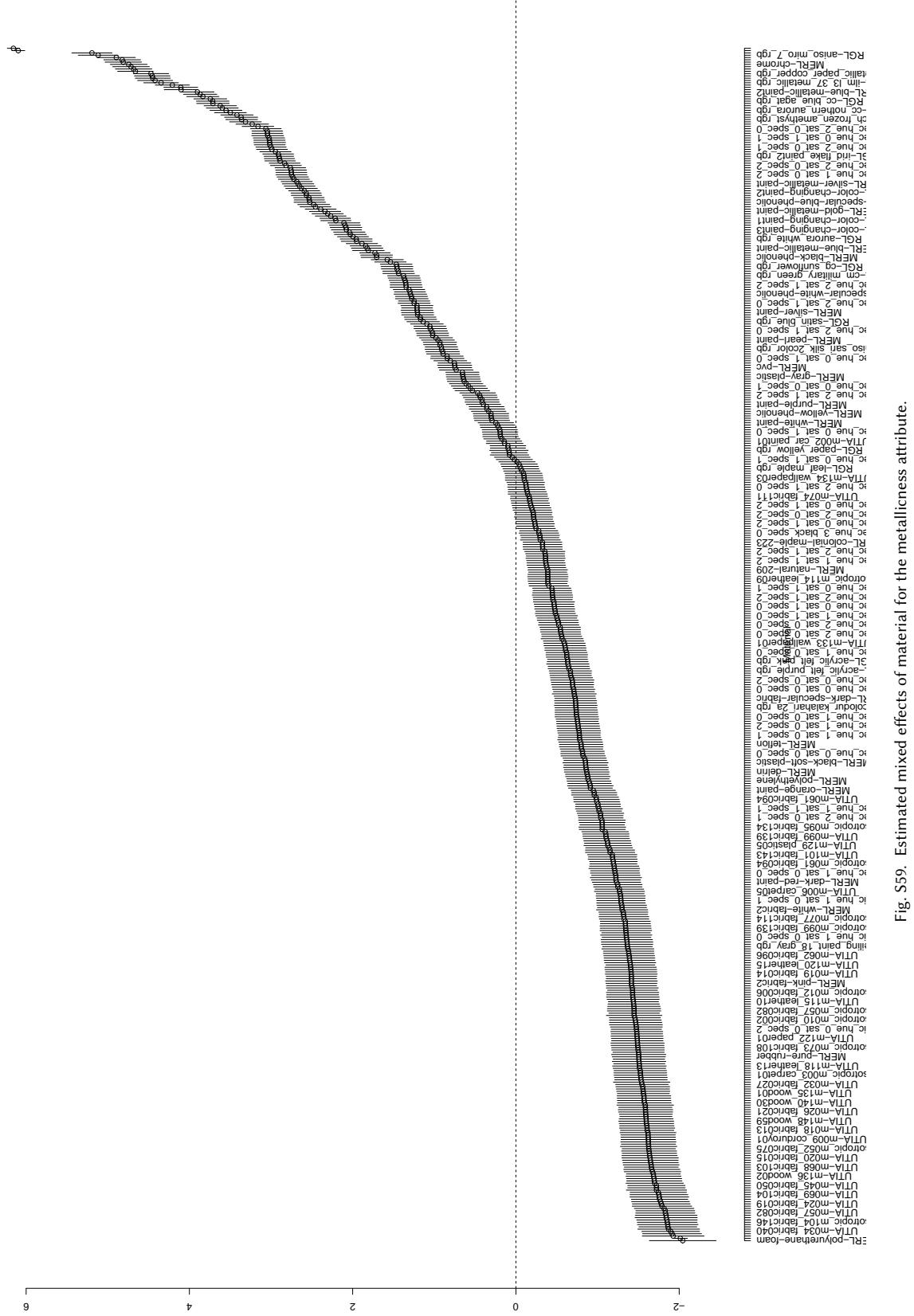


Fig. S59. Estimated mixed effects of material for the metallicness attribute.

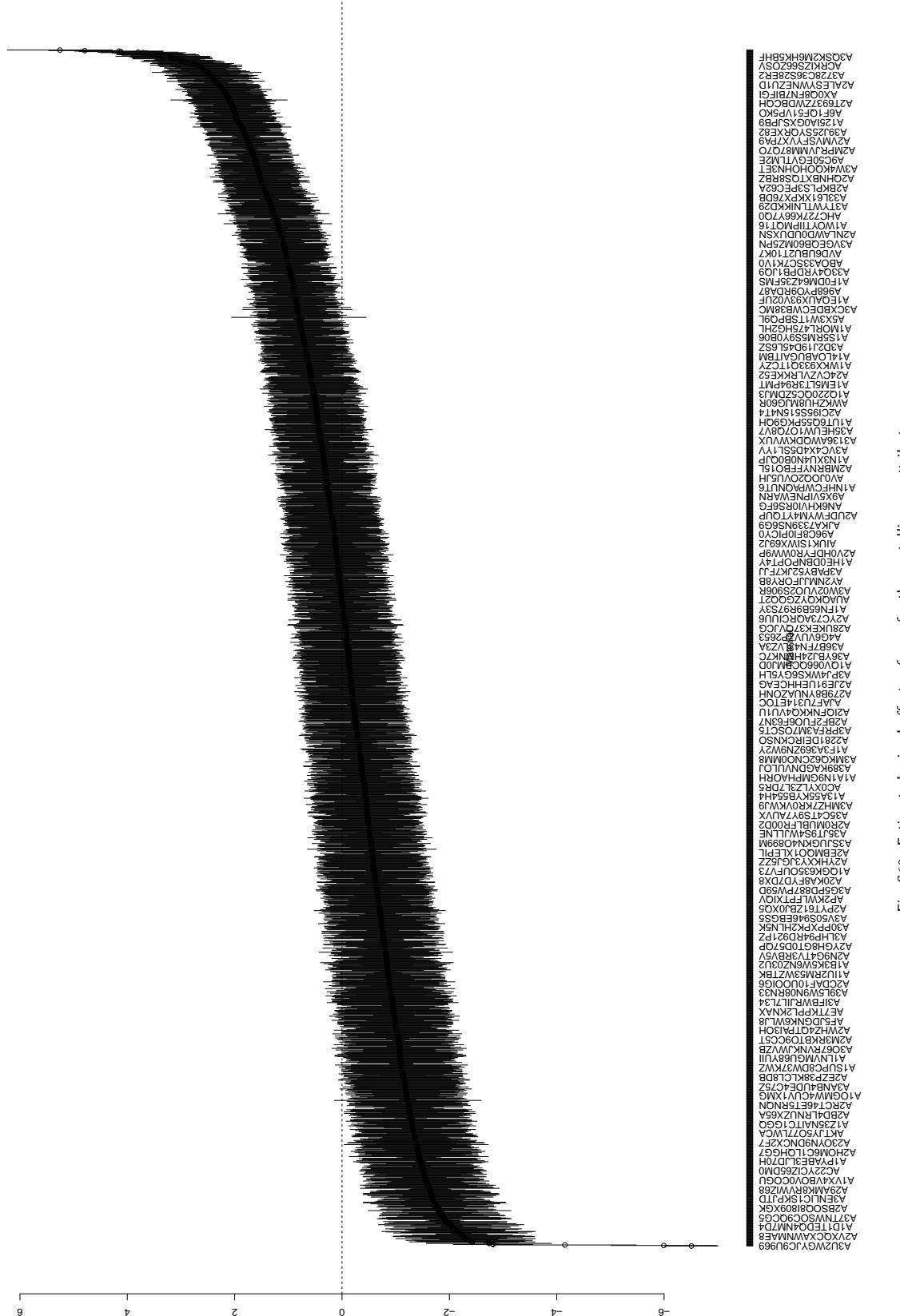


Fig. S60. Estimated mixed effects of user for the metallicness attribute.

S 5. BRDF EDITING

In this section we include additional data on the BRDF editing application described in Section 6.

Fig. S61 shows material appearance editing results given target glossiness 0.1, 0.3, 0.5 and 0.9 from left to right for the gold-metallic-paint in MERL on blob geometry under *St. Peter's Basilica* environment map.

Fig. S62 shows the material appearance editing results given target glossiness 0.1, 0.3, 0.5 and 0.9 from left to right for the gold-metallic-paint in MERL on blob geometry under *St. Peter's Basilica* environment map.

Table S2. BRDF editing results

BRDF environment map geometry	target glossiness	distribution	alpha	eta	k	specular_reflectance
blue-metallic-paint St. Peter's Basilica blob	0.7195 (fitted)	ggx	0.2569	(0.1889, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
	0.1	ggx	0.6248	(0.1889, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
	0.3	ggx	0.4866	(0.1889, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
	0.5	ggx	0.3902	(0.1890, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
	0.9	ggx	0.0752	(0.1890, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
gold-metallic-paint St. Peter's Basilica blob	0.7533 (fitted)	ggx	0.2651	(0.0000, 0.2738, 0.0543)	(0.0118, 0.1124, 0.0000)	(0.2668, 0.2054, 0.0626)
	0.1	ggx	0.6983	(0.0000, 0.2738, 0.0543)	(0.0118, 0.1124, 0.0000)	(0.2668, 0.2054, 0.0626)
	0.3	ggx	0.5258	(0.0000, 0.2738, 0.0543)	(0.0118, 0.1124, 0.0000)	(0.2668, 0.2054, 0.0626)
	0.5	ggx	0.4206	(0.0000, 0.2738, 0.0543)	(0.0118, 0.1124, 0.0000)	(0.2668, 0.2054, 0.0626)
	0.9	ggx	0.0875	(0.0000, 0.2738, 0.0543)	(0.0118, 0.1124, 0.0000)	(0.2668, 0.2054, 0.0626)
green-metallic-paint St. Peter's Basilica blob	0.6549 (fitted)	ggx	0.2817	(0.3277, 0.3072, 0.3423)	(0.1755, 0.3494, 0.0942)	(0.0647, 0.2180, 0.1486)
	0.1	ggx	0.6410	(0.3277, 0.3072, 0.3423)	(0.1755, 0.3494, 0.0942)	(0.0647, 0.2180, 0.1486)
	0.3	ggx	0.4871	(0.3277, 0.3072, 0.3423)	(0.1755, 0.3494, 0.0942)	(0.0647, 0.2180, 0.1486)
	0.5	ggx	0.3750	(0.3277, 0.3072, 0.3423)	(0.1755, 0.3494, 0.0942)	(0.0647, 0.2180, 0.1486)
	0.9	ggx	0.0707	(0.3277, 0.3072, 0.3423)	(0.1755, 0.3494, 0.0942)	(0.0647, 0.2180, 0.1486)
blue-metallic-paint Grace Cathedral blob	0.9	ggx	0.0542	(0.1889, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)
blue-metallic-paint St. Peter's Basilica sphere	0.9	ggx	0.0080	(0.1889, 0.2069, 0.1581)	(0.0959, 0.0831, 0.0632)	(0.0836, 0.0760, 0.1694)

Table S3. Validation on images rendered by analytical models

Surface scattering models	distribution	reflectance	material	diffuseReflectance
diffuse	N/A	gray (102, 102, 102)	N/A	N/A
	N/A	red (254, 39, 18)	N/A	N/A
	N/A	yellow (255, 239, 0)	N/A	N/A
	N/A	green (0, 128, 0)	N/A	N/A
	N/A	blue (102, 153, 204)	N/A	N/A
	N/A	black (24, 24, 24)	N/A	N/A
	N/A	white (245, 245, 242)	N/A	N/A
roughdiffuse	ggx	gray (102, 102, 102)	N/A	N/A
	ggx	red (254, 39, 18)	N/A	N/A
	ggx	yellow (255, 239, 0)	N/A	N/A
	ggx	green (0, 128, 0)	N/A	N/A
	ggx	blue (102, 153, 204)	N/A	N/A
	ggx	black (24, 24, 24)	N/A	N/A
	ggx	white (245, 245, 242)	N/A	N/A
plastic	N/A	N/A	N/A	gray (102, 102, 102)
	N/A	N/A	N/A	red (254, 39, 18)
	N/A	N/A	N/A	yellow (255, 239, 0)
	N/A	N/A	N/A	green (0, 128, 0)
	N/A	N/A	N/A	blue (102, 153, 204)
	N/A	N/A	N/A	black (24, 24, 24)
	N/A	N/A	N/A	white (245, 245, 242)
roughplastic	ggx	N/A	N/A	gray (102, 102, 102)
	ggx	N/A	N/A	red (254, 39, 18)
	ggx	N/A	N/A	yellow (255, 239, 0)
	ggx	N/A	N/A	green (0, 128, 0)
	ggx	N/A	N/A	blue (102, 153, 204)
	ggx	N/A	N/A	black (24, 24, 24)
	ggx	N/A	N/A	white (245, 245, 242)
ward	N/A	N/A	N/A	gray (102, 102, 102)
	N/A	N/A	N/A	red (254, 39, 18)
	N/A	N/A	N/A	yellow (255, 239, 0)
	N/A	N/A	N/A	green (0, 128, 0)
	N/A	N/A	N/A	blue (102, 153, 204)
	N/A	N/A	N/A	black (24, 24, 24)
	N/A	N/A	N/A	white (245, 245, 242)
conductor	N/A	N/A	Gold (Au)	N/A
	N/A	N/A	Silver (Ag)	N/A
	N/A	N/A	Aluminium (Al)	N/A
	N/A	N/A	Copper (Cu)	N/A
roughconductor	ggx	N/A	Gold (Au)	N/A
	ggx	N/A	Silver (Ag)	N/A
	ggx	N/A	Aluminium (Al)	N/A
	ggx	N/A	Copper (Cu)	N/A

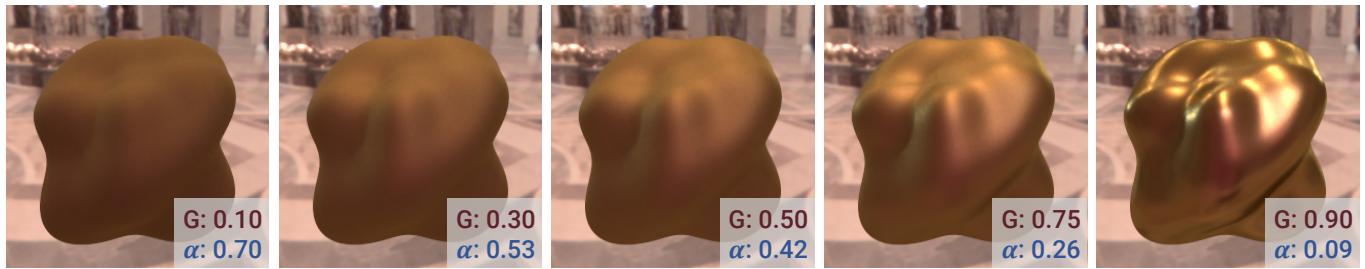


Fig. S61. Material appearance editing results given target glossiness 0.1, 0.3, 0.5, 0.75, and 0.9 from left to right.

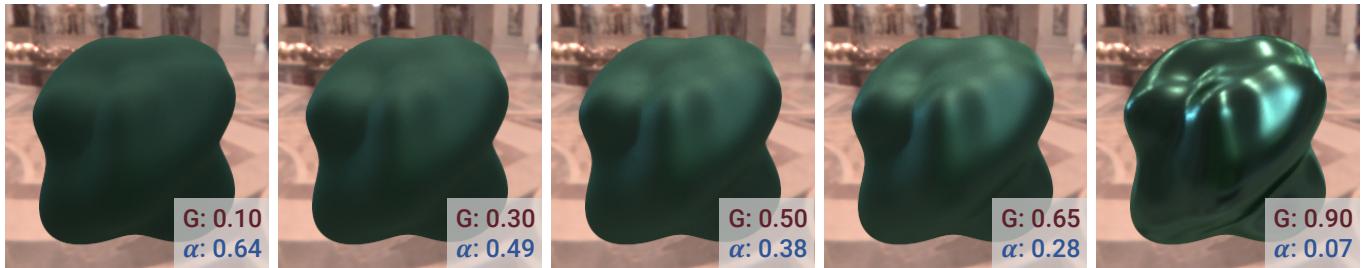


Fig. S62. Material appearance editing results given target glossiness 0.1, 0.3, 0.5, 0.65, and 0.9 from left to right.

REFERENCES

- Barton L Anderson and Juno Kim. 2009. Image statistics do not explain the perception of gloss and lightness. *Journal of Vision* 9, 11 (2009), 10–10.
- Alice C Chadwick and RW Kentridge. 2015. The perception of gloss: A review. *Vision Research* 109 (2015), 221–235.
- Johanna Delanoy, Manuel Lagunas, Ignacio Galve, Diego Gutierrez, Ana Serrano, Roland Fleming, and Belen Masia. 2020. The Role of Objective and Subjective Measures in Material Similarity Learning. In *ACM SIGGRAPH 2020 Posters (SIGGRAPH '20)*. Article 51, 2 pages.
- Jonathan Dupuy and Wenzel Jakob. 2018. An adaptive parameterization for efficient material acquisition and rendering. *ACM Transactions on graphics (TOG)* 37, 6 (2018), 1–14.
- Jiří Filip. 2015. Analyzing and predicting anisotropic effects of BRDFs. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*. 25–32.
- J. Filip and R. Vávra. 2014. Template-Based Sampling of Anisotropic BRDFs. *Computer Graphics Forum* 33, 7 (October 2014), 91–99. <https://doi.org/10.1111/cgf.12477>
- Roland W Fleming and Katherine R Storrs. 2019. Learning to see stuff. *Current Opinion in Behavioral Sciences* 30 (2019), 100–108.
- Juno Kim, Phillip J. Marlow, and Barton L. Anderson. 2012. The dark side of gloss. *Nature Neuroscience* 15, 11 (2012), 1590–1595.
- Max Kuhn et al. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software* 28, 5 (2008), 1–26.
- Manuel Lagunas, Sandra Malpica, Ana Serrano, Elena Garces, Diego Gutierrez, and Belen Masia. 2019. A Similarity Measure for Material Appearance. *ACM Transactions on Graphics (SIGGRAPH 2019)* 38, 4 (2019).
- Manuel Lagunas, Ana Serrano, Diego Gutierrez, and Belen Masia. 2021. The joint role of geometry and illumination on material recognition. *Journal of Vision* 21, 2 (2021). <https://doi.org/10.1167/jov.21.2.2>
- Phillip J Marlow and Barton L Anderson. 2013. Generative constraints on image cues for perceived gloss. *Journal of Vision* 13, 14 (2013), 2–2.
- Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. 2003. A Data-Driven Reflectance Model. *ACM Trans. Graph.* 22, 3 (2003), 759–769.
- Isamu Motoyoshi, Shin'ya Nishida, Lavanya Sharan, and Edward H Adelson. 2007. Image statistics and the perception of surface qualities. *Nature* 447, 7141 (2007), 206–209.
- Shin'ya Nishida. 2019. Image statistics for material perception. *Current Opinion in Behavioral Sciences* 30 (2019), 94–99.
- Jum C Nunnally. 1994. *Psychometric theory 3E*. Tata McGraw-hill education.
- Maria Olkkonen and David H Brainard. 2010. Perceived glossiness and lightness under real-world illumination. *Journal of Vision* 10, 9 (2010), 5–5.
- Fabio Pellacini, James A Ferwerda, and Donald P Greenberg. 2000. Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 55–64.
- Ana Serrano, Diego Gutierrez, Karol Myszkowski, Hans-Peter Seidel, and Belen Masia. 2016. An intuitive control space for material appearance. *ACM Trans. Graph. (SIGGRAPH ASIA 2016)* 35, 6 (2016).
- Robert Sève. 1993. Problems connected with the concept of gloss. *Color Research & Application* 18, 4 (1993), 241–252.
- Lavanya Sharan, Yuanzhen Li, Isamu Motoyoshi, Shin'ya Nishida, and Edward H Adelson. 2008. Image statistics for surface reflectance perception. *JOSA A* 25, 4 (2008), 846–865.
- Tiancheng Sun, Henrik Wann Jensen, and Ravi Ramamoorthi. 2018. Connecting measured brdfs to analytic brdfs by data-driven diffuse-specular separation. *ACM Transactions on Graphics (TOG)* 37, 6 (2018), 1–15.
- Matteo Toscani, Matteo Valsecchi, and Karl R Gegenfurtner. 2017. Lightness perception for matte and glossy complex shapes. *Vision Research* 131 (2017), 82–95.
- Peter Vangorp, Jurgen Laurijssen, and Philip Dutré. 2007. The influence of shape on the perception of material reflectance. In *ACM SIGGRAPH 2007 papers*. 77–es.
- Peter Welinder, Steve Branson, Pietro Perona, and Serge Belongie. 2010. The multidimensional wisdom of crowds. *Advances in neural information processing systems* 23 (2010), 2424–2432.
- Christiane B Wiebel, Matteo Toscani, and Karl R Gegenfurtner. 2015. Statistical correlates of perceived gloss in natural images. *Vision Research* 115 (2015), 175–187.
- Eduard Zell, Carlos Aliaga, Adrian Jarabo, Katja Zibrek, Diego Gutierrez, Rachel McDonnell, and Mario Botsch. 2015. To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 1–12.
- Fan Zhang, Huib de Ridder, Pascal Barla, and Sylvia Pont. 2020. Effects of light map orientation and shape on the visual perception of canonical materials. *Journal of Vision* 20, 4, Article 13 (2020), 18 pages.