Influence of Fresnel effects on the glossiness and perceived depth of depth-scaled glossy objects

Franz Faul Institut für Psychologie, Universität Kiel, Kiel, Germany

Christian Robbes Zentrum für integrative Psychiatrie (ZIP) gGmbH, Klinik für Psychiatrie und Psychotherapie, Kiel, Germany

Fresnel effects, that is, shape-dependent changes in the strength of specular reflection from glossy objects, can lead to large changes in reflection strength when objects are scaled along the viewing axis. In an experiment, we scaled sphere-like bumpy objects with fixed material parameters in the depth direction and then measured with and without Fresnel effects how this influences the gloss impression, gloss constancy, and perceived depth. The results show that Fresnel effects in this case lead to a strong increase in gloss with depth, indicating lower gloss constancy than without them, but that they improve depth perception. In addition, we used inverse rendering to investigate the extent to which Fresnel effects in a rendered image limit the possible object shapes in the underlying scene. We found that, for a static monocular view of an unknown object, Fresnel effects by themselves provide only a weak constraint on the overall shape of the object.

Introduction

Gloss is a subjective quality ascribed to surface materials. It is elicited by characteristic patterns in the proximal stimulus that, under natural viewing conditions, are produced by specularly reflecting surfaces. For the sake of brevity, the term "glossy object" is used in this article for objects with a specularly reflecting surface. Glossy objects reflect a mirror image of their surrounding, which is distorted depending on the form of the surface, and more or less sharp depending on the roughness of the material. Furthermore, and this is the main focus of our study, the intensity of the mirror image varies depending on the angle of the incident light and the index of refraction (IOR) of the material.

We here consider an important subclass of glossy objects that can be described as a combination of a diffusely reflecting base material and a clear transparent overlay. Common examples are opaque plastic objects like billiard balls, wet surfaces or a wood surface that has been treated with clear varnish (see Figures 1A and 1B). Accordingly, the bidirectional reflectance distribution function (BRDF) of this type of material is a combination of the BRDFs of diffuse and diffuse-specular reflection. An important material parameter of the transparent overlay is the IOR, which can take on values between 1.0 (air) and approximately 2.4 (diamond). Values for common materials are 1.33 for water, 1.49 for Plexiglass, and 1.45 to 1.70 for various types of normal glass.

For this material class, spatial variations in the strength of the diffuse–specular reflection that lead to changes in the intensity of the mirror image are described by the Fresnel equations for dielectric surfaces. We therefore refer to them as "Fresnel effects." Figure 1C shows (for unpolarized light) how the strength of the specular reflection depends on the incidence angle ω_i of the light (measured relative to the surface normal) and the IOR of the material: For perpendicular incidence ($\omega_i = 0$), the reflection strength is minimal, and this minimum grows with the value of the IOR. If ω_i increases, the reflection strength rises monotonically up to the maximum value of 1 at 90°. The incoming light that is not reflected specularly from the transparent overlay is transmitted and hits the base material, from which it is partly reflected diffusely.

These regularities imply that, all else being equal, the intensity of the mirror image is low at locations where the surface is viewed frontally and high at locations where the viewing angle is large. However, because the mirrored environment is usually a complex scene and the specular reflection is greater than zero at each surface position, the actual intensity distribution within the mirror image can strongly deviate from this pattern. It is, for example, well possible that the intensity is maximal at a position with perpendicular incidence, if a bright light shines on this low reflecting location. For this reason, Fresnel effects can at best be used as an heuristic to infer a specularly reflecting surface. They are presumably especially useful as a gloss cue under a

Citation: Faul, F., & Robbes, C. (2024). Influence of Fresnel effects on the glossiness and perceived depth of depth-scaled glossy objects. *Journal of Vision*, *24*(13):1, 1–24, https://doi.org/10.1167/jov.24.13.1.

https://doi.org/10.1167/jov.24.13.1 Received June 17, 2024; published December 2, 2024 ISSN 1534-7362

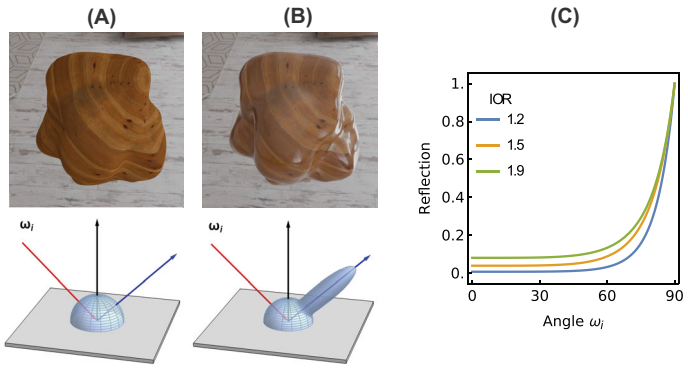

Figure 1. Reflection from a glossy object with a BRDF that combines diffuse and diffuse-specular reflection. (A) Diffusely reflecting wood surface and corresponding BRDF. (B) The same surface with a transparent overlay and corresponding BRDF. The lobe in the BRDF is related to the transparent overlay. Its width corresponds with the roughness of the transparent layer and its length to the strength of specular reflection. For illustrative purposes, the depicted BRDFs are simple approximations using the Phong reflection model (Phong, 1975). This model is *not* used in the investigation. (C) The strength of specular reflection predicted by Fresnel's equations, assuming unpolarized light.

global illumination that is relatively homogeneous and lacks dominant local light sources.

Previous findings

The results reported by Faul (2019) and Faul (2021) suggest that Fresnel effects are indeed used to detect specularly reflecting surfaces and their properties. It was found that ignoring Fresnel effects in the simulation of such surfaces had a detrimental effect on perceived gloss and gloss constancy under changes of illumination and/or object shape. Under a variety of conditions, the gloss impression was stronger and more realistic when the objects were rendered with a "Fresnel-BRDF" (Walter, Marschner, Li, & Torrance, 2007), which correctly models Fresnel effects, than when they were rendered with the Ward-BRDF (Ward, 1992), which ignores them. Another limitation of the original Ward model, compared to the Fresnel-BRDF, is that it is not energy-conserving. However, Faul (2019) obtained essentially the same results using the energy-conserving variant of the Ward-BRDF proposed by Geisler-Moroder and Dür (2010), confirming that the key factor was the omission of Fresnel effects. The perceptual differences in the gloss impressions obtained with these BRDFs were generally of a qualitative nature, that is, it was impossible to achieve a perfect match by varying the reflection strength parameter of the Ward-BRDF. In addition, informal observations suggested that the depth impression of objects is often weaker without Fresnel effects. For example, a sphere seems to be flatter when rendered with the Ward-BRDF instead of the Fresnel-BRDF.

Because Fresnel effects are intimately related to object shape, the aim of the present paper is to more closely investigate a) their role in estimating object depth, and b) their influence on the strength and constancy of gloss impressions. On the one hand, this investigation extends previous research on the influence of Fresnel effects on gloss perception and gloss constancy (Faul, 2019; Faul, 2021). On the other hand, it also complements investigations on the role of gloss for shape perception and vice versa (for recent reviews on this topic see Anderson & Marlow, 2023; Marlow & Anderson, 2024).

There are a number of studies that have explicitly investigated the extent to which gloss information influences depth perception. Fleming, Torralba, and Adelson (2004) considered shape perception in smooth metallic objects that are presented under global illumination. They argue that distortions in the mirror image can potentially be used as a shape cue, and compare it with an alternative shape cue that refers to the distortion of textures on a diffusely reflecting object. It is well-known that motion can improve both gloss perception and gloss constancy (Doerschner et al., 2011; Wendt, Faul, Ekroll, & Mausfeld, 2010). In addition, Dövencioğlu, Ben-Shahar, Barla, and Doerschner (2017) show that the specular flow in moving glossy objects can also lead to a more stable shape impression compared to matte objects. Todd, Norman, Koenderink, and Kappers (1997) combined stereo information with texture and gloss cues and found a small positive effect of gloss on the accuracy of shape recognition. Nefs, Koenderink, and Kappers (2006) considered a single monocular view of an object and tested whether shape perception is different in glossy and matte objects. Somewhat surprisingly, they found that gloss did not influence perceived shape. Adams and Elder (2014) added specular highlights to matte objects with an ambiguous shape that could be interpreted as either convex or concave. The addition of specular highlights led to a bias toward judging the shape as convex, suggesting that gloss has an influence on shape perception. Schlüter and Faul (2019) mainly investigated the case of transparent objects, but as a reference case also opaque glossy object. For this material class, no effect of the mirror image on shape perception was found, which is in line with the results of Nefs et al. (2006). There also exist several computational studies that explore how additional information, for example, motion (Adato, Vasilyev, Ben-Shahar, & Zickler, 2007), stereo information (Muryy, Welchman, Blake, & Fleming, 2013), or the mirror image of a known reference scene (Savarese, Chen, & Perona, 2005) can facilitate the shape reconstruction of specularly reflecting surfaces.

The converse problem, how shape influences gloss perception, was also investigated in several studies. An early study by Nishida and Shinya (1998) used local illumination and the simple Phong reflection model (Phong, 1975). They show that changes in the amplitude of a sinus pattern with fixed glossy material can influence perceived gloss. Later studies that used more realistic objects and illuminations essentially corroborated this effect of shape on perceived gloss (Faul, 2019; Olkkonen & Brainard, 2011; Serrano et al., 2021; Vangorp, Laurijssen, & Dutré, 2007). Another line of research considers the effect of incongruities between shape and highlight position on gloss perception (Beck & Prazdny, 1981; Kim, Marlow, & Anderson, 2011; Marlow, Kim, & Anderson, 2011; Todd, Norman, & Mingolla, 2004). The common finding of these studies is that shifting or rotating highlights relative to physically possible locations weakens or destroys the gloss impression. The results of Marlow, Todorović, and Anderson (2015) point to the necessity of a joint computation of gloss and shape. Wijntjes and Pont (2010) show that it is possible to create illusory gloss impressions with diffusely reflecting surfaces if the shape of the surface is chosen appropriately.

With regard to the influence of Fresnel effects on gloss and shape perception in the class of materials considered here, all these previous studies have some limitations. For example, they consider only localized illuminations, use simplifying BRDFs like the Phong or the Ward model, use another material class like metallic or transparent objects, or refer to additional information, like stereopsis or motion. Therefore, our main focus in the present paper is on how Fresnel effects influence perceived depth and gloss impressions in glossy objects presented under global illumination. More specifically, we investigate how gloss impressions and the perceived depth of smooth, convex "sphere-like" objects change when they are scaled by different factors along the depth direction. This is because the influence of Fresnel effects is expected to be especially strong in this situation. The variation of the presence of Fresnel effects is realized by using either the Fresnel-BRDF or the Ward-BRDF in the rendering.

Outline of the paper

To motivate our approach, we first show why the depth-scaling of glossy sphere-like objects can lead to strong changes in reflection strength that cannot be captured by BRDFs that ignore Fresnel effects.

We then present the results of an experiment, in which such depth-scaled glossy objects were used to investigate the effects of this manipulation on gloss and shape perception. The investigation was restricted to highly glossy objects of low and constant roughness, which reflect a relatively sharp mirror image of the surrounding. In a first subexperiment, the participants were asked to compare the gloss impression of two objects of identical shape but different BRDFs. In the second subexperiment the participants compared objects with identical BRDF (Fresnel or Ward) but different depth scaling.

Finally, we use inverse rendering to investigate the extent to which the presence of Fresnel effects in the rendered image constrains possible shapes in the underlying scene.

Depth scaling of objects and Fresnel effects

Fresnel effects depend on an object's shape, because the reflection angle, which determines reflection strength according to the Fresnel equations, is given by the angle between the viewing direction and the surface normal at the point of intersection with the object. In a frontal view of a sphere, for example, the reflection is minimal in the center (incidence angle $\omega = 0$), then increases monotonically with distance from the center, and reaches its maximum of 1 at the rim (incidence angle $\omega = 90°$).

We refer to this simple case of a sphere to illustrate why the strength of specular reflection generally depends on depth scaling. Scaling the sphere along the viewing axis by a certain factor x leads to a specific ellipsoid. Although the minimum in the center and the maximum at the rim are not affected by this operation, the distribution between these two points changes considerably. For $x > 1$ the sphere is elongated along the viewing axis, which leads to an increase of the area with strong reflection, whereas the compression observed for factors $x < 1$ decreases it.

Figure 2A shows a side view of the assumed situation. In this diagram, the depicted viewing direction corresponds with the grazing angle, that is, it is tangential to the surface and, from the viewer's perspective, the intersection with the surface lies at the rim of the object. Figure 2B shows the same situation for three different depth scaling factors and a viewing angle amounting to 90% of the grazing angle. It is easy to see in which way depth scaling influences the reflection angle and thus the reflectance strength in a given viewing direction.

The two left panels in Figure 3 show the reflection angle and reflection strength as functions of the viewing direction and illustrate how these functions depend on depth scaling. The viewing direction is given as a percentage of the grazing angle, where 0% corresponds with the center and 100% with the rim. The right diagram in Figure 3 shows that the grazing angle hardly

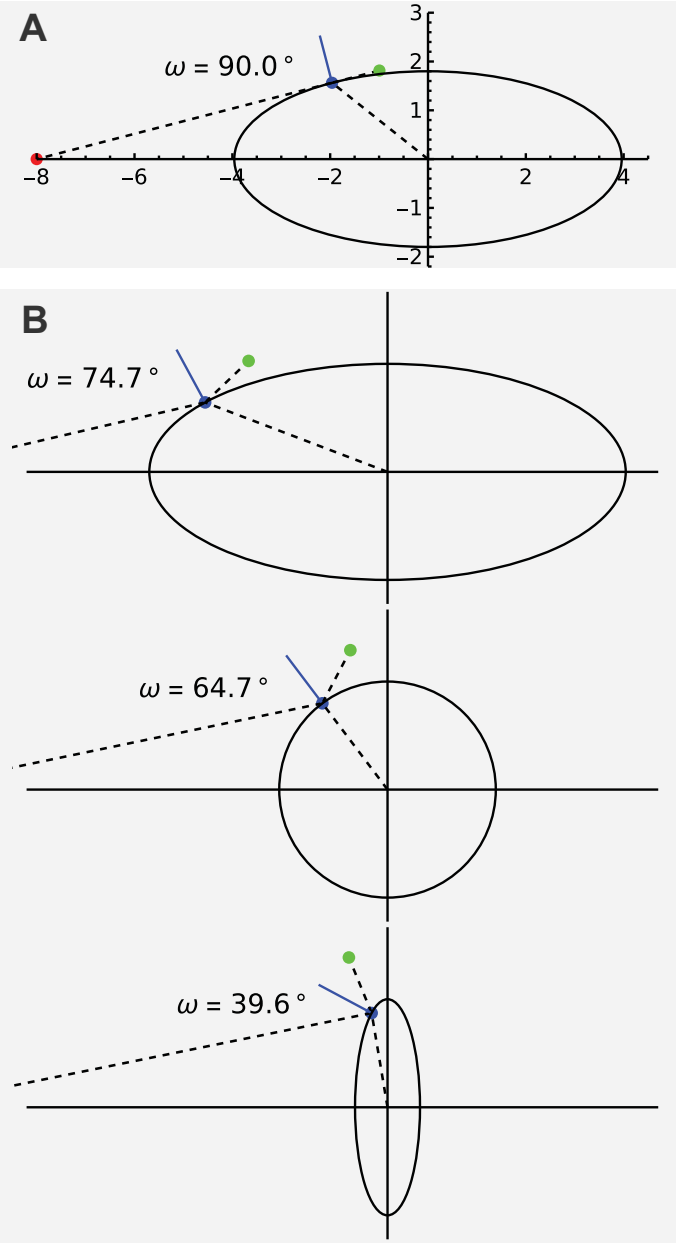

Figure 2. Reflection angle ω depending on the scaling factor of a sphere. (A) The side view of the sphere of radius 1.8 scaled by a factor $\alpha = 2.2$, centered at the origin and viewed from position $z = -8$ (red dot) along the depth axis. The dotted line starting from the viewing position (corresponding with the reflection direction starting from the blue point) is tangential to the object. Thus, here the (identical) angles of incidence and reflection are $\omega = 90°$ relative to the surface normal (blue line). (B) The same situation as in (A), but with scaling factors $\alpha =$ 2.2, 1.0, and 0.3 from top to bottom and a viewing direction corresponding with 90% of the grazing angle. It can be seen why, for a given viewing direction, the reflection angle ω decreases with the scaling factor.

depends on the scaling factor, so that across different scaling factors a given percentage of the grazing angle roughly correspond with the same object position in radial direction. This also means that, despite the perspective projection, the outline of the object remains essentially unchanged under depth scaling.

The left panel in Figure 3 shows how the reflection angle ω changes as a function of the depth scaling factor and the viewing angle. It gives a more complete picture of the relationship already illustrated in Figure 2B for a single viewing direction and three specific depth scaling factors. If these results for the reflectance angle are plugged into the Fresnel equations, assuming an IOR of 1.5 (see Figure 1C), the reflection strengths shown in the middle panel of Figure 3 are obtained. Note that the x-axis in this plot starts at 50% of the grazing angle, because the reflection strength is almost constant below this value.

Overall, these results show that, for a fixed material, the total reflection of approximately convex objects increases systematically as the object expands in depth. Figure 4 visualizes this regularity in renderings of a blob object that is scaled in depth. The upper row shows the physically plausible results obtained with the Fresnel-BRDF and the middle row illustrates the corresponding results obtained with the Ward-BRDF, which ignores Fresnel effects. This comparison confirms that the increase in overall reflection strength is due to Fresnel effects.

These observations raise two questions: 1) Can the relationship between reflection strength and depth scaling be used as an additional gloss-related depth cue? 2) Is this systematic increase in reflectance due to depth scaling correctly accounted for by the visual system or is it erroneously attributed to a change in material?

Experiment

To answer these questions, we conducted an experiment consisting of two subexperiments, in which the depth scaling was the central independent variable.

In the first subexperiment, we compared standard objects rendered with a Fresnel-BRDF with an object that had an identical shape, but was rendered with the Ward-BRDF (condition "Fresnel-Ward"). The reflection strength and the depth scaling factor of the standard object were varied in several steps. At the beginning of each trial, the two objects were presented side-by-side, and the participants had to perform four tasks in the given order: 1) rate the glossiness of both objects separately, 2) rate the relative depth impression of both objects, 3) match the gloss impression of both objects by adjusting the reflection strength parameter of the Ward-BRDF accordingly, and 4) rate the quality of the best achievable match.

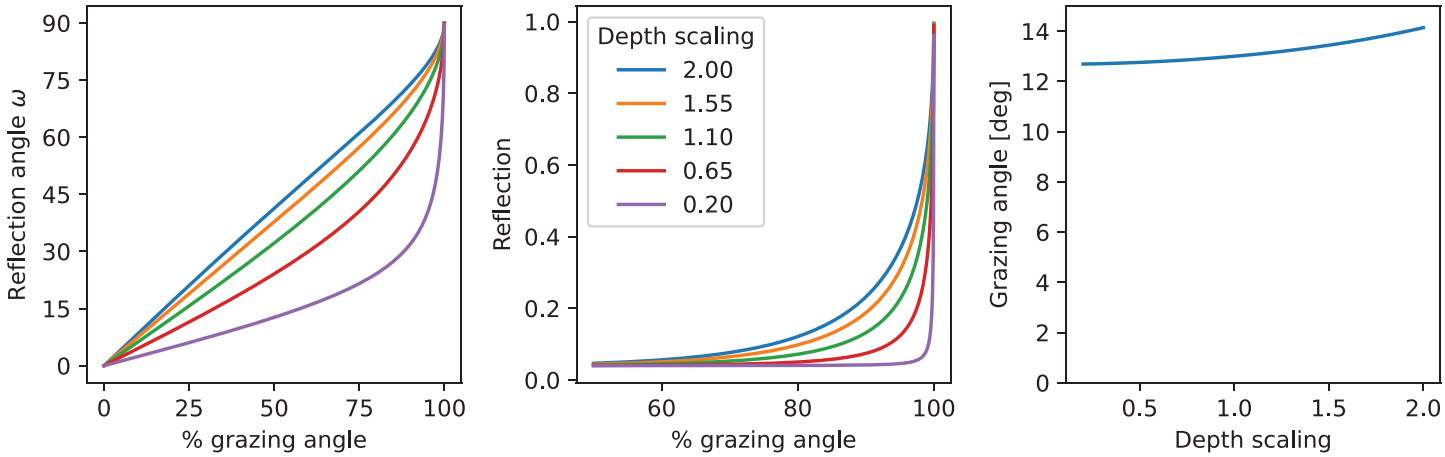

Figure 3. Reflection as a function of the viewing direction for different depth scaling factors. The situation is the same as in Figure 2. (Left) Reflection angle ω. (Center) Reflection strength according to the Fresnel equations for an IOR of 1.5. (Right) Grazing angle depending on the scaling factor.

The second subexperiment was very similar. The main difference was that in each trial the participants compared objects rendered with identical BRDF but different depth-scaling factors. The common BRDF was either the Fresnel-BRDF (condition "Fresnel-Fresnel") or the Ward-BRDF (condition "Ward-Ward"). In this case, the depth scaling factor of the standard object was always fixed to a value of 1.1. The tasks were identical to those in subexperiment 1.

Because the same tasks were used in both experiments, it was possible to conduct them in parallel in a single session. Conducting both subexperiments simultaneously had the advantage that the properties of the standard object varied more, which presumably made it more challenging to use cognitive strategies when completing the tasks. To reduce the likelihood that participants would simply reproduce a response from memory for a recently presented similar stimulus, all 180 conditions were presented in random order. In particular, participants were not aware of which type of subexperiment they were working on in a specific trial. The random order was identical for all participants.

Methods

The type of stimuli used in the experiment and the variation of the main independent variables BRDF and depth scaling are illustrated in Figure 4.

Design and stimuli

In each of the three BRDF conditions ("Fresnel-Ward", 'Fresnel-Fresnel", "Ward-Ward"), we used five depth scaling factors (0.2, 0.65, 1.1, 1.55, and 2.0). To enhance the generalizability of the results, we varied as further variables (see Figure 5) reflection strength (IOR = 1.26, 1.51, and 1.75), object shape ("blob," "sinus-blob"), and the illumination ("outdoor," "indoor"). Overall, this resulted in 3 (BRDF-cond) × 3 (IOR) × 2 (shape) × 2 (illumination) × 5 (depth scaling) = 180 conditions for each participant.

Because our comparisons include a physically impossible situation without Fresnel effects, it is unavoidable to use computer-generated stimuli. This is common praxis in research on material perception, but one must keep in mind that this usually limits the dynamic range of the stimuli and may also influence gloss impressions (Doerschner, Maloney, & Boyaci, 2010). All stimuli were rendered with Mitsuba 0.6 (Jakob, 2010), using the "roughplastic" (Fresnel-BRDF) and "ward" (Ward-BRDF) material with a resolution of 512 × 512 pixel (see section 'Details of the stimulus generation' in the Appendix for more details on the BRDFs and the render parameters). The Mitsuba utility program "mtsutil tonemap" was used to transform the resulting high dynamic range images to sRGB coordinates for display. In the scene, the objects were placed on a dark gray floor with a light gray grid texture, which was visible in the reflected mirror image and was meant to support the depth impression. To obtain gloss impressions of similar strength in both BRDFs, the specular reflection strength parameter ρ_s of the Ward-BRDF was set to $\rho_s = (\text{IOR} - 1) * 0.17$. In Faul (2019), it was found that this transformation provides a good approximation for objects that resemble the blob with depth scaling of 1.

The illumination maps for both the indoor ("brown_photostudio_06") and the outdoor ("konigsallee") illumination were obtained from Polyhaven (https://polyhaven.com). With the outdoor illumination, the light came predominantly from above, whereas with the indoor illumination, the dominant light source was located behind the observer and thus

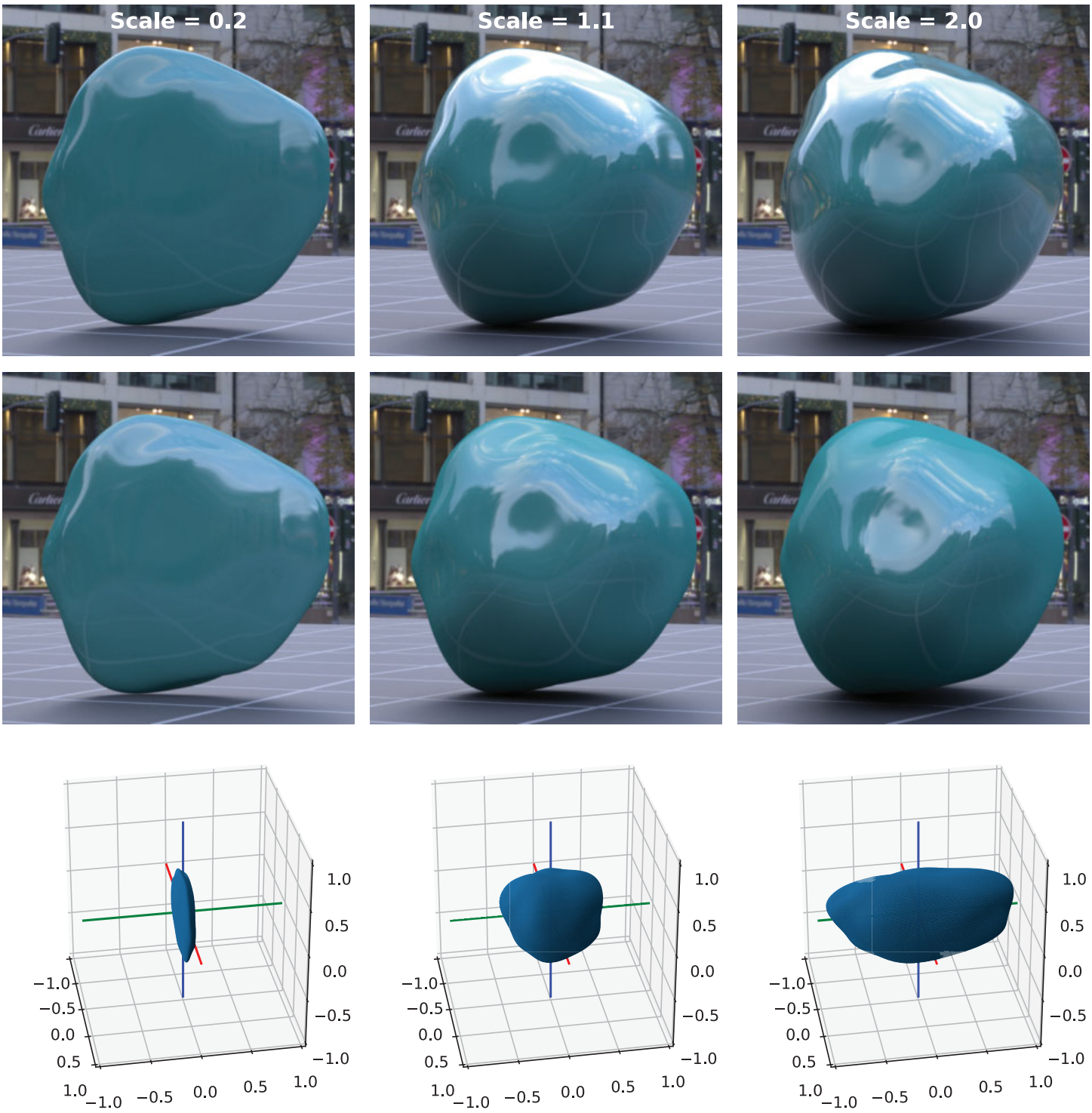

Figure 4. A blob object transformed with three different depth scaling factors (0.2, 1.1, and 2.0), rendered with unchanged material parameters using the Fresnel-BRDF (top), and the Ward-BRDF (middle). The lower row shows the corresponding side view of the object in three-dimensional space. It is located at the center and scaled along the depth axis (green).

illuminated the object frontally (see Figure 5). Because Fresnel effects are especially strong near the rim of the object, we expected them to be more pronounced (and the difference to the Ward-BRDF greater) with the outdoor illumination, which illuminates those areas more brightly than the indoor illumination.

The main reason for including the "sinus-blob" object was that it contains strong local perturbations

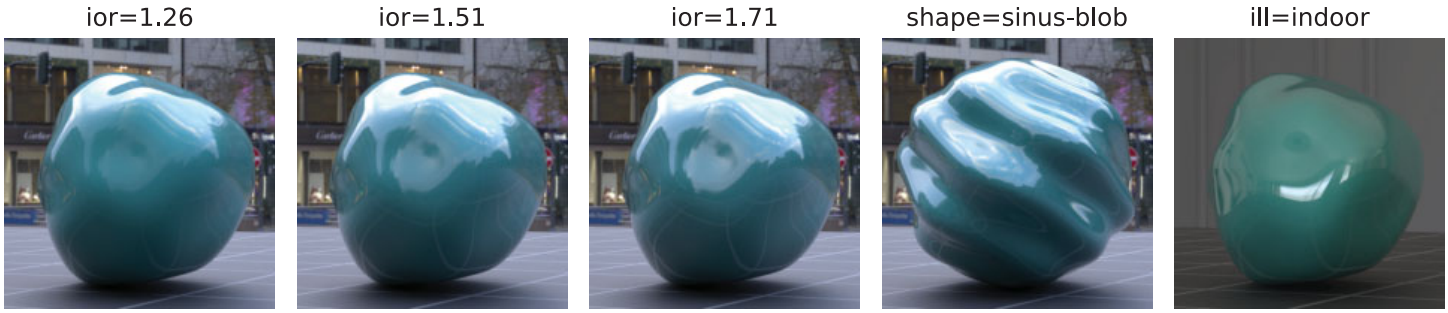

Figure 5. Variations of the stimuli (from left to right): Panels 1 through 3 illustrate the three steps of the IOR, panel 4 the alternative shape "sinus-blob," and panel 5 the alternative illumination "indoor." All stimuli are shown with the Fresnel-BRDF and scaling factor 2.

that lead to a more even distribution of the Fresnel effects across the surface than in the "blob" object, where they are most pronounced near the rim of the object.

Stimulus presentation and instruction

The stimuli were shown side-by-side against a dark gray background on an Eizo ColorEdge CG243W display (24", 1,920 × 1,200 pixel, 60 Hz) connected to an Nvidia Quadro 600 graphics card. The experiment was programmed in Processing 4.0 (https://processing.org).

The fixed standard stimulus was always presented on the left side of the screen, the adjustable test stimulus on the right side. In the first subexperiment, the standard object was always rendered with the Fresnel-BRDF and the test object with the Ward-BRDF, but both had the same depth scaling. In the second subexperiment, the standard object always had a depth scaling of 1.1, whereas the depth scaling of the test object was selected from the five different levels of the depth factor.

At the start of each trial, both objects had nominally the same reflection strength (for the Ward-BRDF, the transformation from IOR to ρ_s described earlier was used to approximate this). The participants first rated the gloss impression of the standard object on a scale from 0 (absent) to 9 (very high). They then did the same for the test object. Next, they had to rate the relative depth of the standard and the test object on a scale with seven steps, with the center position corresponding to "equal." For both directions ("left" vs. "right"), an unequal depth could be further specified with the labels "somewhat larger," "larger," and "much larger." The current settings in all three tasks were reflected in graphical scales shown below the stimuli. After the third rating was finished, the scales disappeared and the participants had to best match the gloss impression in both objects by adjusting the reflection strength of the test stimulus. Possible values ranged from IOR = 1 to IOR = 3 in 100 equal-distant steps (in cases where the Ward-BRDF was adjusted, the corresponding transformed values ρ_s for the Ward-BRDF were used). Because a perfect match is usually not achievable, the last task in each trial was to rate the quality of the best achievable match on a scale with the six levels "impossible," "bad," "somewhat bad," "good," "very good," and "perfect." The number of levels in each of the rating scales was selected based on the expected accuracy of the assessment.

Participants

The mean age of the ten participants was 30.4 (SD = 9.91) years. All were color normal according to the Ishihara test and reported normal or corrected-to-normal visual acuity. All were naive with respect to the purpose of the study.

Results

The results in all four shape × illumination conditions were qualitatively similar, that is, the same general pattern of results depending on depth scaling was observed. For this reason, and because our focus is on the effect of the depth factor, we only report the results averaged across these conditions in the main text. It should be noted, however, that the trends visible in these mean results were in general stronger with the outdoor illumination than with the indoor illumination and slightly more pronounced in the blob than in the sinus-blob condition (as an example, see Figure A2 in the Appendix). More detailed information on the interplay between the experimental factors is given in a statistical analysis using a mixed linear model (see section 'statistical analyses' in the Appendix).

Subexperiment 1

In the first subexperiment, the participants compared objects with different BRDFs but identical shape. The standard object was rendered with the Fresnel-BRDF

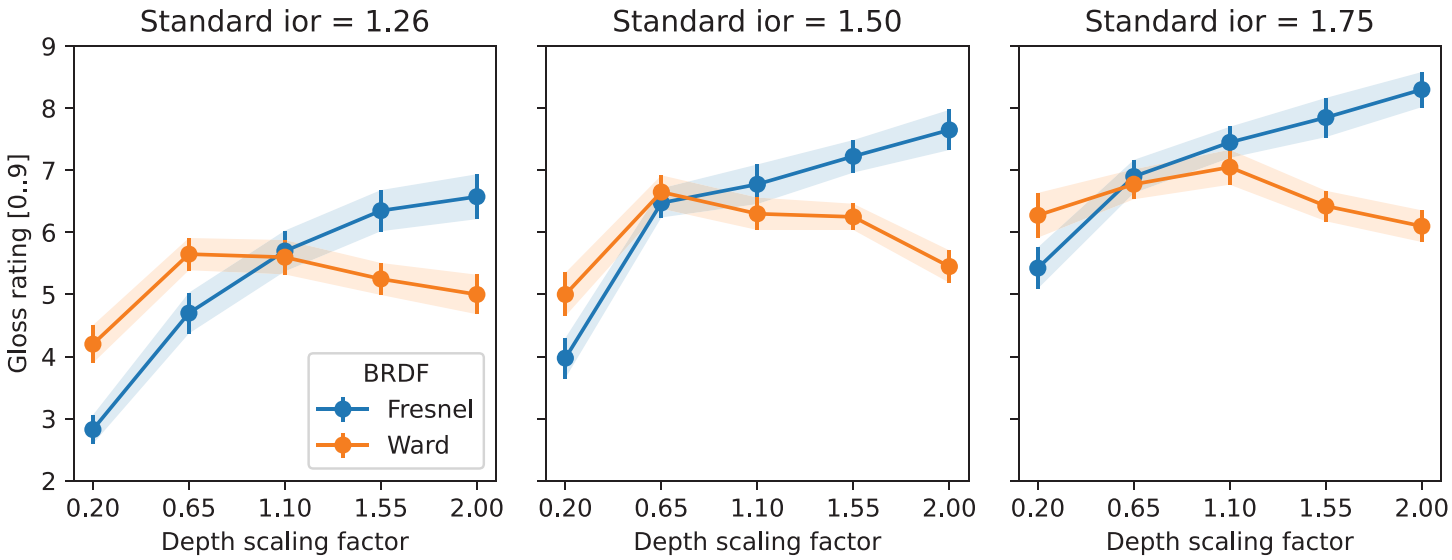

Figure 6. Gloss impression with Ward- and Fresnel-BRDF for fixed material parameters, depending on depth scaling. From left to right, the fixed IOR changes from 1.26 to 1.75. $N = 40$; error bars correspond to $\pm$ 1 standard error of the mean.

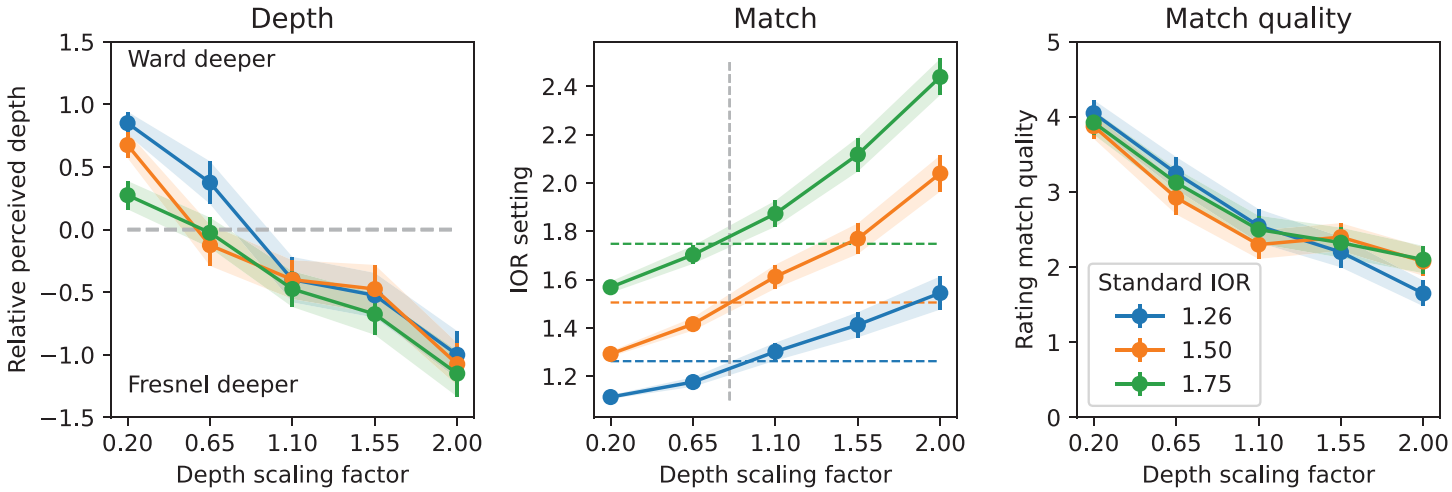

Figure 7. Relative depth rating and gloss matching of objects rendered with Fresnel-BRDF and Ward-BRDF, depending on depth scaling. (Left) Relative perceived depths on a scale from −3 (much larger depth with Fresnel) to 0 (equal) to 3 (much larger depth with Ward). (Center) Reflection strength settings in the Ward-BRDF to approximately match the gloss impressions in standard and test. The reflection strength parameter ρ_S is transformed to a nominal "IOR" value by IOR $= \rho_S/0.17 + 1$. The horizontal dashed lines show the IORs of the standard. (Right) Quality of the match. $N = 40$; error bars correspond to $\pm$ 1 standard error of the mean.

and had a fixed IOR, whereas the adjustable test object was rendered with the Ward-BRDF. At the start of each trial, the parameter ρ_s, which controls the strength of the specular reflection of the test object, was set to $\rho_s = (IOR - 1) * 0.17$. As already mentioned, this value was chosen because Faul (2019) found it produced similar gloss impressions with both BRDFs.

Figure 6 displays the mean results for the first task, which involved separately rating the gloss impression in standard and test, given these initial values. It can be seen that with Fresnel effects, the strength of the gloss impression increases monotonically with increasing depth scaling. When using the Ward-BRDF, the corresponding curve is approximately constant with a slight peak around medium depth scaling.

The second task was to assess the relative depth impression in standard and test object. The corresponding mean results, depicted in the left panel of Figure 7, show that at small depth scaling factors, the higher perceived depth was observed with the

Ward-BRDF, but that the relative depth observed with the Fresnel-BRDF systematically increases with increasing scaling factors. This finding indicates that the increase in perceived depth with increasing depth scaling is greater with Fresnel effects than without them. An impression of how the BRDF influences the changes in gloss and perceived depth with increasing depth scaling can be gained by comparing the top and middle row in Figure 4.

Finally, the participants had to match the gloss impression in test and standard by adjusting the strength of specular reflection in the test and also to judge the quality of the best achievable match. The corresponding results are shown in the middle and right panels of Figure 7. To achieve an approximate match of the gloss impression, it was necessary to considerably change the reflection strength of the Ward-BRDF from its initial value. For small depth scaling factors it had to be lowered, for large factors it had to be increased. The vertical dashed line in the middle panel of Figure 7 indicates the approximate value of the scaling factor at which the initial value for ρ_s could be left unchanged. The monotonic increase of these functions implies that the previously described conversion from IOR to ρ_s is only correct for a single depth scaling factor, or, more generally, that this conversion is shape dependent.

As can be seen in the right panel of Figure 7, the rated quality of the match decreases with increasing depth scaling. This finding indicates that a complete match was not possible with large Fresnel effects.

Subexperiment 2

In subexperiment 2, standard and test had the same BRDF (Ward or Fresnel) but generally differed in depth scaling. The standard always had a depth scaling factor of 1.1, whereas the depth of the test varied in the same five levels used in subexperiment 1.

The results of subexperiment 2 are summarized in Figure 8. At the beginning of each trial, the test was identical to either the standard (for the Fresnel-BRDF) or the test (for the Ward-BRDF) used in one of the trials in subexperiment 1. Thus, the mean gloss ratings for the tests shown in the top row of Figure 8 should ideally be identical to the mean ratings for the corresponding BRDF in subexperiment 1, shown in Figure 6. That this is approximately the case, confirms that these ratings are fairly reliable.

In the matching task, the participants adjusted the strength of the specular reflection in the test object to make the gloss impressions in standard and test as similar as possible. For equal depths of standard and test, choosing the same reflection strength as in the standard should result in a perfect match. The matching results depicted in the second row of Figure 8 show that in this case the subjects indeed reproduced the value of the standard. Would the same setting be chosen for the other depth scaling factors, then this would indicate complete gloss constancy under shape deformation. However, this is clearly not the case. With the Fresnel-BRDF, a much larger reflection strength must be chosen for objects of lower depth and slightly lower values for objects with larger depth. For the Ward-BRDF, slightly larger reflection strengths were chosen for objects with lower *and* higher depth than the standard. Overall, the dependence of reflection strength on depth scaling was lower for the Ward-BRDF, indicating a better approximation to constancy than with the Fresnel-BRDF. It is to be expected that the observed gloss ratings and the reflection strength settings are complementary to each other, because if the gloss rating for the test objects is lower in condition A than in condition B, then a larger reflection strength should be chosen in A than in B to achieve a match. This is indeed observed.

Interestingly, the quality of the match shown in the third row of Figure 8 degraded considerably with increasing difference in depth scaling. The degradation was almost identical for both BRDF and all standard IOR. A possible interpretation is that this indicates that, even when the same BRDF is used, there are qualitative differences in the gloss impression elicited by objects that differ only in depth scaling. An alternative interpretation could be that the participants did not perform the intended rating of the difference in the gloss impression, which requires an abstraction, but instead evaluated the overall similarity of the two stimuli.

Finally, the fourth row of Figure 8 shows the relative perceived depth in standard and test. As expected, the same depth (relative depth = 0) is observed, if standard and test have objectively the same depth scaling factor. Under ideal conditions, that is, perfect depth reconstruction and linear rating behavior, one would expect a linear relationship between the depth scaling factor and relative depth. Actually, we observed negatively accelerated functions with both BRDFs. But with the Fresnel-BRDF, the function at least increases monotonically, whereas the function for the Ward-BRDF is almost flat for depth factors larger than that of the standard. Because other potential depth cues, for example the size of the object's shadow, were identical in both BRDF, this difference can be attributed to Fresnel effects.

This finding is in line with the results obtained in subexperiment 1 with different BRDFs. It also corroborates the casual observation reported in Faul (2019) that the depth impression of an object rendered with Ward-BRDF is often improved when it is rendered with correct Fresnel effects (see also Figure 4).

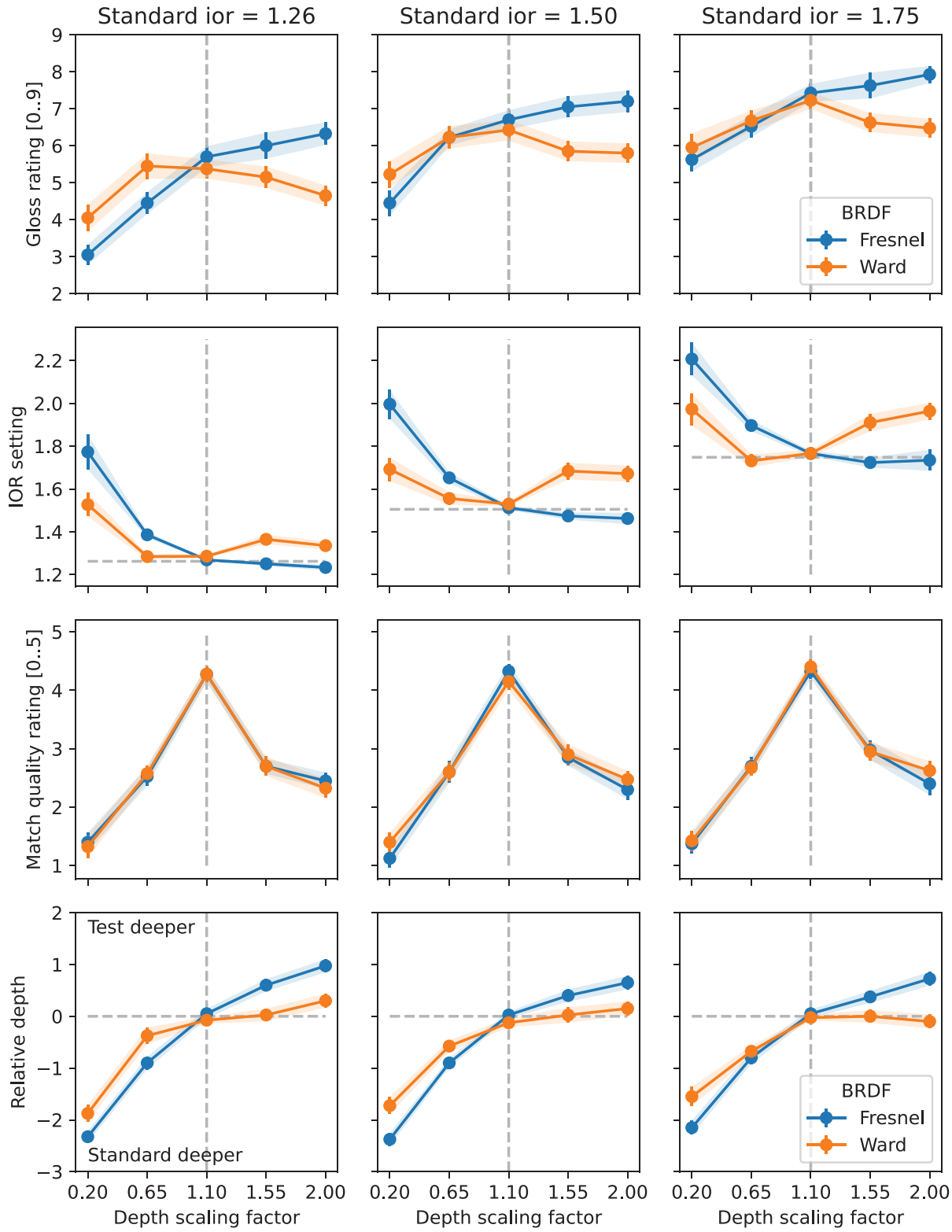

Figure 8. Gloss ratings, gloss matches, match quality, and relative depth as a function of the depth scaling of the test object. Data for Fresnel- and Ward-BRDF in subexperiment 2. From the left to the right panel, the fixed IOR changes from 1.26 to 1.75, whereas within each panel, the material parameters of the standard object are fixed for both the Ward-BRDF and the Fresnel-BRDF. (First row) Gloss ratings with identical material for standard and test. (Second row) IOR settings in the test object leading to an approximate gloss match between standard and test. The dashed horizontal lines indicate the IORs of the standard. (Third row) Quality of the gloss match. (Fourth row) Relative perceived depth of standard and test. The vertical dashed lines indicate the depth scaling of the standard. $N = 40$; error bars correspond with $\pm$ 1 standard error of the mean.

Probing depth ambiguity with inverse rendering

In the experiment, object shape was scaled along the viewing axis (i.e., in depth direction) by factors from 0.2 to 2.0. Although the depth impression clearly increases with real depth, the range of perceived depths is rather narrow, as can be appreciated in Figure 4. When using the Ward-BRDF, this "perceptual compression" is even more pronounced. These informal observations are corroborated by the results of the experiment (cf. lower row in Figure 8). But what are the reasons for this compression?

In a single frontal view of an unknown object, there is usually considerable ambiguity with regard to the depth dimension. For example, Belhumeur, Kriegman, and Yuille (1997) show that, when an unknown diffusely reflecting object is viewed orthographically, there is an implicit ambiguity regarding the object's depth ("bas-relief ambiguity"). This means that, for each image of the object under a given illumination, one can produce an identical image after applying a specific "generalized bas-relief" transformation to the object that alters its shape and albedo.

The perceptual compression of the actual depth of our glossy objects may be a consequence of such an ambiguity. On the other hand, the stimuli used in the experiment contained several monocular depth cues. For example, the object was presented on a floor with a grid texture so that the reflections of the grid lines are visible on the object's surface, and the objects also cast a shadow on the floor that changes with shape. Furthermore, the results summarized in Figure 3 show that the reflection strength of glossy sphere-like objects stands in a monotonic relationship with depth. One may therefore be inclined to assume that this gloss-related information can also be used to partly disambiguate the situation. The difference in perceived depth observed with Fresnel-BRDF and Ward-BRDF can be taken as confirmation of this assumption.

To investigate to what extent gloss information (including Fresnel effects) is useful to disambiguate depth, we used inverse rendering. Inverse rendering is a fitting procedure within photorealistic rendering that aims to adjust parameters of a simulated 3D scene to minimize the difference between a target image and the image rendered based on that scene.

We used the freely available Mitsuba 3 renderer (Jakob et al., 2022), which can be considered state of the art in inverse rendering. Mitsuba 3 no longer supports the Ward-BRDF and the "gridtexture" plugin of Mitsuba 0.6, which were used when rendering the stimuli for the experiment. Because we wanted to use the experimental stimuli also in the simulation, we ported these two plugins from Mitsuba 0.6 to Mitsuba 3. Two recent additions to Mitsuba 3, namely "large steps optimization" (Nicolet, Jacobson, & Jakob, 2021) and "projective sampling" (Zhang, Roussel, & Jakob, 2023) were instrumental for achieving a good geometry fit.

Fitting the Ward-BRDF

A first aim was to explore the extent to which it is possible to reproduce the rendering obtained with the Ward-BRDF (see top left image in Figure 9) when using the Fresnel-BRDF and allowing (almost) arbitrary changes to the geometry of the underlying mesh. In a sense, this is an attempt to deduce what shape a visual system, which uses information from Fresnel effects in scene recognition, would see when confronted with an image rendered with the Ward BRDF, which ignores Fresnel effects.

The fit procedure used was very similar to the one described in the "shape optimization" tutorial in the Mitsuba 3 documentation (https://mitsuba.readthedocs.io/en/stable/src/inverse_rendering/shape_optimization.html). The fit was performed using linear coordinates, that is, without tone mapping. We initialized the mesh with the standard object in the proper scaling and used two views, the frontal view (as in the experiment) and the back view, to constrain the mesh on both sides. Images generated within the fit procedure and the target image both had a resolution of 256 × 256 pixels and were rendered with 512 samples per pixel. We used exemplarily the condition with the largest influence of Fresnel effects on the render result, namely the blob shape, the outdoor illumination, and an IOR of 1.75. Similar results were obtained in other conditions (see also Figure A3 in the Appendix).

The result is shown in the bottom left panel of Figure 9. As expected, a complete reproduction is not possible, because the almost complete absence of specular reflections near the rim that is observed when using the Ward-BRDF cannot be reproduced with the Fresnel-BRDF. Nevertheless, the similarity is astonishingly high, given that only geometric adjustments were allowed. At first glance, the two renderings seem to be almost identical. In any case, the result is much more similar to the rendering with the Ward-BRDF than to the one with the same Fresnel-BRDF (top right panel in Figure 9).

Figure 10A compares the shape resulting from this fit with the correct shape that was used when rendering with the Ward-BRDF. It is evident that the shape alterations are small and would be hard to discern when the two shapes are not positioned in direct juxtaposition. Figure A3 in the Appendix illustrates that this fit is specific for the illumination and it also illustrates the kind of shape alterations resulting from the fit. Given the overall similarity of the shapes, further depth cues as the reflection of the grid lines

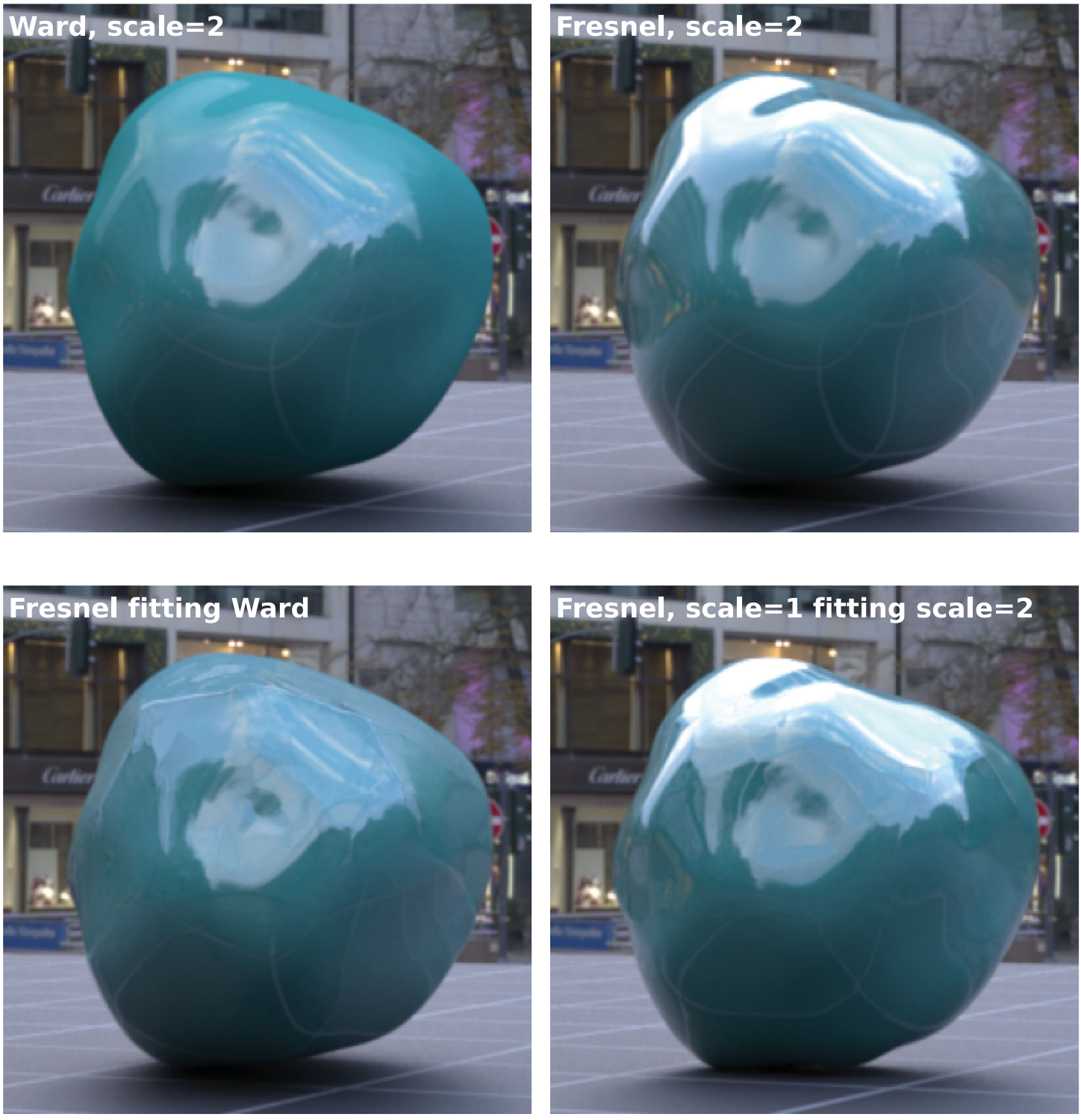

Figure 9. Results of using inverse rendering to find the object's shape for which the rendered image best approximates a target image. (Top) Two target stimuli with depth scaling factor of 2, which are identical except for the BRDF used to render the object (left Ward-BRDF, right Fresnel-BRDF). Both images are rendered with an extended version of Mitsuba 3, but are virtually identical to the ones used in the experiment. The image resolution 256 $\times$ 256 is identical to the resolution used in the fit. (Bottom left) The best result obtained when fitting the rendering with the Ward-BRDF (top left), by changing the geometry of the object rendered with the Fresnel-BRDF (top right). (Bottom right) The result obtained when fitting the object rendered with the Fresnel-BRDF (top right) when using the same BRDF, but starting at a scaling factor of 1.

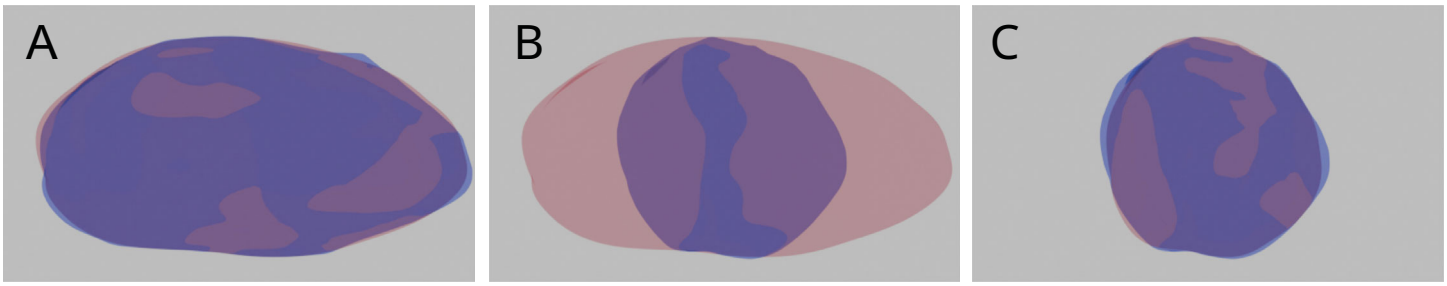

Figure 10. Side view of the shapes resulting from inverse rendering. The corresponding renderings (viewing direction from right to left) are shown in Figure 9. In each panel, the fitted shape is shown as blue transparent object, the reference shape as red transparent object. (A) Fit of a target image rendered with Ward-BRDF. The reference shape is the correct shape with a depth scaling of 2, which was also used as the initial mesh. (B) Fit of the target image rendered with Fresnel-BRDF, when the start mesh has scaling factor 1. The reference shape is the correct shape with a depth scaling of 2. (C) Same as (B), but with the start mesh shown as a reference shape.

drawn on the floor or the form of the shadow are also inconclusive. This means that for a single view of an unknown shape, it can be difficult, to decide whether Fresnel effects are correctly rendered.

This fit is possible because, with the Fresnel-BRDF, the reflection strength of a surface patch depends monotonically on its angle to the viewing direction. Thus, the strength of reflection at a certain image position can be controlled by varying the orientation of the corresponding surface patch: If the reflection strength is too low, the angle must be increased, and vice versa. The only constraint comes from the requirement that the surface should be continuous and reasonably smooth.

Although we have not tested this explicitly, it is to be expected that the reverse case of a fit with the Ward-BRDF to a target rendered with Fresnel-BRDF is more problematic or even impossible, because with the Ward-BRDF there is no consistent coupling of shape and reflection strength. It therefore seems impossible to reproduce the large changes in reflection strength resulting from Fresnel effects without also allowing changes in the illumination.

Fitting with the Fresnel-BRDF

An additional goal was to examine whether the shape of the object can be exactly reproduced with inverse rendering when the *same* BRDF is used. This is, of course, trivially true if the correct shape is used as the starting point in the fit procedure. However, if a different starting point is used, for example, the correct shape with a clearly different depth scaling factor, it could well be that a different shape would result in a similar rendered image, that is, that the fitting would get stuck in a local minimum. This would also indicate an ambiguity in the input with respect to object depth.

For this test, the target image was rendered with the Fresnel-BRDF and a scaling factor of 2 (see top right panel in Figure 9). As the initial mesh for the fit procedure, the same shape was used but with a depth scaling factor of 1. The bottom right panel in Figure 9 shows the rendering obtained for the best fit. Here too, the rendered image from the fit is very similar to the target image at first glance, although on closer inspection some differences can be seen in the shadow region (it is smaller), in the reflections of the grid texture (their positions on the surface are different), and in the outline (parts of surface on the upper outline are missing). However, without the possibility of directly compare target and fit, it would be difficult to recognize these differences.

Figures 10B and 10C compare the shape resulting from the fit with the correct shape and the initial shape, respectively. Again, only very subtle adjustments of the initial mesh were necessary to closely approximate the reflection strength obtained with the correct shape. Conversely, this means that the difference to the correct shape is very large. This result confirms that the available depth cues, in particular Fresnel effects, do not sufficiently constrain the overall shape, but that small changes in shape are sufficient to produce large changes in the reflection strength.

Given this result, the "compression" of perceived depth relative to true depth in a single view of an unknown object is not surprising. Because, without specific assumptions about the true shape, the rendered image is highly variable.

Assuming rigidity

In the previous section, we allowed arbitrary changes in the object's shape to minimize the differences of the rendered image to a target image observed either with a different BRDF or with a different depth scaling factor. Here, we evaluate the fit under the assumption of (partial) "rigidity," that is, when only a depth scaling of the otherwise fixed correct mesh is allowed. To this end, we rendered the object for depth scaling factors

between 0.15 and 3.99. We then computed the loss = mean($|T - C|$) between a target image T rendered for one of five depth scaling factors and comparisons images C rendered across the whole range of depth scaling factors. T and C are vectors resulting from flattening RGB images in linear coordinates, that is, the loss is computed in the same way as in the inverse rendering described earlier. For the target images, either the Ward-BRDF or the Fresnel-BRDF was used. The loss was computed for comparisons that were also either rendered with the Fresnel-BRDF (as in the previous section) or alternatively with the Ward-BRDF.

Rows 1 to 3 in Figure 11 summarize the results for comparisons rendered with the Fresnel-BRDF, rows 4 to 6 show the corresponding results for comparisons rendered with the Ward-BRDF. The loss functions shown in the diagrams (rows 1 and 4) confirm that the minimal loss occurs at the correct depth scaling, when the same BRDF is used in target and comparison (although the loss curve is more shallow for the Ward-BRDF). For the present investigation, the case of different BRDFs in target and comparison is more interesting. Here, the results depend on the choice of the comparisons BRDF: For a Ward target and a Fresnel comparison (row 1), the minimum loss was consistently observed at relatively low depth scaling factors for the comparison object. Furthermore, as the depth scaling factor for the target increases, the loss functions are not only shifted to higher values but also get flatter and contain local minima. Although for a Fresnel target and a Ward comparison (row 4) the loss functions also increase overall with an increasing depth factor for the target, its minimum is still close to the true depth factor and grows monotonically with increasing distance from this minimum.

What is the reason for this asymmetry? A plausible explanation can be derived from examining the rendered images in Figure 11. Row 2 shows that, when using the Ward-BRDF for the target, increasing depth leads to a change in the overall reflection pattern but to almost no change in mean specular reflection. In contrast, when using the Fresnel-BRDF for the target both the reflection pattern changes and the mean specular reflection increases (row 5). Thus, when searching for the best Fresnel comparison for a Ward target with large depth, an increase in the depth of the comparison improves the correspondence in the pattern and simultaneously increases the difference in mean reflectance. This reasoning may suggest that the situation would improve when the images shown in row 2 are replaced with corresponding images that led to the best match to the Fresnel-BRDF in the experiment. However, this does not significantly change the shape of the loss function, because the increases in reflectance made in the matches occur at the wrong locations (predominantly in the center instead of near the rim). In the converse case of a Fresnel target and a Ward comparison, the situation is different. Here the overall reflectance of the Ward comparison hardly changes with depth. Thus, the main factor that influences the loss is the similarity in the gloss pattern, which is most similar at the correct depth.

If one assumes that the visual system uses Fresnel effects as a cue in object recognition, then interpreting an image rendered with the Ward-BRDF is more similar to the case of a Ward target and a Fresnel comparison. The observation that in this case the depth factors of the best match were found to be much lower than the true depth factor may also be related to the reduced depth impression experienced when objects are rendered with a Ward-BRDF.

In this analysis, we have again focused on the case with outdoor illumination and blob shape, where depth scaling leads to the largest difference between Ward-BRDF and Fresnel-BRDF. Changing only the shape to sin-blob, leads to qualitatively similar but slightly less pronounced results. Changing the illumination has a stronger effect, because the indoor illumination has two prominent local illumination sources that produce two highlights near the object's center. Depth scaling leads to large changes in the form of the mirror image of the local highlight. This not only elicits more similar gloss impressions with both BRDFs, but also makes it much easier to find the true depth scaling (see Figure A4 in the Appendix).

Discussion

The main aim of the present study was to explore the influence of Fresnel effects on gloss impressions, gloss constancy, and perceived depth. We first showed that, owing to Fresnel effects, the depth scaling of sphere-like objects can lead to strong changes in the reflection pattern despite fixed physical material parameters. For this reason, this situation seems especially useful to explore the influence of Fresnel effects on gloss-related perceptual properties. In an experiment, we compared gloss and perceived depth observed in depth-scaled objects with fixed material parameters that were rendered with and without Fresnel effects.

A first result was that, with Fresnel effects, the rating of perceived gloss increased strongly and systematically with depth. In contrast, with the Ward-BRDF (i.e., without Fresnel effects), the gloss rating remained approximately constant. Accordingly, the reflection strength parameter ρ_s of the Ward-BRDF had to be significantly increased in the matching task to achieve a similar gloss impression as with Fresnel effects. A first conclusion is that, in the situation realized in the experiment, changes in the reflection strength due to Fresnel effects are not correctly attributed to changes in the shape of the object, but at least partly to changes in the material. One can also conclude that the transformation $\rho_s = (IOR - 1) * 0.17$, which was used

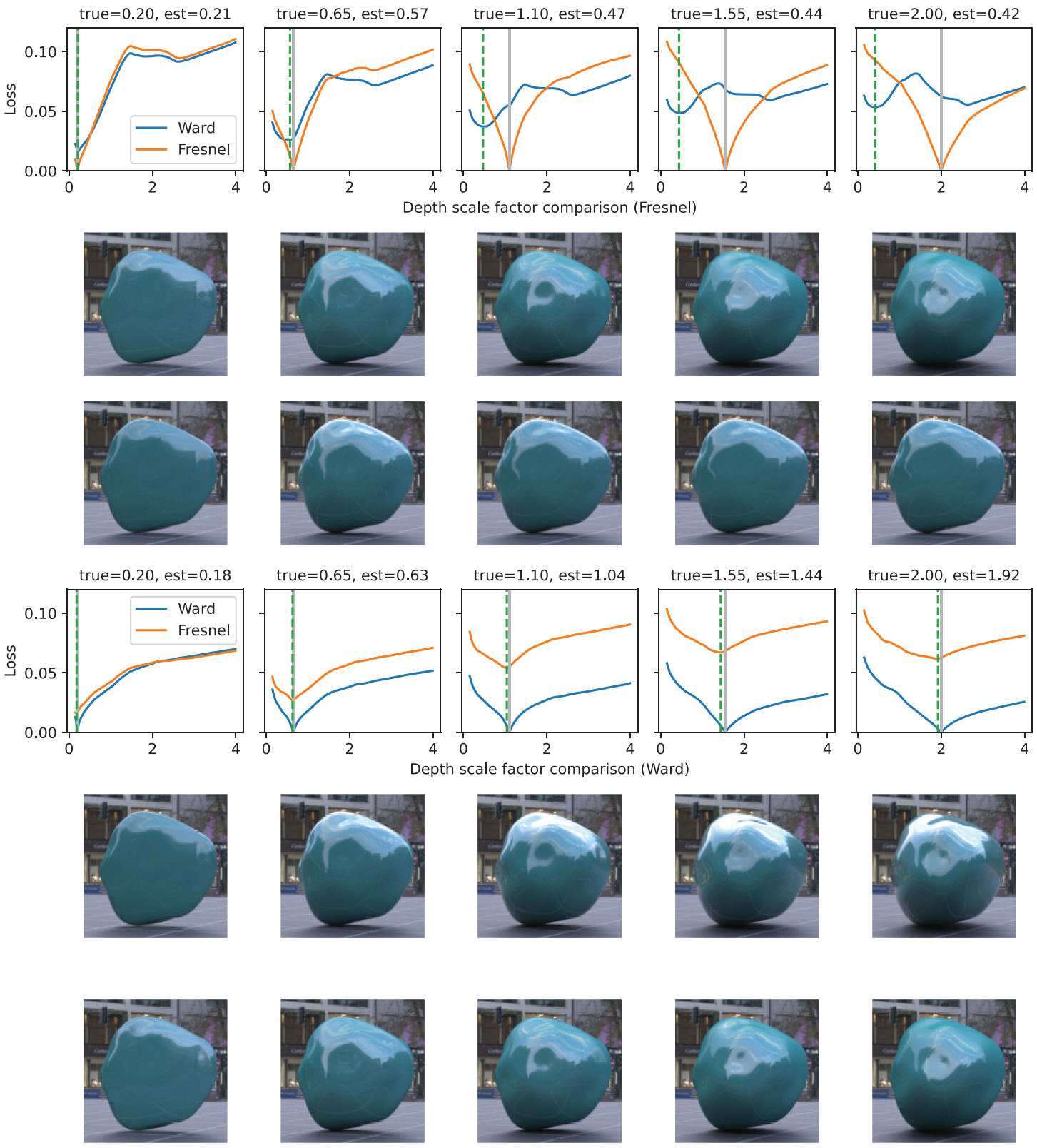

Figure 11. Best fit under the assumption of partial rigidity (i.e., fixed shape up to depth scaling). The diagrams in the first and fourth rows show the loss between target and comparison image as a function of the scaling factor in the comparison. Five true scaling factors are used for the target, which is either rendered with the Ward-BRDF or the Fresnel-BRDF. The comparisons are either rendered with Fresnel-BRDF (rows 1–3), or with Ward-BRDF (rows 4–6). Row 2 shows the target image with Ward-BRDF, row 5 those with Fresnel-BRDF. Rows 3 and 6 depict the comparison with the minimum loss (the corresponding depth scale factor is shown in the plots as dashed vertical line).

in Faul (2019) to achieve similar gloss impressions with both BRDFs, does not generalize to arbitrary shapes. The fact that this transformation depends on shape makes it difficult to approximate the gloss impression of real glossy materials using the Ward-BRDF, especially if in dynamic stimuli the shapes of the objects change over time. The results also suggest that in the situation we investigated, the degree of gloss constancy observed with real materials can be significantly *overestimated* if the Ward-BRDF is used to simulate material reflectance.

Two of the tasks in the experiments were dedicated to evaluating the influence of Fresnel effects on perceived depth. A first task was to judge the relative perceived depth of objects of identical shape that were rendered with different BRDFs. Here, we found that, when the real depth of the objects was increased, the higher perceived depth increasingly shifted toward the object rendered with Fresnel-BRDF. This means that, with Fresnel effects, the depth impression increases faster with real depth than without Fresnel effects. In the second task, the objects were rendered with the same BRDF, and the participants had to judge the relative perceived depth of objects with different real depths. In this case, there exists an objective standard with which the ratings can be compared. We found that the expected increase of perceived depth with increasing real depth was steeper and more regular with the Fresnel-BRDF than with the Ward-BRDF. Although the perceived depth underestimated the real depth with both BRDFs, this result supports the conclusion that shape constancy is higher when Fresnel effects are correctly simulated. Moreover, the findings from both tasks are in line with the informal observation made in Faul (2019) that Fresnel effects often enhance the depth impression of glossy objects.

We also investigated the extent to which regularities inherent in Fresnel effects constrain the shape of the object when (as in our and many other experiments on material perception) a single monocular view of a static object is given. To this end, we tried to predict the true shape from a target image rendered with or without Fresnel effects, assuming a Fresnel-BRDF in the reconstruction. The shape varied during the fit was either fixed to the actual shape except for a depth scaling factor, or allowed to change arbitrarily.

In the former case of partial rigidity, the true shape, that is, the correct depth scaling factor, could be correctly predicted when the target image was rendered with the Fresnel-BRDF: The loss functions comparing target and comparison had a single minimum at the correct depth. Was the target image instead rendered with the Ward-BRDF then the loss function contained local minima and the true depth was systematically underestimated. This result may partly explain why the depth impression is reduced for objects rendered with the Ward-BRDF.

If arbitrary changes to the shape were allowed, only relatively minor local changes to the overall shape were needed to approximately reproduce a target image that was either rendered without Fresnel effects or with a shape that considerably differed from the initial shape used in the fit. The latter result shows that Fresnel effects on its own are not very effective in disambiguating overall shape. This finding is consistent with the informal observation that even with correct Fresnel effects, it is difficult to recognize the true depth of objects when only a single monocular view is available. Furthermore, owing to the dependence of reflection strength on viewing direction, which is the basis of Fresnel effects, it is often sufficient to slightly change the orientation of surface patches to approximate a wide range of reflection pattern, for example those resulting from the Ward-BRDF. Because the exact shape is usually not known, it can be difficult to detect that a physically incorrect BRDF was used in the rendering. This factor may contribute to the fact that even rough approximations to the true BRDF can often produce a robust gloss impression.

Conclusions

Our results with depth scaled glossy objects corroborate previous findings that Fresnel effects can have large effects on gloss impressions and gloss constancy. They also show that Fresnel effects have an effect on perceived depth. Ignoring Fresnel effects in cases where the shapes of objects change over time may lead to unrealistic gloss impressions, to an overestimation of gloss constancy and to a reduced perceived depth. Our investigation with inverse rendering indicates that Fresnel effects alone are not very effective to disambiguate overall shape in a single monocular view of objects.

Keywords: scene perception, image analysis, material perception, three-dimensional surface and shape perception

Acknowledgments

Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - grant number FA425/3-2.

Commercial relationships: none.
Corresponding author: Franz Faul.
Email: ffaul@psychologie.uni-kiel.de.
Address: Institut für Psychologie, Universität Kiel, Kiel 24118, Germany.

References

Adams, W. J., & Elder, J. H. (2014). Effects of specular highlights on perceived surface convexity. *PLoS Computational Biology, 10*(5), e1003576, https://doi.org/10.1371/journal.pcbi.1003576.

Adato, Y., Vasilyev, Y., Ben-Shahar, O., & Zickler, T. (2007). Toward a theory of shape from specular flow. *2007 IEEE 11th International Conference on Computer Vision*, 1–8, https://doi.org/10.1109/ICCV.2007.4408883.

Anderson, B. L., & Marlow, P. J. (2023). Perceiving the shape and material properties of 3D surfaces. *Trends in Cognitive Sciences, 27*(1), 98–110, https://doi.org/10.1016/j.tics.2022.10.005.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48, https://doi.org/10.18637/jss.v067.i01.

Beck, J., & Prazdny, S. (1981). Highlights and the perception of glossiness. *Perception & Psychophysics, 30*(4), 407–410, https://doi.org/10/ffs4zm.

Belhumeur, P., Kriegman, D., & Yuille, A. (1997). The basrelief ambiguity. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1060–1066, https://doi.org/10.1109/CVPR.1997.609461.

Christensen, R. H. B. (2023). Ordinal—regression models for ordinal data. R package version 2023.12-4.1, https://CRAN.R-project.org/package=ordinal.

Doerschner, K, Maloney, L. T., & Boyaci, H. (2010). Perceived glossiness in high dynamic range scenes. *Journal of Vision, 10*(9), 11, https://doi.org/10.1167/10.9.11.

Doerschner, K., Fleming, R.W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. *Current Biology, 21*(23), 2010–2016, https://doi.org/10.1016/j.cub.2011.10.036.

Dövencioğlu, D. N., Ben-Shahar, O., Barla, P., & Doerschner, K. (2017). Specular motion and 3D shape estimation. *Journal of Vision, 17*(6), 3, https://doi.org/10.1167/17.6.3.

Faul, F. (2019). The influence of Fresnel effects on gloss perception. *Journal of Vision, 19*(13), 1, https://doi.org/10.1167/19.13.1.

Faul, F. (2021). Perceived roughness of glossy objects: The influence of Fresnel effects and correlated image statistics. *Journal of Vision, 21*(8), 1, https://doi.org/10/gmdg6t.

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, *4*(9), 10, https://doi.org/10.1167/4.9.10.

Geisler-Moroder, D., & Dür, A. (2010). A new Ward BRDF model with bounded Albedo. *Eurographics Symposium on Rendering, 29*, 1391–1398, https://doi.org/10.1111/j.1467-8659.2010.01735.x.

Jakob, W. (2010). Mitsuba renderer. Retrieved from https://mitsuba-renderer.org.

Jakob, W., Speierer, S., Roussel, N., Nimier-David, M., Vicini, D., Zeltner, T., ... Zhang, Z. (2022). Mitsuba 3 renderer. Retrieved from https://mitsuba-renderer.org.

Kim, J., Marlow, P., & Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision, 11*(9), 4, https://doi.org/10.1167/11.9.4.

Marlow, P. J., Kim, J., & Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *Journal of Vision, 11*(9), 16, https://doi.org/10.1167/11.9.16.

Marlow, P. J., & Anderson, B. L. (2024). Interactions between 3D surface shape and material perception. *Annual Review of Vision Science, 10*, https://doi.org/10.1146/annurev-vision-102122-094213.

Marlow, P. J., Todorović, D., & Anderson, B. L. (2015). Coupled computations of three-dimensional shape and material. *Current Biology, 25*(6), R221–R222, https://doi.org/10.1016/j.cub.2015.01.062.

Muryy, A. A., Welchman, A. E., Blake, A., & Fleming, R.W. (2013). Specular reflections and the estimation of shape from binocular disparity. *Proceedings of the National Academy of Sciences, of the USA, 110*(6), 2413–2418, https://doi.org/10.1073/pnas.1212417110.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2006). Shape-from-shading for matte and glossy objects. *Acta Psychologica, 121*(3), 297–316, https://doi.org/10/b8m226.

Nicolet, B., Jacobson, A., & Jakob, W. (2021). Large steps in inverse rendering of geometry. *ACM Transactions on Graphics, 40*(6), 1–13, https://doi.org/10.1145/3478513.3480501.

Nishida, S., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A, 15*(12), 2951–2965, https://doi.org/10.1364/JOSAA.15.002951.

Olkkonen, M., & Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception, 2*(9), 1014–1034, https://doi.org/10.1068/i0480.

Phong, B. T. (1975). Illumination for computer generated pictures. *Communications of the ACM, 18*(6), 311–317, https://doi.org/10/bkfrm9.

R Core Team. (2024). R: A language and environment for statistical computing. Manual. The R Core Team: Vienna, Austria.

Savarese, S., Chen, M., & Perona, P. (2005). Local shape from mirror reflections. *International Journal of Computer Vision, 64*(1), 31–67, https://doi.org/10.1007/s11263-005-1086-x.

Schlüter, N., & Faul, F. (2019). Visual shape perception in the case of transparent objects. *Journal of Vision, 19*(4), 24, https://doi.org/10.1167/19.4.24.

Serrano, A., Chen, B., Wang, C., Piovarči, M., Seidel, H.-P., Didyk, P., ... Myszkowski, K. (2021). The effect of shape and illumination on material perception: Model and applications. *ACM Transactions on Graphics, 40*(4), 1–16, https://doi.org/10.1145/3450626.3459813.

Todd, J. T., Norman, J. F., Koenderink, J. J., & Kappers, A. M. L. (1997). Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception, 26*(7), 807–822, https://doi.org/10.1068/p260807.

Todd, J. T., Norman, J. F., & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science, 15*(1), 33–39, https://doi.org/10.1111/j.0963-7214.2004.01501006.x.

Vangorp, P., Laurijssen, J., & Dutré, P. (2007). The influence of shape on the perception of material reflectance. *ACM SIGGRAPH 2007 Papers*, 77–es, https://doi.org/10/fnv497.

Walter, B., Marschner, S. R., Li, H., & Torrance, K. E. (2007). Microfacet models for refraction through rough surfaces. *Proceedings of the 18th Eurographics Conference on Rendering Techniques,* 195–206, https://doi.org/10.2312/EGWR/EGSR07/195-206.

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*, 265–272, https://doi.org/10.1145/133994.134078.

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of Vision, 10*(9), 7, https://doi.org/10.1167/10.9.7.

Wijntjes, M.W. A., & Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *Journal of Vision*, *10*(9), 13, https://doi.org/10.1167/10.9.13.

Zhang, Z., Roussel, N., & Jakob, W. (2023). Projective sampling for differentiable rendering of geometry. *ACM Transactions on Graphics, 42*(6), 1–14, https://doi.org/10.1145/3618385.

Appendix

Details of the stimulus generation

All stimuli used in the experiment were rendered with Mitsuba 0.6, all those used in inverse rendering with (an extended version of) Mitsuba 3.5. Both renderers produce nearly identical outputs for identical scene parameters. The "roughplastic" material in these programs implements the BRDF described by Walter et al. (2007), the "ward" material the original Ward-BRDF (Ward, 1992).

In both BRDFs, the roughness parameter α was set to 0.01, a relatively small value that still produces quite sharp mirror images. The Beckmann microfacet distribution was chosen for the Fresnel-BRDF. For a given α, both BRDFs produce similar perceived roughness (see also Faul, 2021). The diffuse reflection was set to rgb = (0.03067, 0.1960, and 0.167). A Gaussian filter with $\sigma = 0.5$ was used to post-process the images. The experimental stimuli were rendered with 512 samples per pixel with the "bdpt" integrator (a bidirectional path tracer).

The original Ward model (Ward, 1992) as implemented in Mitsuba 0.6 (and ported by us to Mitsuba 3.5) was used in its isotropic variant:

$$f(\theta_i, \phi_i, \theta_o, \phi_o) = \frac{\rho_d}{\pi} + \rho_s \frac{\exp(-\tan^2 h/\alpha^2)}{4\pi\alpha^2\sqrt{\cos\theta_i \cos\theta_o}},$$

where θ_i, and ϕ_i denote polar and azimuthal angles, respectively, that together define the directions $\omega_i = (\theta_i, \phi_i)$ and $\omega_o = (\theta_o, \phi_o)$ of the incoming and reflected light, respectively. The variable h in the model equation is the half angle between ω_i and ω_o. The model has three parameters: roughness α, strength of the diffuse reflection ρ_d, and strength of the specular reflection ρ_s.

As the reflection strength parameter ρ_s is arbitrary, that is, has no well-defined relationship with the reflection strength in the Fresnel-BRDF (which depends on the IOR), we used the transformation $\rho_s = (IOR - 1) * 0.17$, because this was found in Faul (2019) to produce similar perceived gloss strengths with both BRDFs. However, a main result of the present study is that such a transformation cannot be universal, but only works under certain restricted (shape) conditions. The reason for this is that in general the perceived gloss strength changes differently for the two BRDFs when the object shape is varied.

Statistical analyses

We analyzed the results of both subexperiments with mixed linear models using the packages "lme4" (version 1.1-35.5, Bates, Mächler, Bolker, & Walker, 2015) and

"ordinal" (version 2023.12-4, Christensen, 2023) for R 4.4.1 (R Core Team, 2024). The "lmer" procedure was used for the matching data, whereas gloss ratings and relative depth ratings were fitted with the procedure "clmm" for ordinal data using a "logit" link.

The design of both subexperiments is balanced and fully crossed. Subexperiment 1 comprises the four factors 1) depth scaling ("dscale"), 2) reflection strength ("ior"), 3) global illumination ("ill"), and 4) shape ("shape"). Subexperiment 2 contains in addition the factor BRDF ("brdf"). From a theoretical perspective, the factors 1 through 4 are best regarded as random factors, as the levels of these factors were selected more or less arbitrarily from many alternatives. However, since factors 2 through 4 have only two to three levels, which makes it difficult to reliably estimate variance components, they were considered as fixed factor.

In subexperiment 1, we started for all dependent variables Y with the model: $Y \sim shape*ill*spec + (1|dscale)$, and in subexperiment 2 with the model $Y \sim shape*ill*spec*brdf + (1 + brdf|dscale)$. Because repeated measures from 10 participants were conducted, we always added a random factor $(1|subj)$ to account for systematic differences between subjects. In some cases, this factor also included random slopes depending on illumination (i.e., $1 + ill|subj$) was used. This allowed us to account for any systematic differences in how participants responded to changes in the illumination.

For each dependent variable tested, we first simplified the fixed effects part of the models. This was done based on estimates of the fixed factors obtained with the start model and the results of likelihood ratio tests between nested models. These simplified final models were used to test our hypotheses. The

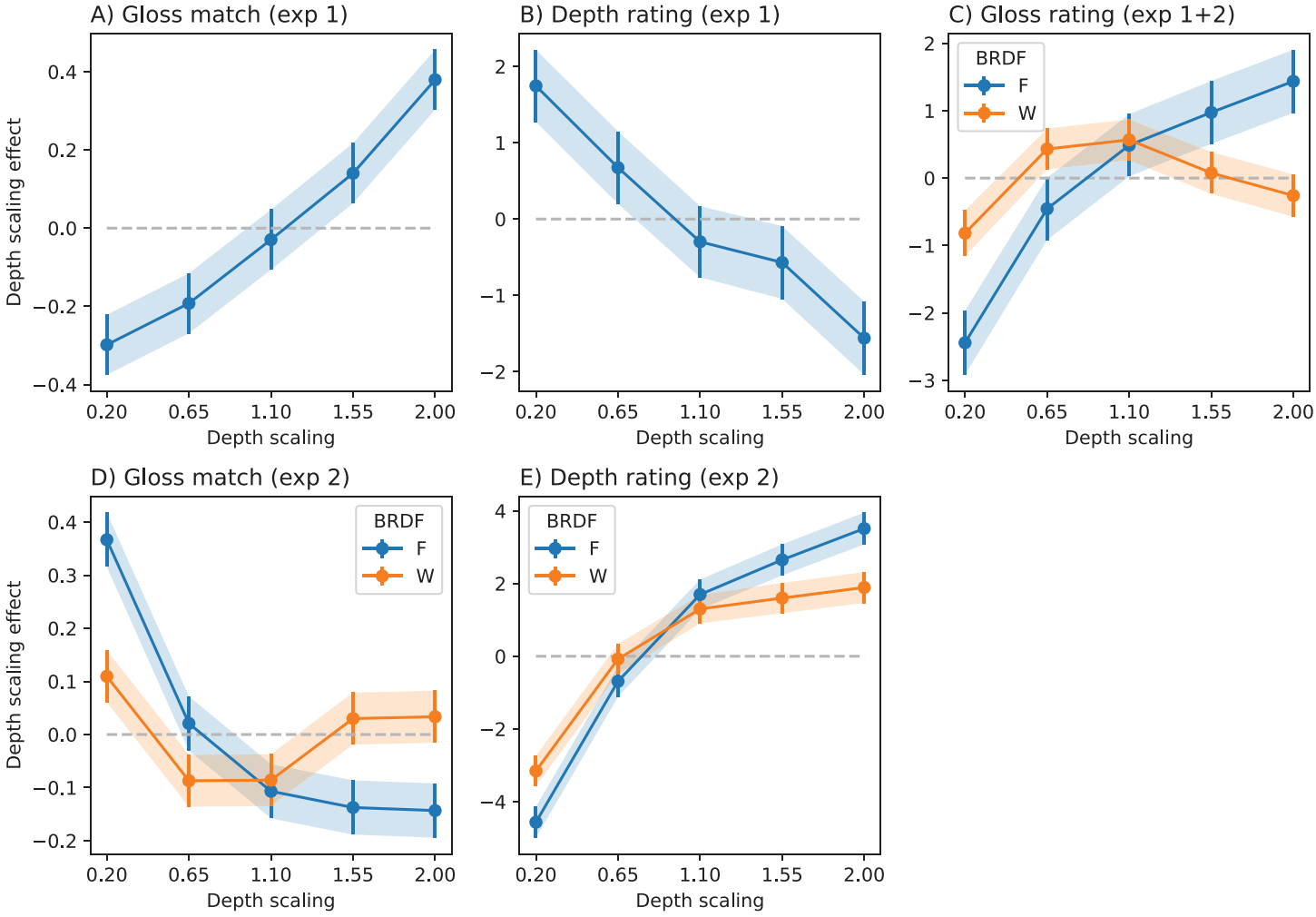

Figure A1. Effects of the random factor "depth scaling" on different dependent variables in subexperiments 1 and 2 as estimated by corresponding mixed linear models (see text for details). The shaded regions correspond to ±1.96 standard deviations of the estimates. (A) Effect on the gloss matches across different BDRF in subexperiment 1, (B) effect on the relative depth ratings in subexperiment 1 (values < 0 mean "Fresnel deeper"), (C) effect on gloss ratings estimated from combined data of subexperiments 1 and 2, (D) effect on gloss matches in subexperiment 2, and (E) effect on relative depth ratings in subexperiment 2.

hypothesis for subexperiment 1 is that the random factor (1|*dscale*) is significant. For subexperiment 2, there are two hypotheses. First that the random factor (1 + *brdf*|*dscale*) is significant, and in particular also the differences between the random slopes estimated for each BRDF. The significance testing was done by conducting likelihood ratio tests between models with and without the corresponding factor. We also report the estimated random effects for the "dscale" factor (see plots in Figure A1), to show how depth scaling affects the settings according to the models. These estimates are in good agreement with the group means shown in Figures 7 and 8.

Gloss matching in subexperiment 1

The final model for the dependent variable Y (= ior set in the match stimulus) is $Y \sim ior + ill + (1|dscale) + (1 + ill|subj)$. The model was fitted with "lmer." A likelihood ratio test against the same model without the "dscale" factor is significant, $\chi^2(1) = 355.1$, $p < 0.001$. The results for fixed and random effects are shown in Table A1, the random effects of depth are visualized in Figure A1A. This analysis indicates that the reflection strength settings in the Ward-BRDF have to increase monotonically with increasing depth scaling to achieve a gloss match with the Fresnel-BRDF.

Relative depth ratings in subexperiment 1

The final model for the dependent variable Y (= relative depth rating) is $Y \sim ior*ill + (1|dscale) + (1 + ill|subj)$. The model was fitted with "clmm." A likelihood ratio test against the same model without the "dscale" factor is significant, $\chi^2(1) = 172.96$, $p < 0.001$. The results for fixed and random effects are shown in Table A2, the random effects of depth are visualized in Figure A1B. This analysis indicates that the perceived depth of the object with Fresnel-BRDF increase with depth scaling relative to the perceived depth of the object with Ward-BRDF.

Groups	Name	SD		
Random effects				
subj	intercept	0.190		
	ill:outdoor	0.184		
dscale	intercept	0.271		
	Est	Std.err	*t*	*p*
Fixed effects				
intercept	1.19	0.137	8.7	<0.001
ior:1.51	0.33	0.026	12.3	<0.001
ior:1.75	0.63	0.026	24.5	<0.001
ill:outdoor	0.25	0.062	4.0	0.003

Table A1. Fixed and random effects for the gloss matches in subexperiment 1.

Groups	Name	SD		
Random effects				
subj	intercept	0.565		
	ill:out	0.825		
dscale	intercept	1.154		
	Est	Std.err	*z*	*p*
Fixed effects				
ior:1.51	−0.81	0.268	−3.0	0.002
ior:1.75	−1.50	0.268	−3.2	<0.001
ill:out	−1.56	0.379	−4.2	<0.003
ior:1.51 × ill:out	1.09	0.377	2.9	0.004
ior:1.75 × ill:out	1.00	0.372	3.7	0.007

Table A2. Fixed and random effects for the relative depth ratings in subexperiment 1.

Gloss rating in subexperiment 1 and 2

We combined the gloss ratings depending on BRDF and depth scaling from both subexperiments to increase sample size. The final model for the dependent variable Y (= gloss rating) is $Y \sim ior*ill*brdf + shape + (0 + brdf|dscale) + (1 + ill|subj)$. The model was fitted with "clmm." A likelihood ratio test against a model without "dscale" factor is significant, $\chi^2(3) = 603.53$, $p < 0.001$. This is also true if only the random slopes of this factor are removed, $\chi^2(2) = 231.5$, $p < 0.001$. The results for fixed and random effects are shown in Table A3 and the random effects of "dscale" depending on the BRDF are visualized in Figure A1C. This analysis indicates that the perceived gloss depends on the depth scaling and that this effect differs between the BRDFs: With the Fresnel-BRDF the gloss increases monotonically with depth scaling, whereas this is not the case with the Ward-BRDF.

Gloss matching in subexperiment 2

The final model for the dependent variable Y (= reflection strength setting) is $Y \sim ior + ill*brdf + (0 + brdf|dscale) + (1|subj)$. The model was fitted with "lmer." A likelihood ratio test against a model without the "dscale" factor is significant, $\chi^2(3) = 340.0$, $p < 0.001$. This is also true for a test against the model, in which only the random slopes depending on "brdf" are dropped, $\chi^2(2) = 130.8$, $p < 0.001$. The results for fixed and random effects are shown in Table A4 and the random effects of "dscale" depending on BRDF are visualized in Figure A1D. This analysis indicates that reflection strength settings depend on the depth scaling and that this effect differs between the BRDFs: With

Groups	Name	SD		
Random effects				
subj	intercept	1.401		
	ill:out	0.757		
dscale	brdf:F	1.404		
	brdf:W	0.546		
	Est	Std.err	*z*	*p*
Fixed effects				
ior:1.51	1.87	0.184	10.2	<0.001
ior:1.75	2.95	0.191	15.5	<0.001
ill:out	1.99	0.303	6.6	<0.001
brdf:W	0.72	0.572	1.3	0.211
shape:sin	0.24	0.073	3.3	<0.001
ior:1.51 × ill:out	−0.79	0.255	−3.1	<0.001
ior:1.75 × ill:out	−1.07	0.259	−4.1	0.002
ior:1.51 × brdf:W	−0.91	0.253	−3.6	<0.001
ior:1.75 × brdf:W	−1.41	0.257	−5.5	<0.001
ill:out × brdf:W	−1.83	0.257	−7.1	<0.001
ior:1.51 × ill:out × brdf:W	0.80	0.359	2.2	<0.024
ior:1.75 × ill:out × brdf:W	1.43	0.361	4.0	<0.001

Table A3. Fixed and random effects for the gloss ratings in subexperiment 1 an 2.

Groups	Name	SD		
Random effects				
subj	intercept	0.051		
dscale	brdf:F	0.217		
	brdf:W	0.088		
	Est	Std.err	*t*	*p*
Fixed effects				
intercept	1.32	0.100	13.2	<0.001
ior:1.51	0.25	0.017	15.2	<0.001
ior:1.75	0.50	0.017	29.9	<0.001
ill:out	0.11	0.019	5.7	<0.001
brdf:W	0.02	0.086	0.4	<0.721
ill:out × brdf:W	−0.07	0.027	−2.7	0.006

Table A4. Fixed and random effects for gloss matches in subexperiment 2.

the Fresnel-BRDF the settings decreases monotonically with depth scaling, whereas this is not the case with the Ward-BRDF.

Relative depth rating in subexperiment 2

The final model for the dependent variable Y (= relative depth rating) is $Y \sim shape + (1 + brdf|dscale) + (1|subj)$. The model was fitted with "clmm." A likelihood ratio test against the same model without the "dscale" factor is significant, $\chi^2(3) = 1018.2$, $p < 0.001$. This is also true when only the random slope of this factor is dropped, $\chi^2(2) = 92.5$, $p < 0.001$. The results for fixed and random effects are shown in Table A5 and the random effects of "dscale" depending on BRDF are visualized in Figure A1E.

Groups	Name	SD		
Random effects				
subj	intercept	0.445		
dscale	intercept	2.961		
	brdf:W	1.139		
	Est	Std.err	*z*	*p*
Fixed effects				
shape:sin	0.426	0.111	3.8	<0.001

Table A5. Fixed and random effects for relative depth ratings in subexperiment 2.

Gloss matching in subexperiment 1 across conditions

In Figure 7 in the main text, we report only the overall gloss matching results. As an illustration of the effect of shape and illumination, Figure A2 shows the corresponding results for all combinations of shape and illumination, presented separately. It can be seen that the overall pattern of results is similar in all combinations, but that the results are more pronounced with the outdoor illumination, which is more homogeneous than the indoor illumination and more intense in regions near the rim, where the reflection is larger than in the center. The differences between the two shapes are less prominent. This is also corroborated by the statistical analysis (see section 'statistical analyses' in the Appendix).

Best fit under a change of illumination

Figure 9 in the main text shows that it is possible to approximate a target image that was rendered with a different BRDF or a different shape by slightly adjusting the geometry. The two top rows in Figure A3 present the corresponding results obtained with the indoor illumination, for which the renderings with different BRDFs are overall more similar. In the lower row, the shapes resulting from the corresponding fit with the outdoor illumination is used. The differences between the middle and the lower row show that the fit is specific to the illumination, and they also illustrate the kind of shape distortions resulting from the fit.

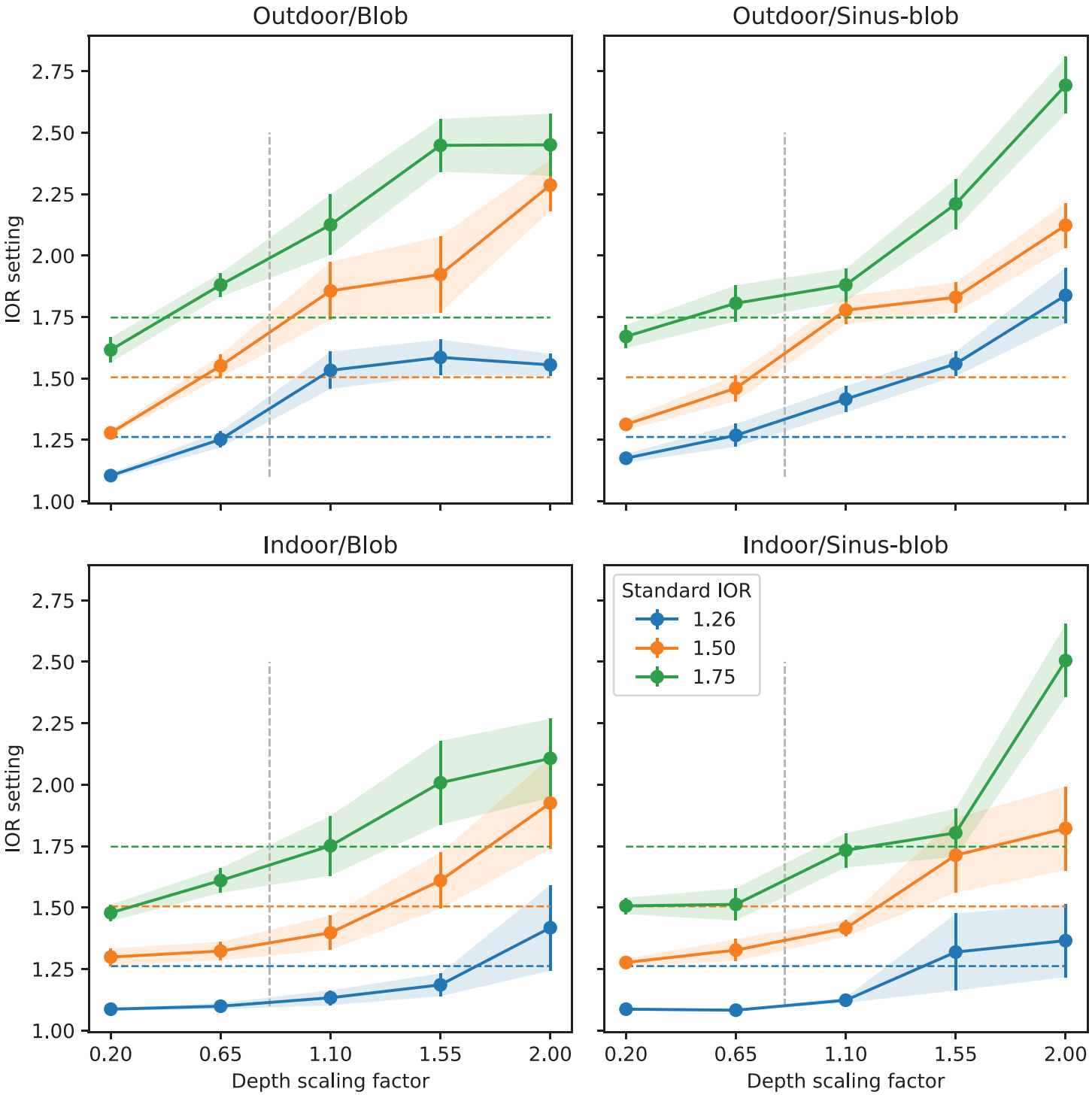

Figure A2. Results of matching the gloss in the standard object rendered with Fresnel-BRDF by adjusting the reflection strength in the comparison object rendered with the Ward-BRDF, for all four combinations of shape and illumination (see Figure 7 for the overall effect).

Best fit under the rigidity constraint for the indoor illumination

Figure 11 in the main text shows the results of searching for the correct shape of an object depicted in a target image rendered with BRDF 1, by minimizing the difference to comparison images, in which the object is rendered with BRDF 2 for varying depths. In this figure, we used a blob object with the outdoor illumination. Figure A4 shows the corresponding results with an indoor illumination. The results for the sinus-blob are not reported as the change in illumination proved to be much more critical than a change in shape.

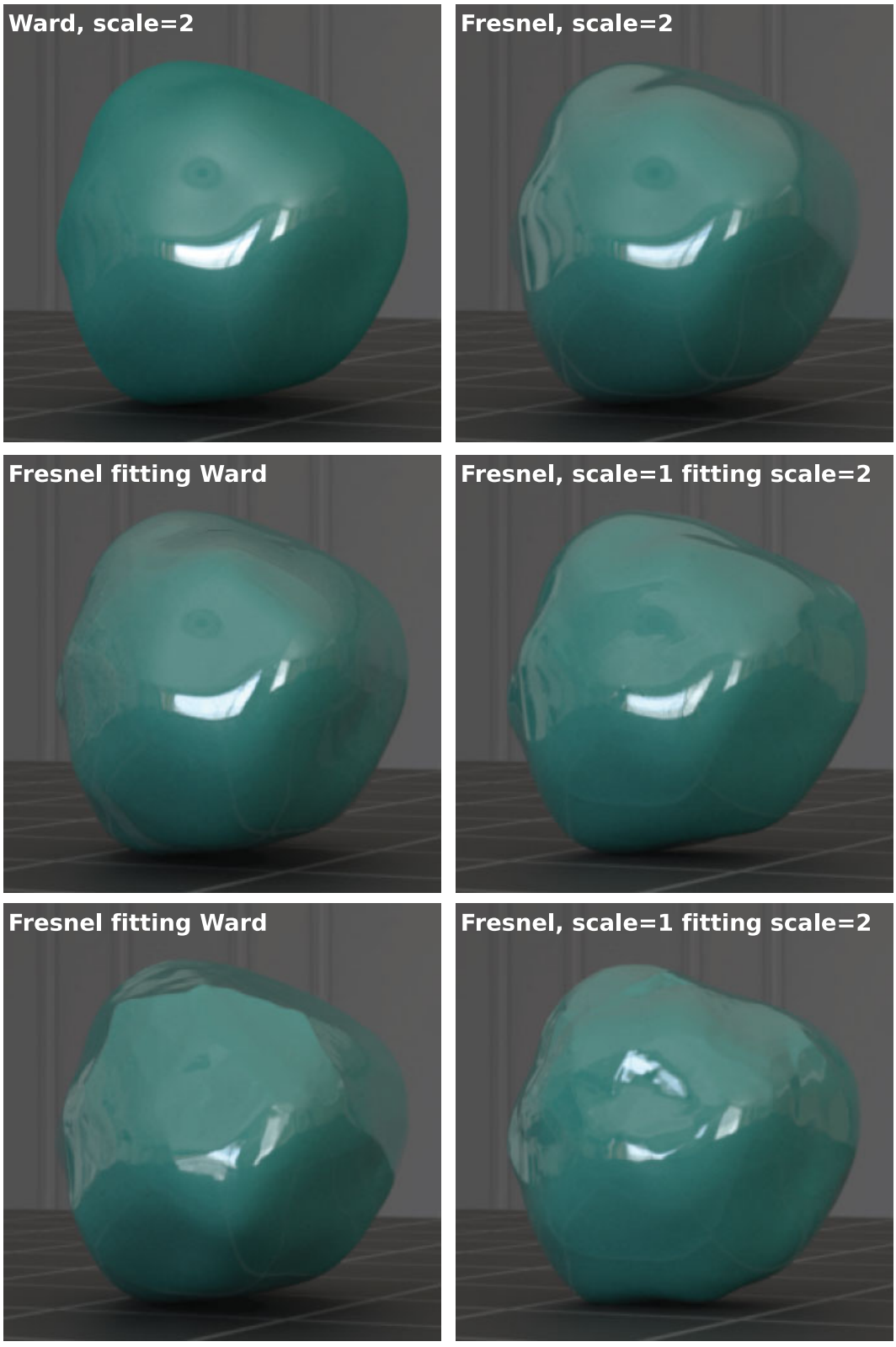

Figure A3. The two top rows are identical to Figure 9 in the main text, with the exception that the fit and the rendering was done with the indoor illumination. In the lower row, the shape resulting from the fit with the indoor illumination is used, that is, fit and render illuminations are different.

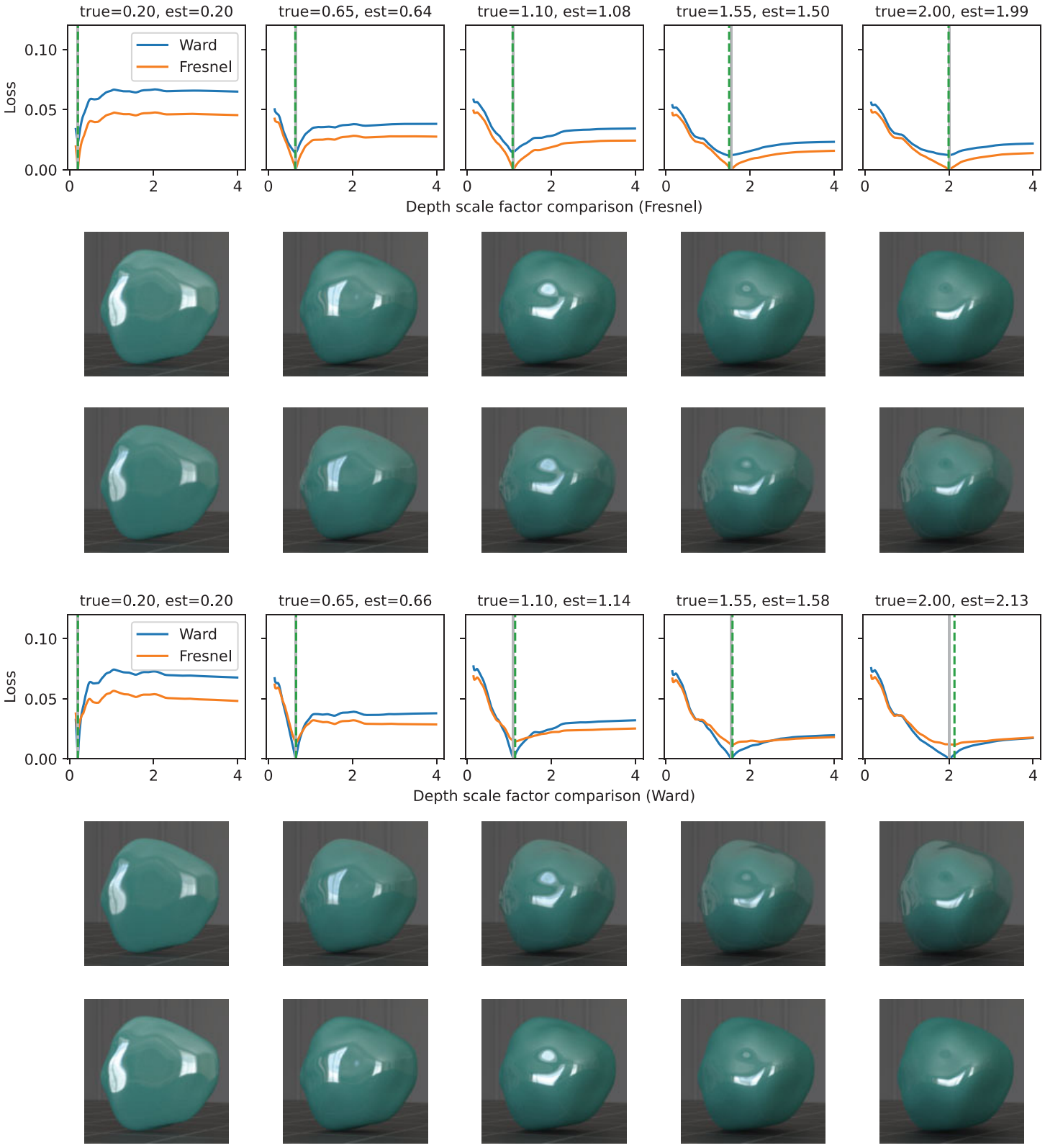

Figure A4. Best fit under the assumption of partial rigidity (fixed shape up to depth scaling). Same as Figure 11 in the main text, but with indoor illumination. See caption of that figure for more information.