



Delft University of Technology

Highlight shapes and perception of gloss for real and photographed objects

van Assen, Jan Jaap R.; Wijntjes, Maarten; Pont, Sylvia

DOI

[10.1167/16.6.6](https://doi.org/10.1167/16.6.6)

Publication date

2016

Document Version

Final published version

Published in

Journal of vision

Citation (APA)

van Assen, J. J. R., Wijntjes, M., & Pont, S. (2016). Highlight shapes and perception of gloss for real and photographed objects. *Journal of vision*, 16(6), 1-14. <https://doi.org/10.1167/16.6.6>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

Highlight shapes and perception of gloss for real and photographed objects

Jan Jaap R. van Assen

Department of Psychology, Justus-Liebig-University
Giessen, Giessen, Germany



Maarten W. A. Wijntjes

Perceptual Intelligence Laboratory, Delft University of
Technology, Delft, The Netherlands



Sylvia C. Pont

Perceptual Intelligence Laboratory, Delft University of
Technology, Delft, The Netherlands



Gloss perception strongly depends on the three-dimensional shape and the illumination of the object under consideration. In this study we investigated the influence of the spatial structure of the illumination on gloss perception. A diffuse light box in combination with differently shaped masks was used to produce a set of six simple and complex highlight shapes. The geometry of the simple highlight shapes was inspired by conventional artistic practice (e.g., ring flash for photography, window shape for painting and disk or square for cartoons). In the box we placed spherical stimuli that were painted in six degrees of glossiness. This resulted in a stimulus set of six highlight shapes and six gloss levels, a total of 36 stimuli. We performed three experiments of which two took place using digital photographs on a computer monitor and one with the real spheres in the light box. The observers had to perform a comparison task in which they chose which of two stimuli was glossiest and a rating task in which they rated the glossiness. The results show that, perhaps surprisingly, more complex highlight shapes were perceived to produce a less glossy appearance than simple highlight shapes such as a disk or square. These findings were confirmed for both viewing conditions, on a computer display and in a real setting. The results show that variations in the spatial structure of “rather simple” illumination of the “extended source” type highlight influences perceived glossiness.

Introduction

All objects in our environment have shape and material properties. We reliably and quickly recognize these material properties (Sharan, Rosenholtz, &

Adelson, 2014) to define the object, to derive the meaning or interactions that can be performed with it. The illumination, object shape, and material properties influence how we perceive an object. We make various kinds of errors perceiving light (Koenderink, van Doorn, & Pont, 2004; Ostrovsky, Cavanagh, & Sinha, 2005; Pont & Koenderink, 2007), shape (Ho, Landy, & Maloney, 2008; Wijntjes, Volcic, Pont, Koenderink, & Kappers, 2009), and materials (Marlow, Kim, & Anderson, 2012; VanGorp, Laurijssen, & Dutré, 2007; Wijntjes & Pont, 2010), especially in situations where very little visual information is available. But even when visual cues are scarce we feel confident in what we perceive (Koenderink, 2001), for instance, for representations of materials in paintings of Vermeer or photographs. Despite the reduction of information in comparison with reality in these cases, we are generally still able to estimate where the light source is, what the shape of an object is, and which material that object has been made of (although these estimates may not be veridical, see Kartashova, Heynderickx, Sekulovski, & Pont, 2014; Kartashova, te Pas, Pont, de Ridder, & Schoemaker, 2015; te Pas & Pont, 2005). Gloss perception is strongly dependent on the illumination, surface shape, and material properties. A minimal requirement for the effect of gloss is the presence of highlights or lowlights (Kim, Marlow, & Anderson, 2012). That only a single local highlight creates the illusion of global gloss can convincingly be observed in graphical user interfaces or cartoons. These highlights come in a variety of shapes but the precise influence of this highlight shape on the appearance of gloss is unknown.

In many studies on gloss perception novel approaches were introduced to specify which properties

Citation: van Assen, J. J. R., Wijntjes, M. W. A., & Pont, S. C. (2016). Highlight shapes and perception of gloss for real and photographed objects. *Journal of Vision*, 16(6):6, 1–14, doi:10.1167/16.6.6.



make a surface look glossy. One of these approaches is to manipulate image statistics based on the idea that our visual system (partly) uses statistical heuristics to infer material properties. For example, images with a positively skewed luminance histogram were found to appear glossier (Motoyoshi, Nishida, Sharan, & Adelson, 2007). Highlights often generate positively skewed histograms. Motoyoshi et al. (2007) proposed that humans use this skewness to make judgments about the glossiness of rough surfaces. However, such luminance histogram skews may have their origin in various combinations of illumination, material properties, and shape. This limits the set of illuminations, material properties, and object shapes that can be evaluated using only image statistics (Anderson & Kim, 2009; Olkkonen & Brainard, 2010). The illumination characteristics, surface shape, and surface structure of an object also influence the spatial structure of its resulting images. This spatial structure offers important cues to the light, material, and shape that are additional to the image statistics (Pont & Koenderink, 2008; Pont, van Doorn, Wijntjes, & Koenderink, 2015).

Another approach is to study the influence of object shape on glossiness perception. Highlights concentrate at points with high curvature (Koenderink & van Doorn, 1980). Vangorp et al. (2007) showed that in the absence of finite curvatures (a polygonal 3D shape), perceived gloss is much lower than a smoothly curved shape. Thus, the object geometry should “afford” the presence of highlights.

Similarly, the influence of illumination on glossiness perception was investigated. Various studies using real-world illumination techniques show that some luminance maps¹ result in consistently higher gloss perceptions than other luminance maps (Olkkonen & Brainard, 2010) and that the perceptual differences are transitive (Doerschner, Boyaci, & Maloney, 2010). Since, for glossy objects, highlight shape is one of the most distinctive properties of real-world illumination, we can safely assume that these studies indirectly show that highlight shape influences the perception of gloss. The statistics of the patterns of real-world illumination show a high degree of variability and complexity and, at the same time, they exhibit a great deal of statistical regularity (Dror, Willsky, & Adelson, 2004). These statistics can be linked to physical properties of highlight shapes like the presence of sharp edges. Dror et al. (2004) stated that one might view illumination patterns as complicated textures with clearly recognizable characteristics.

Fleming, Dror, and Adelson (2003) studied the effect of illumination via a matching experiment in which the observer had to change gloss parameters to match an object with a reference object. They showed that matching performance decreases when the illumination has less real-world characteristics. In particular, they

found that real-world illuminations resulted in higher gloss matching accuracy and reliability than point source highlights, homogeneous Gaussian, and 1/f Gaussian noise illuminations maps. Interestingly, they also found that an extended light source (rectangular in their case) performed on par with the natural illuminations. This raises some questions. For example, does the outline shape (e.g., circular compared with rectangular) of the extended light source influence gloss perception? Furthermore, does the structure (e.g., homogenous or having certain variations) within this extended area influence gloss perception?

In visual arts, like paintings, studio photography, and comic books, various extended highlight shapes are quite common (see Figure 1). Here square, disk, ring, or window shaped highlights are frequently used. Much in line with Fleming et al. (2003), artists manage to create very glossy looking objects using extended highlights. Why do artists use these type of highlights and what determines their choice for a specific shape? The simplicity of producing them likely plays a role: Both for studio photography and drawing or painting, it takes lesser effort to render a simple shape, or use an actual extended light source, than a complete real environment. Did we, over time, get accustomed to these simple highlight shapes, as conventions, or are they intrinsically effective to resolve basic image ambiguities (Mamassian, 2008)?

In this study we varied the outline shape of the extended light source and manipulated the structure within the outline area. Our starting point is the use of simple shapes in photography (Hunter, Biver, & Fuqua, 2011), painting (Miller, 1998), sketching (Eissen & Steur, 2009), and other types of drawings (Johnston & Thomas, 1995). Based on the most common highlight shapes used in paintings, comics, and photography, we choose two general outline shapes: square (often used in illustration, see Mazur & Danner, 2014) and circular. We found that common variations of these two are a window shape (often used in paintings, see Miller, 1998) and ring (like a ring flash in photography). These two variations are similar in outline as the square and disk, respectively, but have additional inner structure. To further investigate the influence of this inner structure, we also produced two highlights that are markedly more “complex” than the window and disk. First, we created an area consisting of many small disks, covering approximately the same area as the square highlight. This causes many highlights, or spatial variations, within the extended source, that are not present in the initial four stylized highlights. Secondly, we created an “abstract” highlight that consists of a rather random outline and various areas with different intensities, not found in the other more binary highlight shapes. We also question whether the conditions in which these stimuli are shown might have an effect. Our highlight shapes are

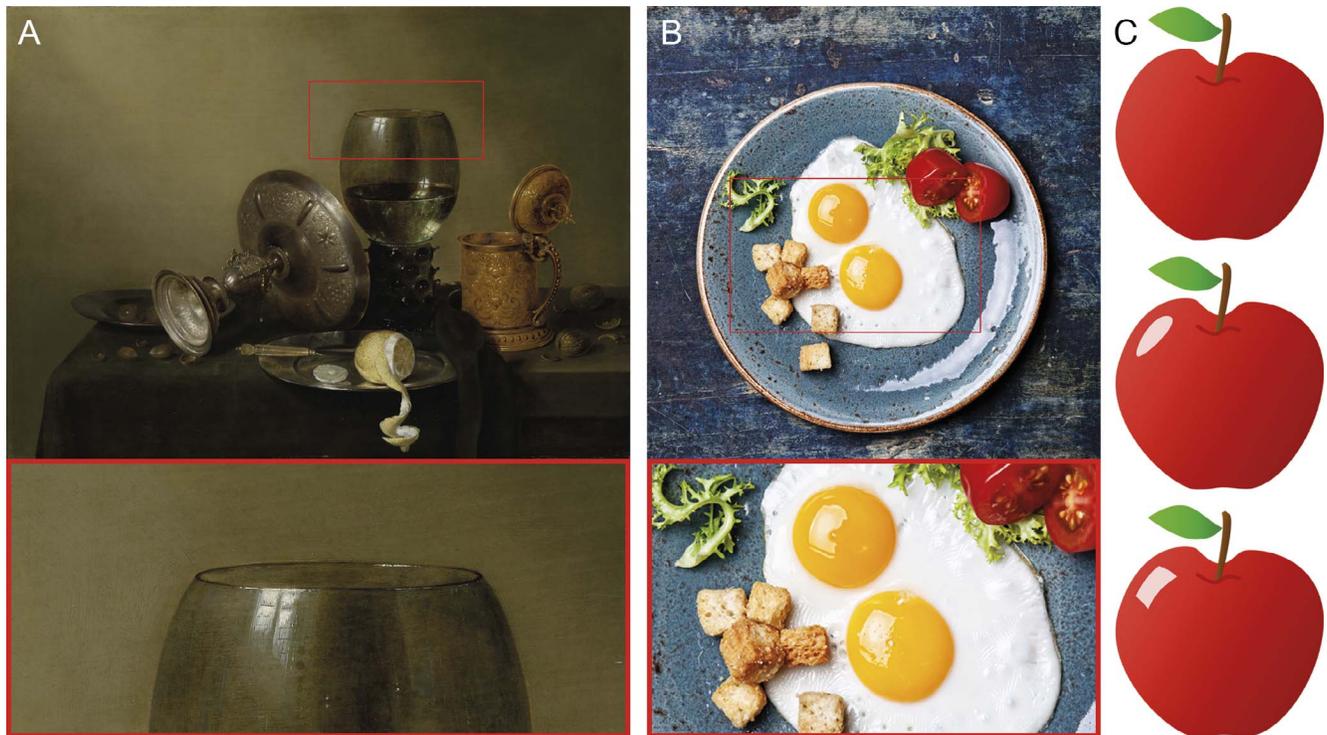


Figure 1. A few examples of highlight shapes. (a) Is an example of a still life painting with a window highlight shape, *Stilleven met vergulde bierkan*, Willem Claesz, Heda, 1634. (b) Is an example of highlight shapes used with studio photography, by Lisovskaya Natalia. (c) Is a simple example of highlights added to cartoon drawings and how the illusion of gloss is created.

stylized and commonly used in two-dimensional art. This raises the question whether they give convincing gloss impressions in two-dimensional representations only, or also in three-dimensional representations of real objects.

Methods

To measure the effect of highlight shape on gloss perception, we tested gloss rating and gloss comparison. We conducted three experiments in order to study the consistency of the effects over presentation conditions and measurement methods. The first experiment (both rating and comparison) took place on a calibrated CRT monitor (Electron 22 Blue III; LaCie, Paris, France) with grayscale photographs of the stimuli. The second experiment (both rating and comparison) took place on the same CRT monitor with color photographs of the stimuli. In the third and last experiment (rating only) the real physical stimuli were shown.

Stimuli

Since the stimuli were made in a real setting using a light box we will precisely describe how the scene was

made, photographed, and presented on a monitor during the experiments.

Apparatus

The basic illumination setup was a 100 cm cube with an upper panel that diffusely lighted the inside of the cube. The light panel consists of white opal glass backlit by fluorescent tubes in such a way that the resulting illumination was homogeneous. In Figure 2 we depict the setup in side views. Spheres were placed on a transparent plexiglass tube fixed on the back wall so the construction holding the sphere was not visible from the viewing hole in the front wall. The sphere was not placed exactly in the middle of the cube but at 45 cm high and at 61 cm viewing distance. This was done to create highlights with a reasonable size and position on the object. The sides and bottom of the box were covered with an opaque dark gray fabric. Below the top surface, which consisted of a lighting panel, we put a frame to attach the masks of the different highlight shapes. The viewing hole was big enough to look through with both eyes and part of the light source on top was visible.

Two experiments were performed with photographs made of objects in the light box presented on a CRT screen in a darkened room and one using real objects in the light box. The onscreen experiments were programmed in MATLAB R2012a with the use of the

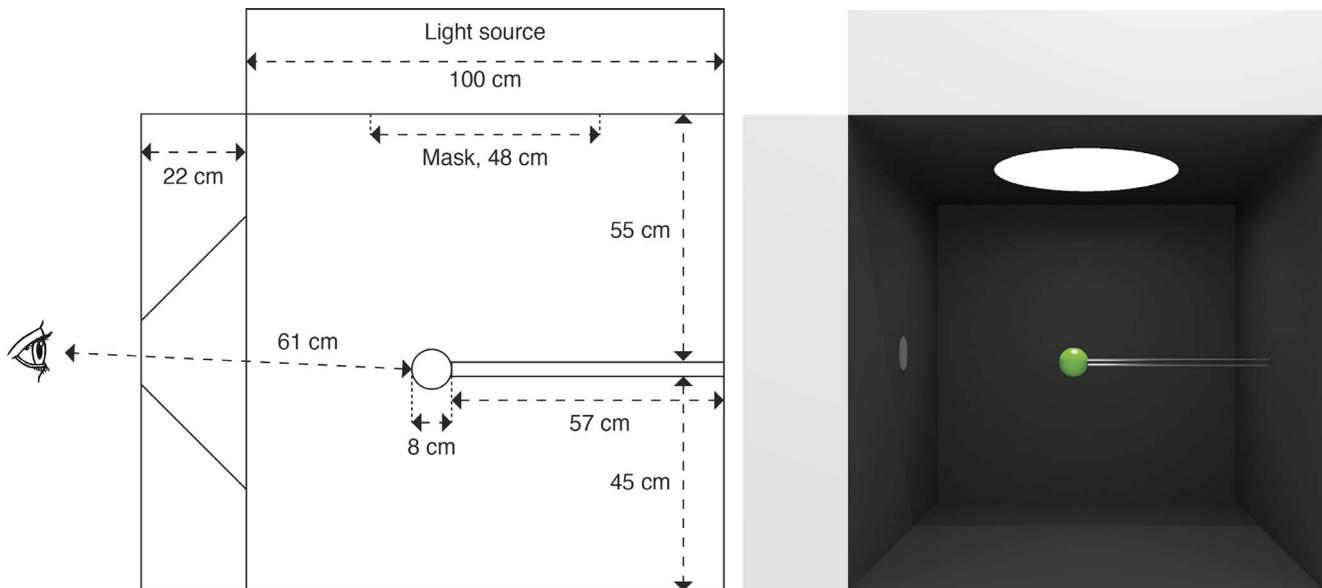


Figure 2. Side view of the light box setup and a rendering giving an impression how it looks in real life.

Psychtoolbox Version 3.0.10 library (Brainard, 1997; Pelli, 1997). The grayscale version of the experiment took place on a CRT screen with a calibrated color profile. A grayscale was measured with a Minolta Luminance Meter LS-100 (Minolta, Osaka, Japan) in the light box and a photograph of that grayscale was projected on the CRT monitor. Using the Mac OS X

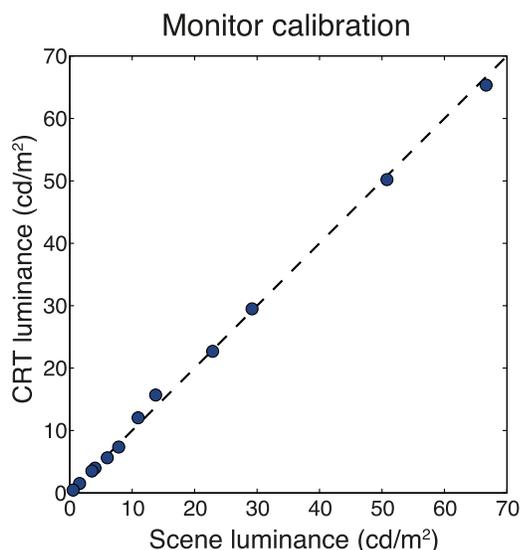


Figure 3. The correlation of the 12 luminance levels of the Fotowand Greystep Card 4962 measured in the light box with the disk mask and on the CRT monitor ($r = 0.99$). Since the observers viewed the stimuli horizontally the greyscale card was oriented vertically in the light box (where the primary lighting came from above, so the horizontal illuminance was much stronger than the vertical) and its luminance values stayed in the dynamic range of the CRT monitor.

Display Calibrator Assistant and the luminance meter the screen was calibrated up to a correlation of 0.99 between measurements from the light box and the CRT monitor. This high correlation was possible because the relatively low luminance levels in the light box were exactly in the reproducible range of the CRT monitor. Figure 3 shows the correlation of luminance levels between the real light box setting and the reproduced stimuli on the monitor. This figure also shows that the luminance levels in the real scene from a black to white surface on a standard grayscale varied between 0.58 and 66.7 cd/m^2 using the disk mask. It is also important to note here that the luminance levels on the very shiny spheres locally reached higher luminance levels. Everything above 65.3 cd/m^2 (monitor limit) is clipped, which applies to the inner core of highlights on the glossiest finishes. For the experiment in color the default factory settings were used, which also enabled us to know whether results of the first experiment can be extrapolated to common conditions, in which we view arbitrary pictures in screens with arbitrary and/or default settings.

Highlight shapes

Six different highlight shapes were used: Disk, Square, Ring, Window, Dots, and Abstract. We created these highlight shapes on spheres in the center of the box by putting masks under the top lighting panel. See Figure 4 for the different highlight masks. The square shape was 48 by 48 cm. The disk shape had a diameter of 48 cm. The ring had a diameter of 48 cm on the outside and 28 cm on the inside, making the ring 10 cm wide. The highlight dots were made in the shape of a 48 cm by 48 cm square where dots of 25 mm in diameter were

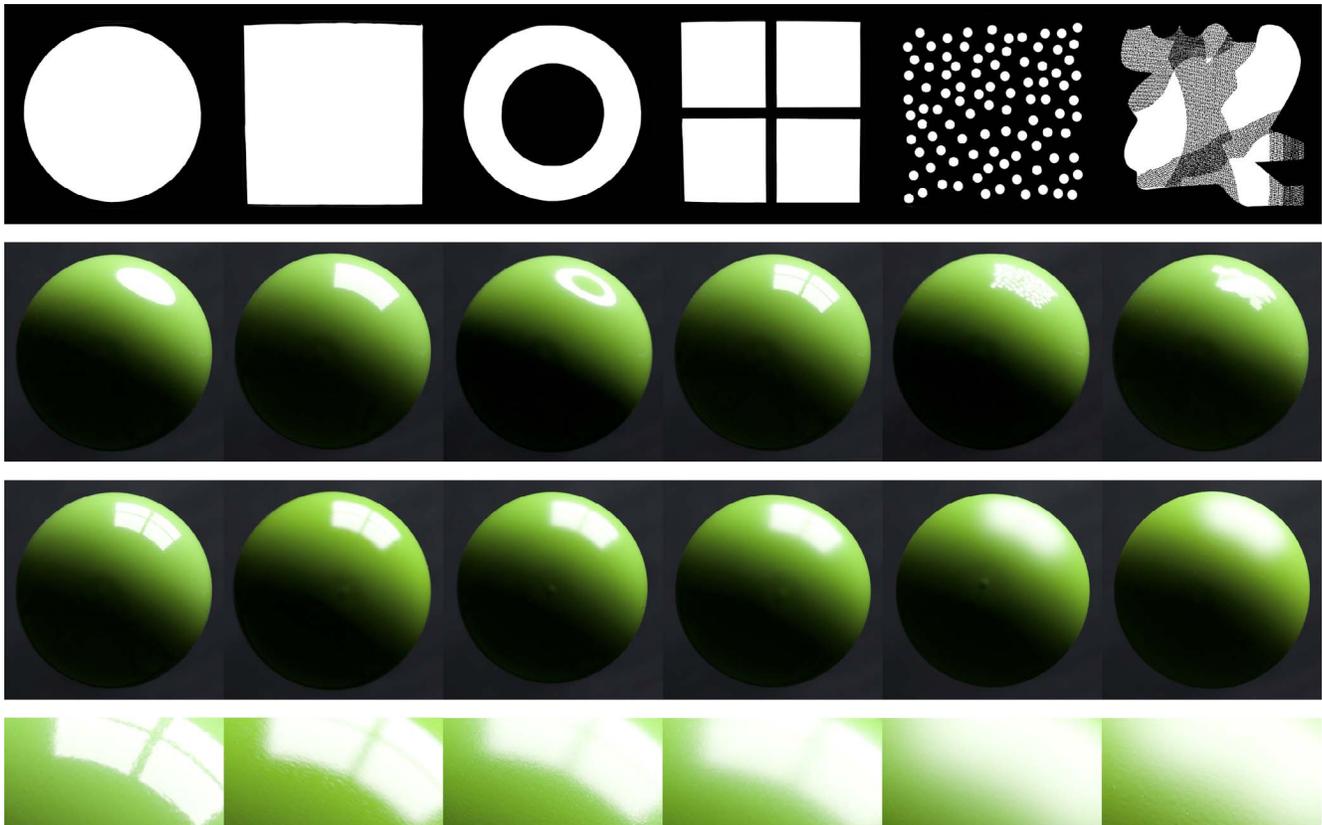


Figure 4. The first row shows the highlight masks: Disk, Square, Ring, Window, Dots, and Abstract. The second row shows the six different shapes on a sphere where the gloss is constant. The last two rows show a constant highlight shape but different gloss coatings and a magnified version. Note: Since these images look like computer renderings, we would like to emphasize that they are photographs of real objects in a controlled environment. Figure 10 is a big version of one of the stimuli showing some of the imperfections that are common in real, but not (yet) in rendered, images.

randomly perforated out of the cardboard. The abstract highlight simulates a more complex shape with different levels of light intensity and a combination of straight and round shapes. The window shape was the same as the square shape only with 34 mm thick muntins crossing through the middle. The relative dimensions of the muntins of the window and of the size of the ring were based on images found using regular web search queries like “ring flash” or “window.” The “complexity” of our stimuli was defined subjectively in an informal experiment and checked according to a ranking of the shapes. There was agreement that the Disk and Square were the least complex of the six shapes; the Ring and Window, medium; and the Abstract and Dots, the most complex of these shapes.

Objects

The highlights were created on six spheres, 80 mm in diameter. The spheres were glass Christmas balls with a black coating inside the sphere preventing light “bleeding through” the objects. We applied four types of green paint on the outer surface in the color RAL 6018. One

type was high gloss alkyd acrylic-based paint (“Belton”, Peter Kwasny GmbH, Gundelsheim, Germany). The other three paint types were high gloss, satin, and matte nitrocellulose-based Selemix (PPG Industries, Pittsburgh, PA). Using layers of different finishes finally resulted in a scale of six different gloss levels. About 15 spheres were spray-painted using combinations of layers with different paint types from which a set of six spheres was selected based on the gloss levels and the absence of visible surface imperfections. The lower row of Figure 4 shows the six different spheres with the window shaped highlight. With six highlights and six spheres the total number of stimuli was 36; see Figure 9 for an overview of all stimuli.

Photography

The pictures were taken with a Canon EOS 5D Mark II (Canon, Inc., Tokyo, Japan) and Canon EF 24-70 mm F2.8 L USM II lens. The pictures were saved in Canon Camera Raw format. All the pictures were taken at ISO 200–70 mm; $f/5.6$ – 0.3 s. We performed a photometric calibration by slightly adjusting the light intensity

(exposure) and color temperature in Adobe Photoshop Lightroom 3.4.1 (Adobe Systems, San Jose, CA) using a Fotowand Greystep Card 4962 (Fotowand Technik, Sudwalde, Germany), which was photographed with each highlight shape. The exposure was calibrated for the different masks to match the DN 0.9 gray step to 50% gray or RGB [128,128,128]. This exposure correction was applied to the photos of the spheres for each mask, normalizing the illuminance over the different masks. The color temperature was set using the Canon firmware and a white balance lens cap. Afterwards we made small corrections using Lightroom in the order of 50 to 100 Kelvin using photographs of a Labsphere (Labsphere, Inc., North Sutton, NH) white reflectance standard. The photos were cropped to 600×600 pixels and rotated 25° to the right for a more generic orientation of the light source in the right upper corner instead of exactly from above. The files were saved in the TIFF format. The conversion from RGB to grayscale took place in MATLAB. MATLAB uses the following conversion weights: $0.2989 \times R + 0.5870 \times G + 0.1140 \times B$.

Procedure

For the first two experiments, which took place on a CRT monitor, observers were placed in a darkened room. First, there was a training session with pictures and paintings of real objects where the observer had to point out the difference between glossy and matte objects. The last two slides in the training session were actual stimulus images, first the maximum and minimum gloss level spheres and then a pair of spheres with only one gloss level step between them. This was done to verify whether the observer fully understood the concept of glossiness and to provide frame of reference for the range of glossiness. During this training session feedback was given about whether the answers were correct or incorrect.

Then the observer started with the paired comparison task. In each trial, two images of spheres were shown next to each other and the observer had to press the left or right arrow key on the keyboard to choose the glossiest sphere of the two. The total amount of possible pairs is $N(N - 1) / 2$ or $36(36 - 1) / 2$, which is 630 pairs for the comparison task. The pairs were presented in a random order and also randomly projected left or right. All the possible pairs were presented once. On average this task took approximately 45 min per observer.

The rating task, always performed after the comparison task, had 144 observations in total (36 stimuli were repeated four times). The observer was presented with a stimulus picture and a rating bar ranging from “very matte” to “very glossy” at the bottom of the screen. With the mouse the observer could place a dot on the rating bar and confirm his or her answer with the space

key on the keyboard. The rating bar had seven ticks, although participants could set the level on a continuous scale, which was translated to a 0–20 score scale. This task took approximately 15 min per observer.

The experiment with color stimuli was performed exactly the same way as the experiment with grayscale stimuli, using the same two tasks as described previously.

For the experiment with the real spheres in the light box only a rating task was performed. The amount of stimuli was limited for this experiment because of the time it took to change the real spheres and highlights. Three highlights instead of six were used, which resulted in 18 stimuli: the disk, abstract, and window highlights with six different gloss levels. The highlight shapes were chosen based on the results of the first two experiments. These stimuli were repeated three times resulting in 54 trials. This experiment was performed in a dimly lit room (luminance was around 0.3 cd/m^2), which was just enough light to fill in the paper form. The scene luminance levels ranged from 0.58 to 70 cd/m^2 . Using the adaptation curves reported by Hecht, Haig, and Chase, (1937) and taking the time into account between each presented scene we can safely assume there was not enough time to fully adapt to mesopic vision. The minute to 70 s between each presented scene was not long enough to prevent rapid recovery to the higher light intensities of the scene.

The third experiment also started with a short training session, which was performed on a laptop to see if the concept of gloss was clear to the observer. The training session was the same as with the previous experiments where scenes with multiple objects were presented on a laptop screen, only the sphere stimuli were now presented in the light box instead of on the monitor. To give the observer a frame of reference for the range of glossiness levels, the first picture of the training session depicted two spheres with the maximum and the minimum gloss finishes. After the training session, the 54 trials were rated. A stimulus was prepared and checked by the experimenter after which the observer could look through the viewing hole with both eyes and mark a cross on a rating bar on the form. This rating bar was the same as the rating bar used in the digital version of the first two experiments. In some cases multiple observers took part in the experiment simultaneously with strict instructions to cover their answer sheets and not talk about what they saw in the light box during the experiment. In total this experiment took around 75 min per (group of) observer(s).

Observers

In each of the three experiments a different group of observers participated. In the first experiment with

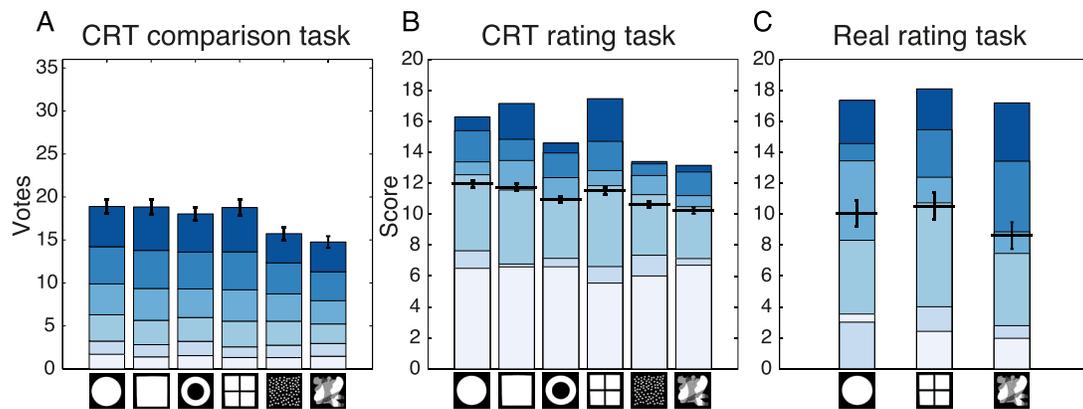


Figure 5. Scores for each gloss finish (dark blue is glossy; light blue is matte) for the different highlight shapes and the three different tasks. (a) Is based on the comparison task performed on the CRT monitor. Note that the scale here is based on the mean amount of votes the highlight shape received with a possible maximum of 35. (b) Shows the mean scores given by participants for each gloss level on a 0–20 scale during the rating task performed on the CRT monitor. The black line represents the mean over all gloss levels. (c) Is based on the rating task performed with real stimuli. The black line represents the mean over all gloss levels.

grayscale stimuli, 11 observers participated of which seven were female and four were male with an average age of 22 years ($SD = 3.1$). In the second experiment performed in color on a CRT monitor, 11 observers participated of which three were female and eight were male with an average age of 20.7 ($SD = 3.3$). Three females and five males with an average age of 23.4 ($SD = 2.0$) participated in the third experiment with real stimuli. All observers had normal or corrected-to-normal vision. All observers gave written consent prior to the experiment and most were novice observers who were paid for participating.

Results

We first analyzed whether using color or grayscale stimuli had a significant effect on the perceived gloss. Second, we analyzed whether the highlight shapes influenced gloss perception differentially. Finally, we analyzed whether the three experiments gave consistent results throughout.

Differences between color and grayscale stimuli

The data of the three experiments were analyzed in three separate two-way repeated-measures analysis of variance (ANOVAs), with Greenhouse–Geisser corrections, where applicable, because of sphericity. However, to reveal possible differences in the perception of gloss between color or grayscale stimuli we added a between subjects factor to the ANOVA for the experiments performed on a monitor. The results show that the effect of color on the pairwise comparison is nonsignificant (p

> 0.05) for both the comparison and rating task. This is perhaps not completely unexpected because of the close to monochromatic nature of the colored stimuli. Since these findings show that there is no significant effect of color we will combine these datasets in the further analysis.

Comparison task

The stimuli of the comparison task were given a score based on how many times they were chosen to be glossier in the pairwise comparison. Since there are 35 possible pairs the maximum score possible to attain was 35 for one stimulus. The results, using a repeated-measures ANOVA, show that perceived glossiness was significantly affected by highlight shape, $F(2.353, 47.064) = 8.598$, $p < 0.0005$, and, as expected, the gloss finish significantly affected perceived glossiness, $F(1.605, 32.108) = 68.159$, $p < 0.0005$. There was also a significant interaction between highlight shape and glossiness, $F(5.529, 110.59) = 7.554$, $p < 0.0005$. Figure 5a shows the scores for each highlight shape. Here, we can see that, overall, the more complex shapes, namely the dots and the abstract highlight, resulted in less glossy percepts than the other shapes. Figure 6 summarizes the pairwise comparisons (Bonferroni corrected) between the highlight shapes that were found to be significant.

Rating task

The repeated-measures ANOVA applied on the rating task data confirms the results of the comparison task in that highlight shapes are influencing perceived glossiness, $F(3.001, 60.028) = 10.753$, $p < 0.0005$, gloss finish influences perceived glossiness, $F(1.638, 32.753) =$

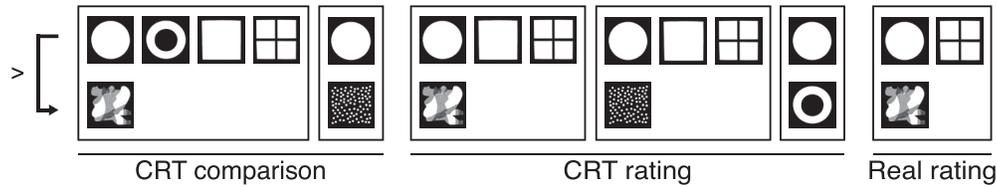


Figure 6. A representation of all significantly different pairwise comparisons found between the highlight shapes for each method, e.g., the disk, ring, square, and window highlight were perceived to be significantly more glossy than the abstract highlight for the comparison task performed on the CRT monitor.

64.067, $p < 0.0005$, and there is a significant interaction between highlight shape and gloss finish $F(6.822, 136.44) = 6.822, p < 0.0005$. Figure 5b shows the scores for each highlight from a 0-to-20 score scale. Figure 6 summarizes the results: the disk, square, and window highlights resulted in significantly increased glossiness perception compared with the abstract and dots' highlight shapes; for the rating task we found one more significant relation between the highlight shapes, namely that the disk shape resulted in increased glossiness perception compared with the ring.

Rating task in light box

The results for the experiment in the real setting showed that there was a significant relation between highlight shape and perceived glossiness, $F(2, 14) = 18.844, p < 0.0005$. The same applies for the relation between gloss finish and perceived glossiness, $F(1.558, 10.907) = 119.026, p < 0.0005$, and the interaction between highlight shape and gloss finish, $F(10, 70) = 5.834, p < 0.0005$. The results of the experiment with real stimuli in the light box confirm that the disk and window highlights were again perceived to be significantly glossier than the abstract highlight shape; see

Figure 5c and Figure 6. Note that only these three shapes were tested in this experiment.

Rating task compared with the comparison task

The results from the ANOVAs showed that there were statistical differences between the comparison and rating task performed on a CRT monitor. Figure 7a shows the relation between the comparison and rating data with the mean of all highlight shapes. The results from the comparison task were rescaled to a 0–20 scale. The general structure for the two different tasks is very similar, indicating that a rating task could serve as a good alternative for testing psychophysical stimuli compared with the more proven pairwise comparison.

Rating task performed on a CRT monitor or in the light box

Figure 7b shows the results of the rating task performed on a CRT monitor (both colored and grayscale stimuli) and the results of the rating task performed with real stimuli in the light box. The main difference between the two graphs is that for real

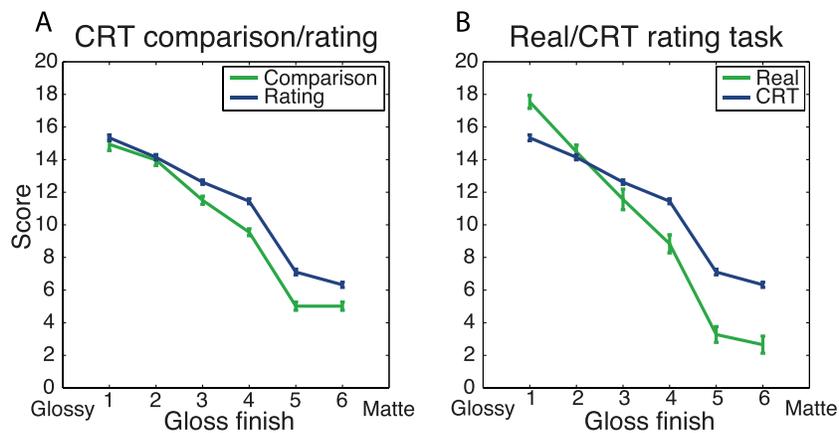


Figure 7. Showing the scores as a function of gloss finishes for the different experiments. (a) Shows the difference between the comparison task performed on a CRT monitor and the rating task on a CRT monitor. Note here that the comparison task data was rescaled from 0–35 to 0–20 to fit the rating task data. (b) Shows the gloss scores for the rating task experiments performed on a CRT monitor and with real stimuli in the light box.

stimuli a wider range of the score scale was utilized by the observers than for the stimuli on the CRT screen. The general pattern of the data as a function of gloss level is consistent in all experiments: a gradual decrease of perceived gloss between level 1 and 4, then a steep decrease from 4 to 5, followed by a less steep decrease between 5 and 6.

Conclusions and discussion

We tested the influence of highlight shape on gloss perception. Six highlights were used varying from simple shapes, like a disk or square, to more complex shapes, like our dots or abstract shape. These six highlights were combined with six spheres having different gloss levels, resulting in 36 different stimuli (see Figure 9). The results show that highlight shape does significantly influence the perception of glossiness. This was confirmed via a 2AFC task performed on a CRT monitor, a rating task performed on a CRT monitor, and a rating task using real stimuli in a light box. The different types of experiments show the same qualitative trends in the results, meaning that photographed stimuli presented on a CRT monitor performing a rating task or using real stimuli with a rating task resulted in the same general patterns and relative differences. Since the step sizes of the graphs in Figure 7 seem to rescale in a consistent manner for all experiments, they probably reflect physical differences between the glossiness of the spheres. However, observers did use a bigger range of the rating scale if they judged the glossiness of the real objects than if they judged the stimuli on a computer screen.

What might explain the last result is that the stimuli on the computer were presented faster after each other (1 s delay) than in the lighting box; preparing a stimulus in the light box took around a minute. Another difference is that on the computer the observer was presented with one rating bar at a time whereas with the light box experiment the observer could oversee a maximum of 18 answers on the answering sheet. These two differences might have influenced the frame of reference of the observer. Using real stimuli instead of images on a computer screen also provided the observers with a higher dynamic range, stereo cues, and full resolution. This extra information might have made them more confident to use the extremes on the rating scale or indeed might have made the stimuli look less glossy for the lowest level and more glossy for the highest levels, compared with the on-screen stimuli. Other studies have also shown that high dynamic range and stereo cues can make glossy stimuli look more glossy (Lichtenauer, Schuetz, & Zolliker, 2013; Obein,

Knoblauch, & Viénot, 2004; Philips, Ferwerda, & Luka, 2009; Sakano & Ando, 2010).

According to Ferwerda, Pellacini, and Greenberg (2001) and Anderson and Kim (2009) the two main image features influencing the perception of gloss are the contrast of the reflected image and the distinctiveness of the reflected image. Our stimuli were created in a controlled environment—a diffuse light box covered with dark gray fabric except for the highlight mask. The highlighted objects were smooth glass objects with different paint finishes. Thus, the differences between our stimuli were mainly influenced by the distinctiveness of the reflected image, while the contrast was rather constant. This can explain the results as a function of gloss level, but not the differences between the highlight shapes. Marlow et al. (2012) have demonstrated that the perception (and misperception) of gloss is well predicted by the way that each illumination field modulates the size, contrast, sharpness, and depth of specular reflections. The contrast and depth of the highlights was rather constant for our stimuli. Sharpness indicates a similar image feature as distinctness of image. We kept the highlight outer sizes similar for our stimuli, but we cannot exclude that the coverage might partly explain the results that the complex highlights resulted in less glossy appearances.

Ferwerda et al. (2001) argue that other visual qualities like sheen and haze might be needed to describe certain qualities of real glossy and or shiny objects. These qualities were also included in the framework suggested by Hunter and Harold (1987) in which there are at least six visual properties related to apparent gloss: specular gloss, contrast gloss, distinctness-of-image gloss (DOI), haze, sheen, and absence-of-texture gloss. The first three properties were basically discussed in the former paragraph, and haze and sheen do not apply to our stimuli. Since our paint coatings showed some imperfections (see Figure 10), we cannot exclude that absence-of-texture gloss might have influenced our results somewhat. However, we carefully selected our stimuli and rejected objects that showed too many imperfections and therefore believe that this did not influence our main conclusions.

As mentioned in the Introduction, Fleming et al. (2003) found that real-world illumination resulted in equally reliable and accurate gloss perception as an extended rectangular light source, while other artificial highlights caused less reliable and accurate gloss perception. We wanted to investigate this further by testing variations of the extended highlight light source that are often used in artistic practice such as photography, painting, and illustration. Note that we were particularly interested in what type of illumination would result in the highest gloss, and did not test the “reliability” or “accuracy” as Fleming et al. (2003) did. Furthermore, we wanted to know whether integrating

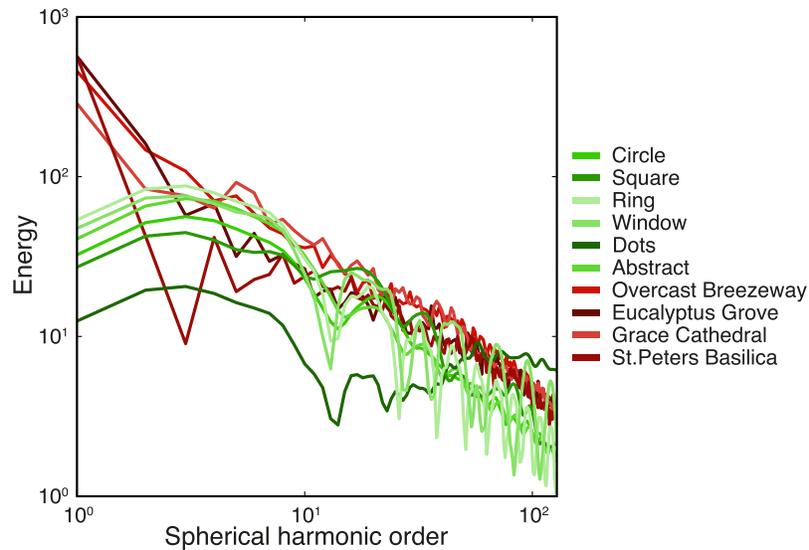


Figure 8. Results of the spherical harmonic decomposition performed on our six different luminance maps and four real-world environment maps. The real-world environment maps have red colors and our luminance maps have green colors. It is clearly visible that the low order's energy is lower in our luminance maps than in the real-world environment maps, because of the restricted directionality and binary character of our stylized extended sources. The higher order's energy however shows similar patterns for all but the Dots illumination.

certain real-world characteristics within this extended area would influence gloss perception. First, we found that there are hardly any differences among the four basic highlights (Disk, Square, Ring, Window). The only difference we found was the Disk being perceived more glossy than the Ring in the CRT rating task. Here, the outline of the light source is similar, but the inner structure differs. Secondly, we found rather robustly that the two complex highlights resulted in lower gloss estimation than the four basic highlights. Although we did not use actual real-world illumination, we intended to integrate some aspects of real-world illumination in these two complex artificial light sources. The motivation behind this was to explore whether an extended light source that consists of real-world illumination elements would cause higher perceived gloss than merely a simple extended light source.

Although we designed our light sources with good intentions, it is still difficult to assess how close they resemble actual real-world illumination. To get a rough impression of how close our artificial “real-world” light sources resembled actual real-world illumination we performed a spherical harmonic decomposition of the luminance maps, a technique often applied in studies into the luminous environment (Doerschner, Boyaci, & Maloney, 2007; Dror et al., 2004; Mury, Pont, & Koenderink, 2007). We included all six illuminations and four real-world environment maps made by Paul Debevec (1998). A spherical harmonic decomposition can be seen as the spherical version of a Fourier decomposition. As shown in Figure 8, our six illuminations contain less low-order energy in com-

parison with the real-world environment maps. The restricted directionality and binary character of our stylized extended sources are the cause for this decreased low order energy. The spectral energy of the six extended sources and the real-world environment maps is rather similar beyond the second order component, with the Dots illuminant as an outlier. Based on this spherical harmonics analysis, we cannot explain our results. As Fleming et al. (2003) stated: “Mimicking the power spectrum of real-world illumination is insufficient to create a compelling impression of gloss.” However, they also state that “By contrast, extended edges and a predominant direction of illumination tend to lead to good impressions of gloss.” The lack of such clearly defined “extended edges” in our abstract and dots highlights might explain our results.

One commonality between our simple extended light sources and natural illumination is that they both contain a clearly recognizable shape that is deformed. Perhaps, the visual system assumes that the environment contains somewhat regular objects, and is able to identify their deformation as a highlight indicating gloss. The two complex highlights both have less clearly defined outlines (or “extended edges”). It should be noted that a deformed luminance map does not affect the perception of 3D specular shape (Fleming, 2004). However, this does not necessarily mean that the perception of surface quality is similarly permissive with respect to deformations. Fleming et al. (2003) noted, “Although higher-order regularities found in the environment are likely to facilitate realism, they are not

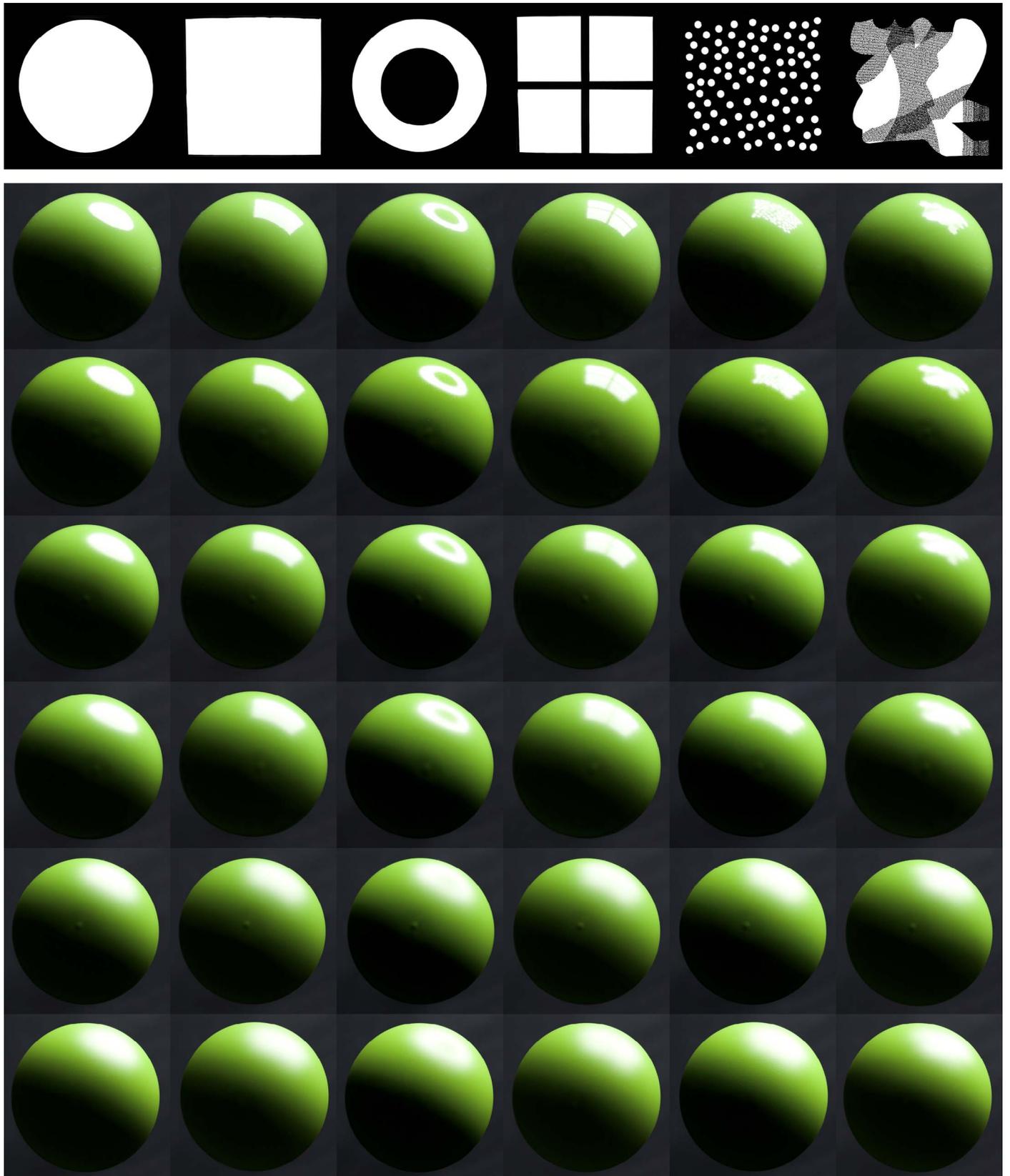


Figure 9. The stimuli: from left to right the six highlight shapes and from top to bottom the six gloss levels.

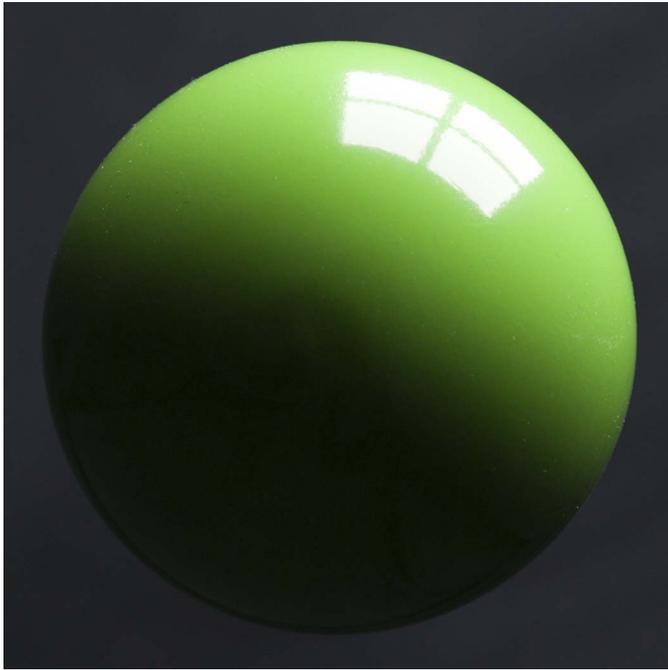


Figure 10. A stimulus printed at a bigger size. Note that on this scale it is visible that this is a photograph and not a 3D rendering from the imperfections on the surface of the sphere.

required for compelling impressions of surface reflectance.” However, deforming the whole environment may have a different effect than deforming the only visible shape in the environment. Using a simple regular shape as extended light source and deforming it according to the geometry of the object is indeed common artistic practice. We suggested in the Introduction that this convention could originate from either practical reasons (easy to render) or based on a heuristic based on the visual system. Our empirical findings and their robustness for viewing conditions suggest that the latter explanation is likely part of the reason, although we cannot rule out the contribution of practical advantage.

Much theory about vision is based on regularities in the natural environment, and in natural images. Indeed, it is biologically very plausible that our visual system is based on these natural statistics. However, we are not only surrounded by other animals, trees, buildings, etc.; we are also surrounded by depictions. Although these depictions find their origin in our natural environment, they do not always comply to the same rules. Therefore, we should not rule out the possibility that our visual system is also tuned to interpret depictions that are markedly dissimilar to the natural environment. Although very tentative, we should also consider that our finding is partly based on the visual conventions we see in the depictions around us. In other words, the artists may base their conventions (Johnston & Thomas, 1995; Mamassian,

2008; Mazur & Danner, 2014; Miller, 1998, Phillips, Mazarrella, & Docter, 2014) on elements of the visual system; their conventions may, in turn, serve as input to our perceptual learning. A regularly shaped highlight may not only be effective because we can interpret its deformation easily and attribute it to gloss, but also because we have learned to interpret it as gloss.

Although our current empirical findings cannot be used to support this hypothesis, it is in our belief certainly an interesting direction for future research.

Keywords: gloss, highlight, shape, illumination, natural scenes

Acknowledgments

This work has been funded by the EU FP7 Marie Curie Initial Training Networks (ITN) project PRISM, Perceptual Representation of Illumination, Shape and Material (PITN-GA-2012-316746). We thank our reviewers and the editor, David Brainard, for the elaborated critical discussions, which significantly improved the scientific argumentation of this paper.

Commercial relationships: none.

Corresponding author: Jan Jaap R. van Assen.

Email: mail@janjaap.info.

Address: Department of Psychology, Justus-Liebig-University Giessen, Giessen, Germany.

Footnote

¹ We use the term luminance map instead of light field. A light field is the radiance distribution throughout the empty space of a three-dimensional (3D) scene. Thus, it is dependent on position and direction. Luminance is radiance, weighted spectrally according to a “normal observer.” A luminance map is a map of a local measurement (in a certain position) of the luminance. Thus, it is only dependent on direction.

References

- Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision*, 9(11):10, 1–17, doi:10.1167/9.11.10. [PubMed] [Article]
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Debevec, P. E. (1998). Rendering synthetic objects into

- real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proceedings of ACM SIG-GRAPH, 1998*, 189–198.
- Doerschner, K., Boyaci, H., & Maloney, L. T. (2007). Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. *Vision Research, 47*(28), 3409–3423.
- Doerschner, K., Boyaci, H., & Maloney, L. T. (2010). Estimating the glossiness transfer function induced by illumination change and testing its transitivity. *Journal of Vision, 10*(4):8, 1–9, doi:10.1167/10.4.8. [PubMed] [Article]
- Dror, R. O., Willsky, A. S., & Adelson, E. H. (2004). Statistical characterisation of real-world illumination. *Journal of Vision, 4*(9):11, 821–837, doi:10.1167/4.9.11. [PubMed] [Article]
- Eissen, K., & Steur, R. (2009). *Sketching: Drawing Techniques for Product Designers* (12th ed.). Amsterdam, The Netherlands: BIS Publishers.
- Ferwerda, J. A., Pellacini, F., & Greenberg, D. P. (2001, June). Psychophysically based model of surface gloss perception. In *Photonics west 2001-electronic imaging* (pp. 291–301). Bellingham, WA: International Society for Optics and Photonics.
- Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5):3, 347–368, doi:10.1167/3.5.3. [PubMed] [Article]
- Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision, 4*(9):10, 798–820, doi:10.1167/4.9.10. [PubMed] [Article]
- Hecht, S., Haig, C., & Chase, A. M. (1937). The influence of light adaptation on subsequent dark adaptation of the eye. *Journal of General Physiology, 20*, 831–850.
- Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Biological Science, 19*(2), 196–204.
- Hunter, F., Biver, S., and Fuqua, P. (2011). *Light science and magic: An introduction to photographic lighting* (4th ed.). New York: Focal Press.
- Hunter, R. S., & Harold, R. W. (1987). *The measurement of appearance* (2nd ed.). New York: Wiley.
- Johnston, O., & Thomas, F. (1995). *The illusion of life: Disney animation*. New York: Disney Editions.
- Kartashova, T., Heynderickx, I., Sekulovski, D., and Pont, S. C. (2014). Relating physical and visual global light structures. In *Proceedings of experiencing light 2014*. (pp. 118). Presented at International Conference on the Effects of Light on Wellbeing, November 10–11, 2014, Eindhoven, The Netherlands.
- Kartashova, T., te Pas, S. F., Pont, S. C., de Ridder, H., & Schoemaker, M. (2015). The visual light field in paintings of museum Prinsenhof: Comparing settings in empty space and on objects. In B. E. Rogowitz, T. N. Pappas, & H. Ridder (Eds.), *Human vision and electronic imaging XX Vol. 9394*, Proceedings of SPIE International Society for Optical Engineering (pp. 93941M-1-93941M-10). Washington: SPIE. (TUD).
- Kim, J., Marlow, P. J., & Anderson, B. L. (2012). The dark side of gloss. *Nature Neuroscience, 15*, 1590–1595, doi:10.1038/nn.3221.
- Koenderink, J. J. (2001). Multiple visual worlds. *Perception, 30*, 1–7.
- Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta, 27*, 981–996.
- Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shad(ow)ed random Gaussian surfaces. *Perception, 33*, 1405–1420.
- Lichtenauer, M. S., Schuetz, P., & Zolliker, P. (2013). Interaction improves perception of gloss. *Journal of Vision, 13*(14):14, 1–13, doi:10.1167/13.14.14. [PubMed] [Article]
- Mamassian, P. (2008). Ambiguities and conventions in the perception of visual art. *Vision Research, 48*, 2143–2153.
- Marlow, P. J., Kim, J., & Anderson, B. L. (2012). The perception and misperception of specular surface reflectance. *Current Biology, 22*, 1909–1913.
- Mazur, D., & Danner, A. (2014). *Comics: A global history, 1968 to the present*. London: Thames & Hudson.
- Miller, J. (1998). *On refecton*. London: National Gallery Publications Limited.
- Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447*, 206–209.
- Mury, A. A., Pont, S. C., & Koenderink, J. J. (2007). Light field constancy within natural scenes. *Applied Optics, 46*, 7308–7316.
- Obein, G., Knoblauch, K., & Viénot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision, 4*(9):4, 711–720, doi:10.1167/4.9.4. [PubMed] [Article]
- Olkkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision, 10*(9):5, 1–19, doi:10.1167/10.9.5. [PubMed] [Article]

- Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception, 34*, 1301–1314.
- Pelli, D. G. (1997). The VideoToolbox for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.
- Philips, J. B., Ferwerda, J. A., & Luka, S. (2009, January). Effects of image dynamic range on apparent surface gloss. In *Color and imaging conference, Vol. 2009, No. 1* (pp. 193–197). Bellingham, WA: Society for Imaging Science and Technology.
- Phillips, F., Mazzarella, J., & Docter, P. (2014). Specularity and shape from line drawings. *Journal of Vision, 14*(10):729, doi:10.1167/14.10.729. [Abstract]
- Pont, S. C., & Koenderink, J. J. (2007). Matching illumination of solid objects. *Perception and Psychophysics, 69*, 459–468.
- Pont, S. C., & Koenderink, J. J. (2008). Shape, surface roughness, and human perception. In M. Mirmehdi, X. Xie, & J. Suri (Eds.), *Handbook of texture analysis* (pp. 197–222). London, UK: Imperial College Press.
- Pont, S. C., van Doorn, A. J., Wijntjes, M. W. A., & Koenderink, J. J. (2015). Texture, illumination, and material perception. *Proceedings SPIE 9394, Human Vision and Electronic Imaging, XX*, 93940E1–10, doi:10.1117/12.2085953.
- Sakano, Y., & Ando, H. (2010). Effects of head motion and stereo viewing on perceived glossiness. *Journal of Vision, 10*(9):15, 1–14, doi:10.1167/10.9.15. [PubMed] [Article]
- Sharan, L., Rosenholtz, R., & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision, 14*(9):12, 1–24, doi:10.1167/14.9.12. [PubMed] [Article]
- te Pas, S. F., & Pont, S. C. (2005, August). A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In *Proceedings of the 2nd symposium on applied perception in graphics and visualization*. (pp. 75–81). New York: ACM.
- Vangorp, P., Laurijssen, J., & Dutré, P. (2007). The influence of shape on the perception of material reflectance. *ACM Transactions on Graphics 26, 3 (Proceedings of ACM SIGGRAPH 2007)*, 77, 1–77:9.
- Wijntjes, M. W. A., & Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *Journal of Vision, 10*(9):13, 1–12, doi:10.1167/10.9.13. [PubMed] [Article]
- Wijntjes, M. W. A., Volcic, R., Pont, S. C., Koenderink, J. J., & Kappers, A. M. L. (2009). Haptic perception disambiguates visual perception of 3D shape. *Experimental Brain Research, 193*, 639–644.