# The Effects of the Physical Characteristics of Light Probes on Image Light Perception

Sònia Fanlo Garcia

Student number 0930415

Faculty of Social and Behavioural Sciences, Utrecht University

Master's Thesis Applied Cognitive Psychology

Main supervisor Dr. Susan te Pas

Second supervisor Christoph Strauch

May 7th, 2021

Universiteit Utrecht

## ABSTRACT

The physical light field becomes apparent when it interacts with objects within a space. Human observers are able to primarily infer three light properties relevant in lighting design (1) Overall diffuse light, (2) Directed light and (3) Brilliance. Previous studies mainly used Lambertian light probes to gauge light in space by taking cues from the probes' lighting characteristics. However, Lambertian probes fail to capture Direction and Brilliance, which allows for glossiness and atmosphere perception. We tested whether differences in material, shape and surface structure of light probes influenced the ability of observers to gather information and interpret three light properties in natural scenes. Three different light probes were used: A Lambertian sphere, a black shiny sphere and a golf ball. Participants were shown black and white images of one of the probes photographed in a scene. Some images had the probe originally photographed in the scene, but in other images the probe was swapped for that of a different scene, creating different combinations of probe and scene. Participants had to determine whether the probe "fitted" the scene. This project aims to consolidate light probes as a design tool and help lighting professionals correctly represent a space's illumination using the light probe that best captures the light properties at play in a scene. Overall, black shiny probes have proven to be better at helping determine light than Lambertian spheres. Our hypotheses that one probe type is better than others in helping infer certain light properties in scenes with similar features cannot be proven, but significant results at the image level suggest a conditional relationship of this matter might exist.


## INTRODUCTION

The elements and principles of design are the framework upon which good design is created, evaluated and communicated (Nielson & Taylor, 2007). Although there are inconsistencies about which concepts compose those elements, a survey by Boucharenc (2006) stated that point, line, space, proportion, light, color and rhythm are the concepts most used in design education across twenty-two different countries. In this paper, we are going to be looking at light in natural scenes. As Kartashova et al (2019) said: In order to consciously manage the looks of a scene, one should keep in mind the intricate model of all light interactions. Gershun first talked about light and its importance in design as a result of the growth in illuminating engineering. There is a physical light field that becomes apparent as it interacts with the objects within a space (Gershun, 1936; Schirillo, 2013): It helps accentuate surfaces, materials and structural features. As such, understanding how light is perceived in a space is key to the general construct and success of a design.

Different studies tried to measure and quantify both the physical and perceptual light field and their effects in natural scenes. Koenderink et al (2003) measured the direction of irradiation from texture while Mury et at (2007, 2009) focused on extending the physical light field's understanding of higher order components in the layout of natural scenes. Kartashova et al (2016, 2018) made a reconstruction of the spatial structure of the perceived light field and developed a toolbox to analyze the spatial distribution of light and its properties' variation though different scenes.
Koenderink et al (2007) studied the differences between the physical and visual light field and found that human observers have expectations about the way an object would appear in a space, thus perceived in the visual light field, but Kartashova et al (2016) later found that those expectations are simplified versions of the (physical) light field.

Human observers are sensitive to the three low-order light properties -intensity, direction and diffuseness- (Cuttle, 2003). Because light is intangible and only visible when shown in objects' surfaces, knowledge about light properties and light transmission on different surfaces was needed.
Light probes are objects used as a tool to help infer light in a space (Koenderink et al., 2007). Most studies about human perception of light properties consisted of experiments run in an artificial setting where light probes are typically used as a tool to help gauge the light in a scene by inferring the lighting seen in the probe (Kartashova et al., 2016), (Kartashova et al., 2018), (Koenderink et al., 2007). Different shaped objects, such as penguin statues (Koenderink et al., 2007), spinners or bowling pins (Kartashova et al., 2018) have been used as light probes.

White matte spheres, also called Lambertian probes, are the type of light probe most commonly used for light perception studies. However, Lambertian light probes might be problematic due to their matte condition, as they fail to correctly capture the higher-order properties of the light, that allow for glossiness and atmosphere perception (Mury et al., 2007; Kartashova et al., 2019). The question inevitably arises as to whether a different type of probe -textured or non-matte- would be more effective in providing the necessary information to gauge and interpret the illumination in a space. Similar work was done by Xia et al. (2014), in comparing a smooth and a rough probe optically placed in a natural scene using a mirror, and concluded that the rough probe was better at identifying mismatches between direction and diffuseness between scene and probe.

In this project we investigate if a variation in the shape, material and surface structure of a light probe affects how well humans perceive light in a natural scene -not a laboratory-.
Three different light probes will be tested: a white matte probe (Lambertian), a black shiny probe, and a golf (styrofoam) ball. Our goal is to research what light properties in a scene are easier to be identified with a specific type of probe.
The hypothesis for this project are the following:

- *Hypothesis 1:* White matte probes (Lambertian) are best at capturing ambient light (Pharr et al., 2016).
- *Hypothesis 2:* Golf balls are best at capturing directed light (Koenderink & Pont, 2003; Xia et al., 2014).
- *Hypothesis 3:* Black shiny probes are best at capturing brilliance (Kartashova et al., 2016; Kartashova et al., 2018).

Knowing what light properties are at play in a particular scene conveys important implications for designers and lighting professionals, as it allows them to better understand the role certain lighting plays in a space (Kartashova et al., 2019). Does it highlight or hide, how do reflective materials, ceiling lights and geometrical shapes translate in the space? And most importantly, how do observers perceive them?
Our goal is to consolidate light probes as a design tool and to provide a guideline for lighting professionals on what type of light probe is best suited for presenting their projects based on the desired illumination and specific features of a space.

# METHODS

This research was approved by the Ethical Review Board of the Faculty of Social and Behavioural Sciences of Utrecht University, number 21-0614.

## Participants

We gathered data from 20 participants with normal or corrected-to normal vision. Participants' age ranged between 18 and 26 years old. All participants were naïve as to the goal of the experiment and were given little instructions[1]. None of the participants had an art or design background.
Recruitment was made through SONA Systems (25) and Facebook groups of Utrecht University (6). All participants were students of Utrecht University and received either Financial compensation (8€ per hour) or participation credits (1PPU) for the experiment.
The total number of participants recruited was 31. Of those, 4 did not complete the experiment and 7 completed the task in under 9 minutes. That means they spent less than 3 seconds per image and thus results cannot be trusted to be anything but answered at chance level. Those 7 participants were also not included in the data.
Because of participants cruising though the experiment at high speed, image screen time was modified mid-recruiting so that images stayed in the screen for 3 seconds before response buttons popped up. 8 of the final 20 participants performed the experiment with timed screens.

## Stimuli

All images were provided by Sylvia C.Pont from the Faculty of Industrial Design Engineering at TUDelft[2].
Three types of light probes were photographed in different scenes for the experiment: a white matte Lambertian sphere, a black shiny sphere, and a styrofoam "golf ball".

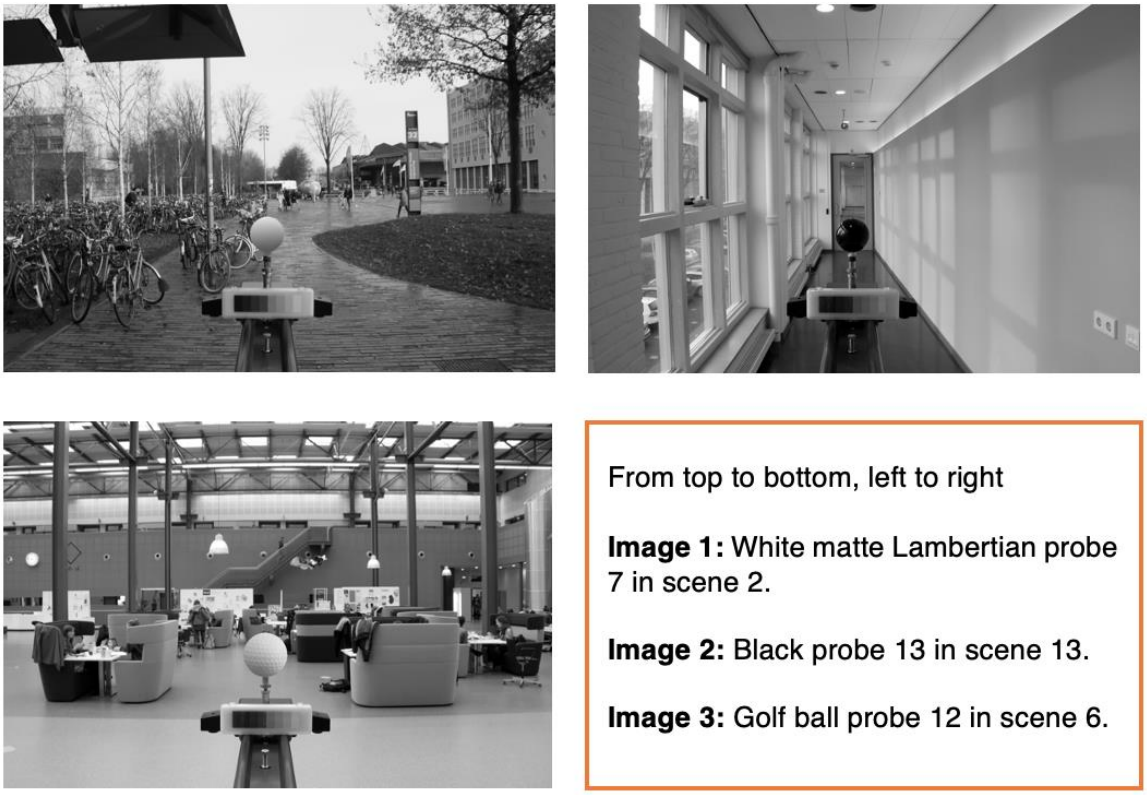



**Figure 1.** Example of images used in the experiment of the three probe types tested in different scenes.

[1] See Appendix 1 for Instructions of the experiment
[2] See Appendix 2 for all Stimuli arranged by scene

Each probe was photographed in 13 scenes, so there are 13x3= 39 original images. The other images were a manipulation of the original images where the probe originally photographed in a scene was artificially set in another scene. For example, an image with the golf ball probe of scene number 12 was placed in scene number 6. A total of 156 images were shown in the experiment. 39 were the original photographs and 117 were combinations of a probe placed in a different scene (13 scenes x 3 probes x 3 combinations of each probe). All images were named and will be referred to by: ProbeType_NumberScene-NumberProbe[3].

## Setup

We used the Gorilla Experiment Builder (www.gorilla.sc) to create and host a one-session online experiment. Data was gathered between February 12th 2021 and March 7th 2021.

Participants signed up to the online experiment using the Recruitment platform SONA Systems. They could take part in the experiment whenever they wanted as long as it was at least 24h after signup. Participants had to log in to their account in the recruitment website, and click the experiment post in the platform where they were provided with a direct link to the experiment.
For participants recruited in Facebook groups, they directly contacted the main researcher and were given access through an email shot policy.

Participants took part in the experiment using their own computer or laptop. Mobile devices and tablets were not allowed to avoid problems with image resolution.
The experiment was built with a feature to automatically display full-screen and all participants needed to adjust their screen resolution size to that of a credit card by adjusting it to a white-block figure. This ensured standard ratio size (width to height) of 177/100 for all images shown across all participants.

The experiment began with a practice test round consisting of 5 images. The images were selected with the criteria to be easily identifiable either because they were original images (3) or because they were clearly manipulated and considered easy to identify as not fitting (2). However, participants didn't receive any feedback about their practice test answers.
After that, the experiment began and participants were shown 156 images in randomized order. There were two 1 minute breaks in between the experiment.

All images in the experiment appeared on the screen for 3 seconds before two buttons popped up below: FITS & DOES NOT FIT. This was to ensure a minimum time was spent looking at the images and not randomly passing through. The experiment had a time limit of 3:30h to be completed, although all participants spent between 11 and 90 minutes. It is worth noting the participant who spent 90 minutes (p.103) is an outlier and all other participants were under the 38 minute mark.

## Procedure

### Coding[4]

All images were first coded on difficulty by the main researcher and their supervisor according to visual inspection to make predictions about image performance.
A four-step difficulty scale was created and assigned each step numbers from 1 to 4. Number 0 corresponds to an image being the original photograph, meaning the probe corresponds to the scene.

(0) Original, (1) Easy to identify, (2) Moderate, (3) Difficult and (4) Chance level.

The scale was based on the difference comparison of light properties between probe and scene. Images where the probe had one light property extremely different than the scene were coded as 1, and images with multiple subtle differences in light properties for probe and scene were sometimes coded as 4.

[3] Note that there is no scene or probe number 11, but there are 13 scenes and probes so the last ones are coded as scene and probe 14.
[4] See Appendix 3 for Coding of all images.

Therefor the amount of light properties not matching between sphere and scene was not the driver for the difficulty scale, but rather the perceived difference strength between them. Both researchers separately coded all the images to ensure higher reliability. Afterwards, coding for all images were compared and when more than 1 point difference between criteria existed, the image was discussed between the two coders to reach a more harmonic conclusion. The final coding is the average of the two researchers code, which explains the half points attributed to some images.

Furthermore, images were also briefly described in three categories according to three attributes or cues of light (Kartashova et al., 2019), each enhanced by one of the probes.
Those categories are *Diffuseness, Direction* and *Brilliance*, and correspond to white probes, golf probes and black probes respectively.
The category with the strongest effect, or the most noticeable cue to determine if a probe fitted the scene was written down first. Not noticeable differences in categories were left out of the classification.

## Grouping[5]

Afterwards, all images were grouped by one of the researchers in four categories based on the location and lay-out of the scene.
Groups were made as a means to classify scenes based on their spatial characteristics and dominant light properties. Having scenes with similar layouts allows for better extraction of result implications when making recommendations for the different spaces.

Four Groups were made:
(1) Outside scenes, (2) Inside scenes with big windows, (3) Inside scenes with open spaces, (4) Inside corridor scenes.

## Variables[6]

Variables extracted from raw data are:

- *Participants* (p101 to p120) coded for anonymity
- *ANSWER* correct response to the image
- *Correct* (0 or 1) binary variable to check if what participants answered matches ANSWER
- *Response* answers given by participants
- *ProbeType* to act as a predictive measure of what probe condition: Black=1,White=2, Golf=3
- *NumScene* refers to the number of scene of the image (1-14 except 11)
- *NumProbe* refers to the original probe that has been placed on a scene (1-14 except 11)
- *ImageName* in the format of ProbeType_NumberScene.NumberProbe
- *Response time* as an excluding variable to detect outliers

Variables created are:

- *Coding* (0-4) which corresponds to the four-step difficulty scale to predict image performance according to visual inspection
- *Grouping* (1-4) captures what space properties are characteristic of a scene

[5] See Appendix 4 for scenes in Group classification.
[6] See Appendix 5 for a Variables table.

# ANALYSIS

Data was analyzed using IBM SPSS 23 Statistics.

## Participant

Before starting the analysis, raw data was formatted and cleaned, creating a Mean Success Rate measure for each image based on the average number of the participants' responses.
*Mean Success Rate (MSR*) is the average of Response/Number of participants computed per each image. This variable ranges from 0 to 1 and will act as an output measure.
*Mean Success Rate per participant (MSRparticipant)* has also been computed to determine how sensitive participants are in determining the correct fit of the probes in a scene.

The 20 participants are represented on the X axis. The MSRparticipant scale is represented on the Y axis. There is a green line at the 0,5 level indicating participants answered correctly for half of the images.

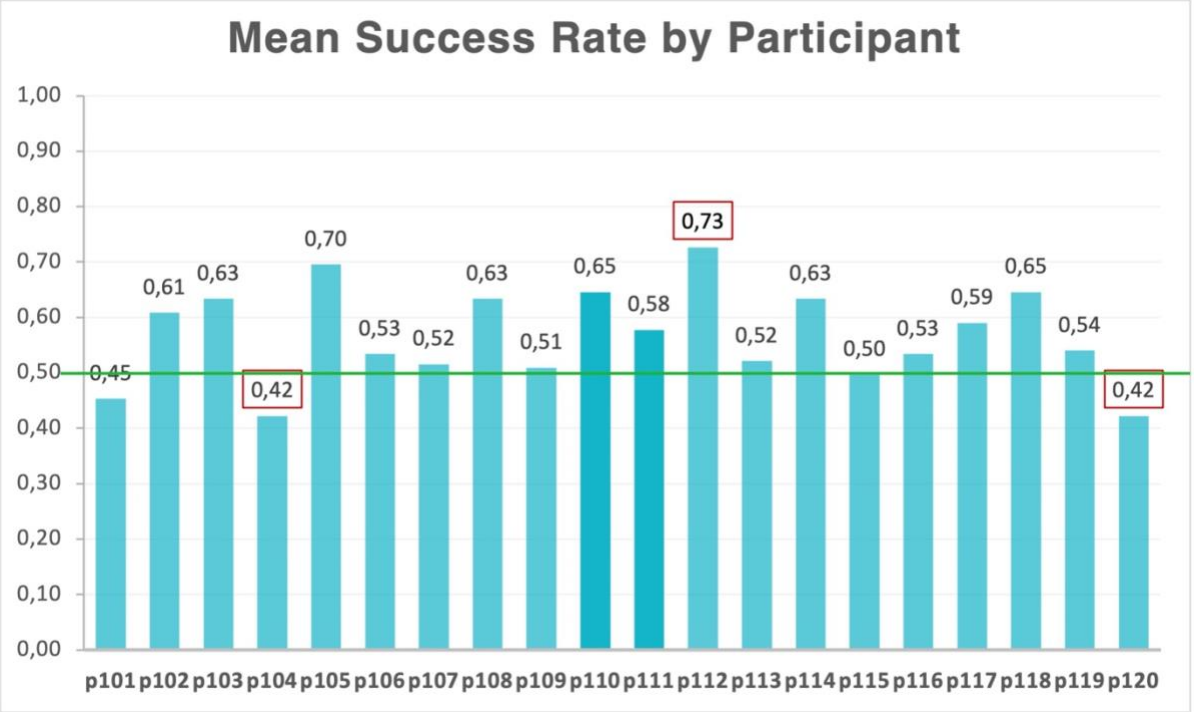



**Figure 2.** Graph of the Proportion of correct images answered by each participant.

As it can be seen, participants differed in their sensitivity to determine light probes' fit in different scenes. Participant p112 had the best performance while participants p104 and p120 had the worse with a MeanSuccessRate value of 0,42. This result indicates they answered correctly to 42% of all images, worse than if done at chance level.
We also checked how many times participants answered FITS compared to DOES NOT FIT. Seven participants answered FITS less than 50% of the time, compared to the correct amount of times the FITS answer was correct which is 25%. The other 13 participants answered FITS with a frequency of between 50 and 65% of the time.

## Image

First, it was important to illustrate how each image was evaluated by all participants to check how they performed and what images were easier or harder to judge. To do so, a frequency table with the Mean Success Rate per image (MSRimage)was created.

The table below has the *MSRimage* which corresponds to the average proportion (between 0 and 1) of participants that correctly identified an image. The closest a number is to 1 the better the image was judged.
Each MSRimage number has a bar with the names of the images that performed at that level. The higher the bar in each MSRimage point the more images were correctly judged at that proportion.
The number of images correctly judged for each proportion is englobed in the *Frequency* row below.

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | black_4-1 | black_1-1 | | | | | | |
| | | | | | | | | | | | | | black_4-4 | black_1-6 | | | | | | |
| | | | | | | | | | | | | | black_5-3 | black_2-5 | | | | | | |
| | | | | | | | | | | | | | black_5-4 | black_3-3 | | | | | | |
| | | | | | | | | | | | | | black_8-5 | black_4-13 | | | | | | |
| | | | | | | | | | | | | | black_8-8 | black_5-2 | | | | | | |
| | | | | | | | | | | | | black_2-2 | black_9-6 | black_6-6 | | | | | | |
| | | | | | | | | | | | | black_9-10 | black_9-7 | black_6-9 | | | | | | |
| | | | | | | | | | | | | black_14-2 | black_10-10 | black_9-9 | | | | | | |
| | | | | | | black_6-7 | | | | | black_3-9 | golf_2-2 | black_13-13 | black_12-7 | black_1-3 | | | | | |
| | | | | | | black_8-9 | | | | | black_6-6 | golf_2-5 | black_14-14 | black_12-12 | black_2-7 | | | | | |
| | | | | | | black_14-12 | | | | | black_8-14 | golf_3-10 | black_14-5 | golf_1-6 | black_5-5 | | | | | |
| | | | | | | golf_3-5 | | | black_4-10 | | black_8-5 | golf_4-10 | golf_1-8 | golf_2-4 | golf_3-3 | black_1-8 | | | | |
| | | | | golf_3-9 | golf_7-6 | | | | black_7-6 | | black_13-10 | golf_12-2 | golf_5-2 | golf_4-4 | golf_8-14 | black_2-4 | | | | |
| | | | | golf_6-8 | golf_7-8 | | | | black_10-4 | | golf_1-3 | golf_13-12 | golf_5-5 | golf_6-6 | golf_10-7 | black_7-8 | black_7-7 | | | |
| | | | | golf_8-9 | golf_10-10 | | | | golf_1-1 | | golf_2-7 | white_2-5 | golf_13-10 | golf_9-9 | golf_12-12 | black_7-14 | black_10-13 | | | |
| | | | | golf_9-6 | white_2-7 | | | | golf_6-12 | black_3-10 | golf_8-5 | white_3-10 | white_2-4 | golf_12-7 | golf_14-14 | black_10-7 | golf_4-1 | | | |
| | | | | golf_12-14 | white_6-7 | black_12-14 | golf_1-8 | | golf_7-7 | black_12-2 | golf_8-8 | white_4-1 | white_4-10 | white_1-6 | white_6-6 | black_13-12 | golf_4-13 | | | |
| | | | | golf_14-12 | white_6-9 | black_13-5 | golf_14-2 | | golf_9-10 | white_1-3 | golf_10-4 | white_4-4 | white_4-13 | white_2-2 | white_7-7 | golf_3-3 | golf_10-13 | | | |
| | | | golf_5-3 | white_7-8 | white_9-7 | golf_6-7 | white_1-8 | | white_3-9 | white_3-6 | golf_14-5 | white_5-2 | white_10-7 | white_7-14 | white_8-8 | golf_7-14 | golf_13-13 | | | |
| | | black_3-5 | white_5-3 | white_8-9 | white_12-2 | golf_9-7 | white_6-12 | | white_10-4 | white_10-10 | white_8-5 | white_5-4 | white_10-13 | white_12-12 | white_13-13 | white_3-3 | white_1-1 | | | |
| golf_13-5 | white_13-5 | white_14-2 | white_9-6 | white_12-14 | white_12-7 | white_7-6 | white_9-10 | | white_8-14 | white_13-12 | white_14-12 | white_9-9 | white_13-10 | white_14-5 | white_13-13 | white_5-5 | white_14-14 | black_6-12 | golf_5-4 | |
| **MSRimage** | **0,05** | **0,1** | **0,15** | **0,2** | **0,25** | **0,3** | **0,35** | **0,4** | **0,45** | **0,5** | **0,55** | **0,6** | **0,65** | **0,7** | **0,75** | **0,8** | **0,85** | **0,9** | **0,95** | **1** |
| **Frequency** | **1** | **1** | **2** | **3** | **9** | **13** | **5** | **5** | **10** | **6** | **13** | **16** | **22** | **22** | **13** | **10** | **8** | **1** | **1** | **0** |

**Figure 3.** Frequency table of the proportion of correctly identified images. Thee proportions are each the average correct response by all participants.

In a slightly darker color are the 5 images used for the practice trial of the experiment.
The mode of the table is between 0,65 and 0,7 with 22 images each. The median of images (80,5/160) is found between 0,55 and 0,6.

Because every image was evaluated by all participants, Repeated-Measures ANOVA tests were performed for each image. As a Factor, we added the Type of Probe, coded as follows: Black=1, Golf=2, White=3.
When sphericity was violated the Greenhouse-Geisser correction was applied. Post-hoc tests were conducted when results were significant at p=0,05 with a Bonferroni correction.
Significant differences due to a main effect of probe were found in 13 images[7]. After Post hoc tests[8], only 12 images presented at least one significant difference in the evaluation of MeanSuccessRate between the types of probes.

**Table 1.** Output of Pairwise comparisons for images with significant images after a Repeated-Measures ANOVA for each image. In Image Sig. there is the significance level found in the images during each ANOVA test. The three other columns correspond to the pairwise comparisons between all probe types.

| **Pairwise Comparisons of significant images** | | | | |
|---|---|---|---|---|
| | | Mean Difference b | | |
| Image | Image Sig. | Black-Golf | Black-White | Golf-White |
| 1.1 | ** | 0,250 | -0,150 | -0,400 b |
| 2.7 | ** | 0,200 | 0,450 b | 0,250 |
| 5.4 | ** | -0,300 b | 0,050 | 0,350 b |
| 6.8 | ** | 0,450 b | 0,400 | -0,050 |
| 6.12 | ** | 0,450 b | 0,500 b | 0,050 |
| 7.7 | * | 0,400 b | 0,100 | -0,300 |

[7] See Appendix 6 for a table with statistical results of Image comparison by ProbeType.
[8] See Appendix 6 for Post hoc tests for each significant image.

| | | | | |
|---|---|---|---|---|
| 7.8 | ** | 0,500 b | 0,550 b | 0,050 |
| 8.9 | ** | 0,500 b | 0,500 b | 0,000 |
| 9.6 | ** | 0,400 b | 0,450 b | 0,050 |
| 9.7 | * | 0,300 | 0,350 b | 0,050 |
| 12.7 | ** | 0,000 | 0,400 b | 0,400 b |
| 13.5 | * | 0,300 | 0,250 | -0,050 |
| 14.2 | * | 0,200 | 0,450 b | 0,250 |
| *pvalue <0,05; **pvalue <0,01 | | | | |
| b. The Mean Difference is significant at the ,05 level | | | | |

## Probes

To check the accuracy of our predictions[9] regarding image performance, a linear regression model for each of the three probe types was conducted.
A visual inspection of the difficulty to correctly deterimine probe fit in a scene was done for each image using a 0 to 4 step scale and encompassed in the Coding variable.

The dependent variable for the regression MeanSuccessRate for each image, and the independent variable was our Coding for each probe type, that acts as a predictive measure.
In the horizontal axis is represented the predicted difficulty of an image. The higher the number the more difficult it was coded. The vertical axis corresponds to the proportion of correct answers each image received.

For all black images, data is not linear but there is a clear downward tendency. This means images coded as more difficult were also the more difficult for participants to correctly determine.
For golf probes data is much more noisy but there is also a downward tendency. Our coding doesn't predict outcomes very well.
Finally, for white probes we again see that downward tendency and even thought there is still some variability data is less dispersed.

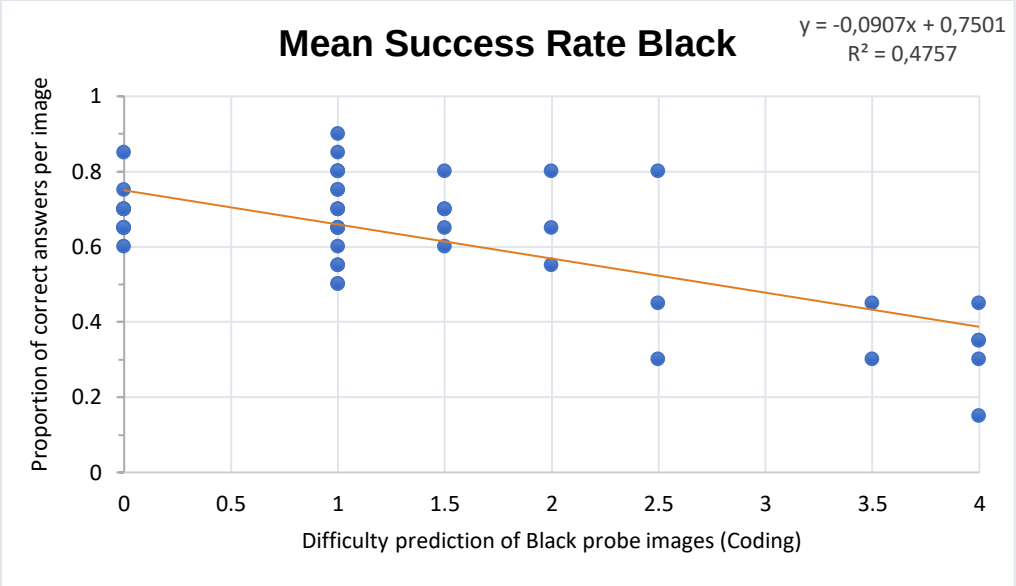



**Figure 4.** Linear regression graph between our predictions of performance for each Black probe image and the actual performance extracted from the amount of correct answers from all participants for each image.

[9] See Appendix 7 for all output under Coding Regression Output by Probe.

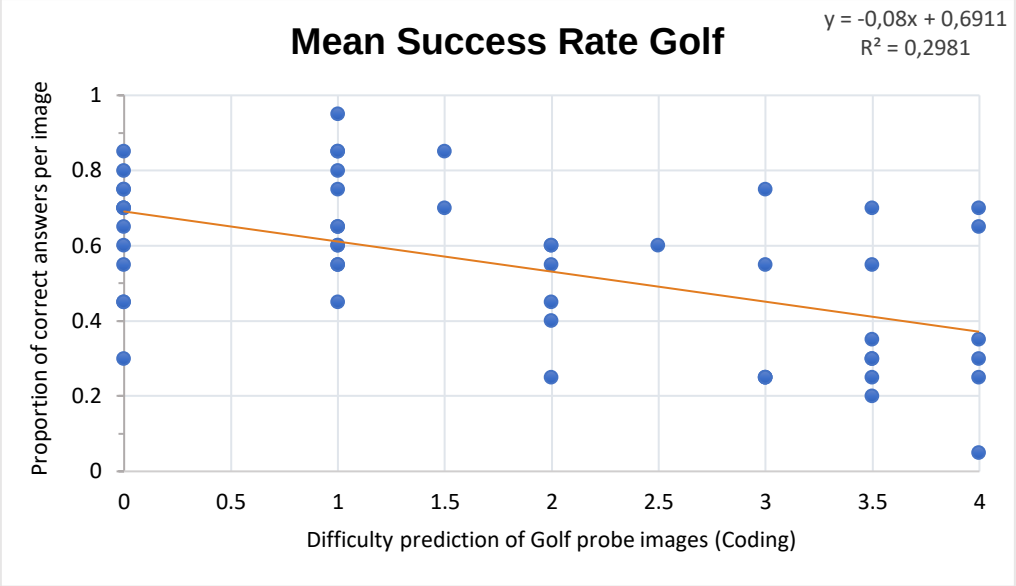



**Figure 5.** Linear regression graph between our predictions of performance for each Golf probe image and the actual performance extracted from the amount of correct answers from all participants for each image.

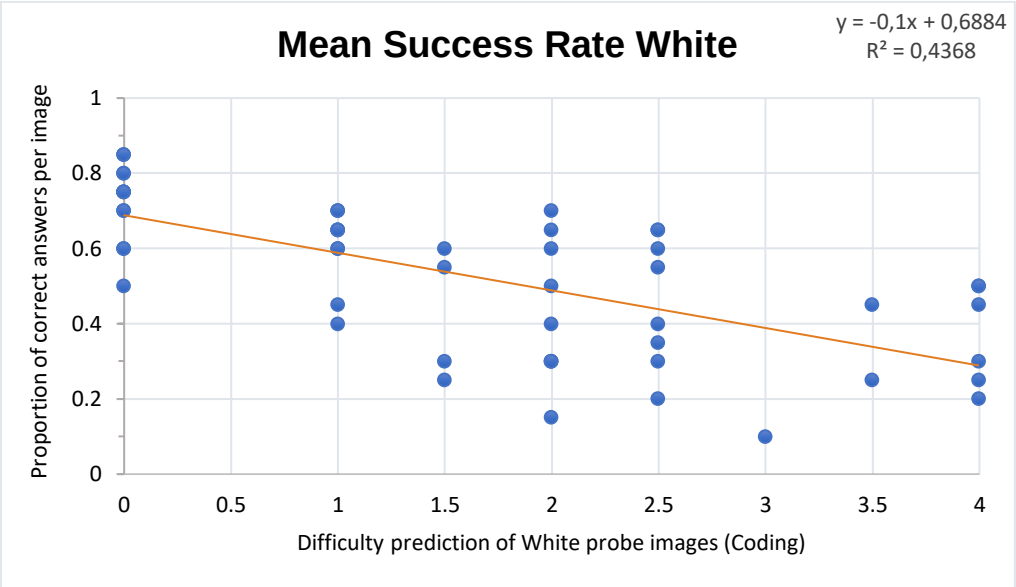



**Figure 6.** Linear regression graph between our predictions of performance for each White probe image and the actual performance extracted from the amount of correct answers from all participants for each image.

## Groups

A two-way ANOVA test was conducted to examine the effect of different probe types and Groups on the average image performance score. $F(6,144) = 0{,}778$, $p = 0{,}589$.

There were 2 independent variables, ProbeType and Group, and the average performance of each image (MSRimage) was the dependent variable.

- ProbeType englobes three conditions: 1=Black, 2=Golf, 3=White.

- Group consists of 4 conditions and refers to the main scene characteristics present in an image: 1=Outdoor scenes, 2=Indoor with big windows, 3=Indoor open spaces, 4=Indoor corridors.

**Table 2.** Table showing the significance of a Two-way ANOVA model with ProbeType and Group as independent variables and image performance (MeanSuccessRate image) as the dependent variable.

| **Two-way ANOVA** | | | | | |
|---|---|---|---|---|---|
| Dependent Variable: MeanSuccessRate image | | | | | |
| Source | Type II Sum of Squares | df | Mean Square | F | Sig. |
| Corrected Model | 0,836[a] | 11 | 0,076 | 2,139 | 0,021* |
| Intercept | 47,208 | 1 | 47,208 | 1329,195 | 0,000** |
| ProbeType | 0,316 | 2 | 0,158 | 4,443 | 0,013* |
| Group | 0,343 | 3 | 0,114 | 3,220 | 0,025* |
| ProbeType*Group | 0,166 | 6144 | 0,028 | 0,778 | 0,589 |
| Error | 5,114 | 156 | 0,036 | | |
| Total | 55,930 | 155 | | | |
| Corrected Total | 5,950 | | | | |
| a. R Squared= 0,140 (Adjusted R Squared =0,075) | | | | | |

There is a statistically significant difference in MeanSuccessRate of images between ProbeTypes (p=0,013) and between Groups (p=0,025). However, there is not a statistically significant interaction (ProbeType*Group) effect in the dependent variable MSRimage.
Because the Interaction term (ProbeType*Group) is not significant, but the individual variables are (p<0,05), we will directly look at the Post-hoc test results[10] with Turkey Multiple Comparisons table for both Group and ProbeType.
In the Multiple Comparisons for Groups, there were no statistically significant differences between any of the four Groups, so we have not included the table. For the ProbeType, we find significant differences (p=0,011) between ProbeType 1 (Black) and ProbeType 3 (White).

**Table 3.** Table of Multiple Comparisons between each ProbeType after showing main differences in the variable during the Two-way ANOVA.

| **Multiple Comparisons of ProbeType** | | | |
|---|---|---|---|
| Probe | Sig. | Mean Difference (I-J) | Std. Error |
| Black-Golf | 0,077 | 0,0808 | 0,03696 |
| Black-White | 0,011* | 0,1077 | 0,03696 |
| Golf-White | 0,747 | 0,0269 | 0,03696 |
| Based on observed means.<br>The error term is Mean Square(Error) = 0,036. | | | |
| *. The mean difference is significant at the ,05 level. | | | |

[10] See Appendix 8 for all the Results of Multiple comparisons from the Two-Way ANOVA.

We plotted the Probe Type and Group to have a more visual representation of the variables' behavior for each level.

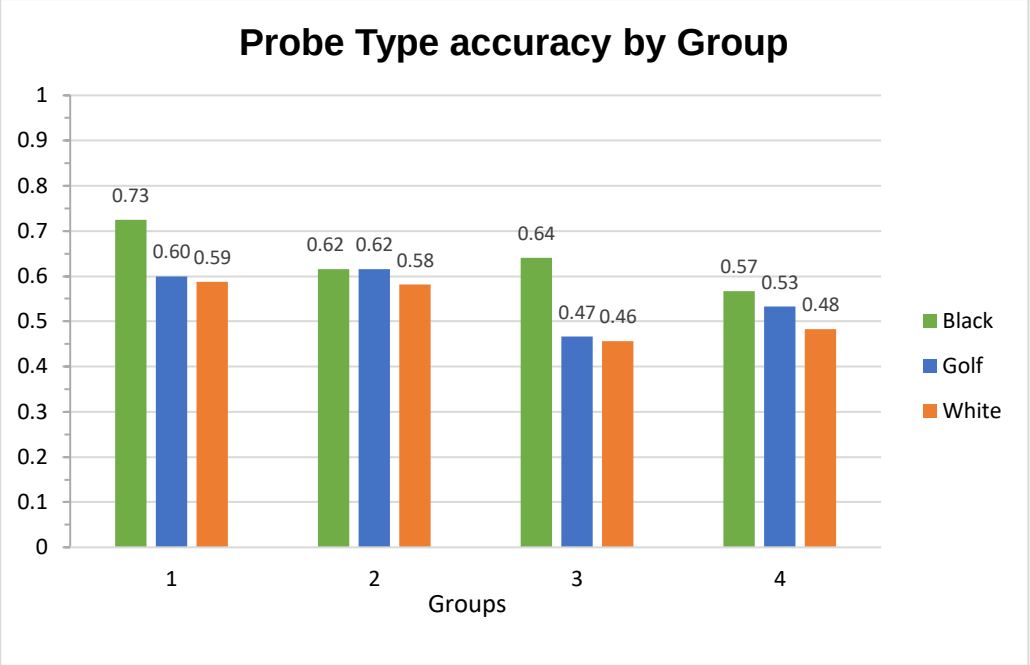



**Figure 7**. Graph representing the interaction between the three different probe types for each group of scenes. Note, ProbeType numbers are as follows: 1=Black, 2=Golf, 3=White. And Groups represent, respectively, 1=Outdoor scenes, 2=Indoor with big windows, 3=Indoor open spaces, 4=Indoor corridors.

## RESULTS AND DISCUSSION

We conducted an experiment to investigate if light probe's physical differences had an effect on the cues individuals take when judging a scene because of the difference in the light properties most relevant appearance. This is, to our knowledge, the first experiment to test probes with different texture and gloss against the more mainstream Lambertian matte sphere.
The task to answer whether a probe fits the scene's illumination is quite standard in these type of studies and allows the task to be simplified.

Participants were not able to judge correctly in all cases, and some participants performance was worse than that others. Judging by the results of the amount of times participants answered FITS, it cannot be ruled out this performance is the best anyone can do, since for 7 out of the 20 participants the response to FITS was answered less than half the time and for all participants it was lower than 2/3 of total answers (66%). This result suggest a clear awareness of the task at hand.

However, we could hypothesize individual's performance was not the best it can be due to certain elements that can be improved by implementing them in the experimental design phase:

- Paying attention to certain information but not to other
- Confusion in stimuli due to extreme similarity and randomized order of appearance
- Low target prevalence

It can be argued that by giving more detailed instructions and providing feedback during the practice trial, the first cause could be improved. As for the confusion in stimuli, the fact images were not shown following an order either by scene or probe type likely confused the participants during the task.

As an anecdote, after finishing the experiment participant p106 mentioned several images were repeated, proving they were not aware of -at least- minor differences between probes. This participant answered 53% of images correctly.
For the low target prevalence, participants were unaware of the proportion of FITS/DOES NOT FIT images show (39/117). As a result, and similar to signal detection theory in vigilance tasks (Wolfe et al., 2007), they tended to equalize the numbers for misses and false alarms to 50/50, resulting in an increase to their response to FITS.

Regressions of each ProbeType show that data is very noisy and doesn't follow a clear straight line but there is an overall downward tendency between the MeanSuccessRate of images and our Coding predictions which suggests there is a mild causal relationship, and it is backed up by statistical data in the Regressions done for our Coding and each Probe Type ANOVA (see Appendix 7). As shown in Figures 5,6,7, the images given a higher score in Coding based on visual inspection were also the images participants found harder to determine the fit in, regardless of the probe used.
Given the previously mentioned obstacles participants might have experienced, these results are encouraging in supporting our predictions determining the difficulty of an image based on mismatches and noticeably different light properties between probe and scene and support both Koenderink et al (2007) and Kartashova et al (2016) statement that human observers have a strong idea of how objects might appear in a space, but that expectation is a simplified representation of the (physical) light field.

In Figure 3, the frequency table with all images ordered by MeanSuccessRate performance shows no image was correctly judged by all participants. It is also interesting that the worst performing image is 13.5 for the Golf probe and White probe version. However, even though image 13.5 showed differences between probe types, after post hoc tests those differences proved not to be significant.
At the individual image level, main differences are found in the assessment of Black probes which performed consistently better compared to White probes and Golf probes in 10 out of 12 images with significant differences. This finding is consistent with the results from our Two-way ANOVA, where we found Black probes were 17% better at helping infer light in a scene than White probes in overall images ($p=0{,}011$). This result indicates Black shiny probes stand out from the other probes and should be incorporated as a tool for designers when showing their projects.
The outliers between the significant images were image 1.1, where White probes were better than Golf probes, and image 5.4 where Golf balls performed significantly better than both Black and White probes. These two images correspond to Group 1 and Group 2, respectively.
All outdoor scenes were grouped in Group 1. The dominant light property in those spaces is diffuseness, best captured by White probes according to our hypothesis. Similarly, Group 2 englobes indoor scenes with big windows in which light comes from multiple different directions. Our hypothesis was Golf probes would be best for capturing this light property, as Xia et al (2014) proved.
Although no significant differences were found in the interaction between ProbeType and Group, results of this kind suggest Black probes might not be the best probe for all types of scenes and light properties most dominant in a scene could indeed have an influence to the type of probe best suited to infer a space's light properties.

It is interesting that 7 of the significant images belong to scenes englobed in Group 3 under indoor open scenes (scenes 6,7,8,9) characterized by strong brilliance, with mild diffuse light and strong directed light from focus lighting. In these images Black probes being significantly better performing than at least one other probe type. For image 9.7, Black probe performs significantly better than White only, for all others Black probes outperformed Golf probes and for images 6.12, 7.8, 8.9 and 9.6 Black probes proved better than both Golf and White probes.
For significant images 14.2 and 12.7, both belonging to Group 4 characterized by indoor corridors, Black probes performed significantly better than White probes, but no other differences were found. These results are interesting because both these images were coded as easier to identify in the Black version, and harder for White probes. When the Black probe version of both images is shown, the mismatch between probe and scene is extremely clear due to Brilliance. For image 14.2, when the Black probe is presented an outside scene can be identified in the probe with only one main illumination coming directly from behind. For image 12.7, the Black probe shows multiple bright spots coming from

all across the ceiling that are not found in the scene. These findings are consistent with our third hypothesis, that Black probes are best at capturing Brilliance.
It is worth noting that even though we did not find significant differences in the interaction term between ProbeType*Group in the Two-way ANOVA, differences can be appreciated from the Figure 7 graph of *ProbeType Accuracy by Group* in which each ProbeType was represented for each group.
Both Group 1 and Group 3 stand out, with Black probes (Type 1 in Blue) clearly performing better than their counterparts. Our results about the different Groups encompassing scenes with similar characteristics are in line with the findings of Kartashova et al. (2016) that state the perception of the luminous environment is dependent on the geometric characteristics of a scene, as well as the materials and amount of objects placed in it.

The original question of whether a certain probe type was better for judging a scene's illumination based on its most relevant light properties cannot be proven but our results are encouraging and, at the image level, support out hypotheses. Further research needs to be done in this area to test for variations in scene-specific illumination perception with different light probes. This subject conveys great implications for design and lighting professionals' presentation of their projects and would consolidate the use of different scene-specific light probes for visualization of a space.
It is our belief that by improving the information and feedback given to participants at the begging of the experiment, performance will prove less noisy and research could better control for chance-level responses.

## CONCLUSIONS

In conclusion, our results show Black probes are generally better at helping determine light than the more conventional and widespread Lambertian White matte spheres.
Unfortunately our hypotheses that one probe type is better than others in helping infer certain light properties in scenes with similar features cannot be proven, but significant results at the image level suggest a conditional relationship of this matter might exist.

We hope our work can be used as a starting point as further research is needed in this area to determine if there are indeed significant differences in the performance of different probes for specific scenes, as well as determining what type of probe should be used by designers and lighting professionals when presenting a space with a particularly dominant light property.

# REFERENCES


1. Adams, E. (2013). The elements and principles of design. *iJADE, 32.2,* 157-175. https://doi-org.proxy.library.uu.nl/10.1111/j.1476-8070.2013.01761.x
2. Cuttle, C. (2003). Lighting by Design. Architectural Press.
3. Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics (5th ed.). SAGE Publications.
4. Gershun, A. (1939). The light field (translated by Moon, P.H. & Timoshenko, G.) Journal of Mathematics and Physics, 18, 51-151.
5. Kartashova, T., De ridder, H., te Pas, S.F. & Pont, S.C. (2015). The visual light field in paintings of Museum Prinsenhof: comparing settings in empty space and on objects, Proc. SPIE 9394, Human Vision and Electronic Imaging XX, 93941M. https://doi.org/10.1117/12.2085030
6. Kartashova, T., De ridder, H., te Pas, S.F. & Pont, S.C. (2018). Visual light zones. *Perception, 9(*3), 1-20. https://doi-org.proxy.library.uu.nl/10.1177/2041669518781381
7. Kartashova, T., De ridder, H., te Pas, S.F. & Pont, S.C. (2019). A toolbox for volumetric visualization of light properties. Lighting Res. Technol., 51, 838-857.
8. Kartashova, T., te Pas, S.F., de Ridder, H. & Pont, S.C. (2019). Light Shapes: perception-Based Visualizations of the global light transport. *ACM Transactions on Applied Perception*. 16,1, Article 4 (January 2019), 17 pages.
9. Koenderink , J.J. & Pont, S.C.(2003). Irradiation direction form texture Journal of the Optical Society of America A 20, 1875-1882. https://doi-org.proxy.library.uu.nl/10.1364/JOSAA.20.001875
10. Koenderink, J.J., & Pont, S.C., van Doorn, A.J., Kappers, A.M.L., Todd, J.T. (2007). The visual light field. *Perception, 36,* 1565–1610. https://doi.org/10.1068/p56772
11. Mury, A.A., Pont, S.C. & Koenderink, J.J. (2007). Light field constancy within natural scenes. Optical Society of America. *Applied Optics,* 46(29), 7308-7316. https://doi-org.proxy.library.uu.nl/10.1364/AO.46.007308
12. Mury, A.A. The light field in natural scenes. (2009). PhD thesis. *Technical University of Delft.*
13. Mury, A.A., Pont, S.C. & Koenderink, J.J. (2009). Structure of light fields in natural scenes. Optical Society of America. *Applied Optics,* 48, 5386-5395. https://doiorg.proxy.library.uu.nl/10.1364/AO.48.005386
14. Schirillo, J.A. (2013). We infer light in space. *Psychonomic Society,* 20, 905–915. https://doi-org.proxy.library.uu.nl/10.3758/s13423-013-0408-1
15. Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General, 136*(4), 623–638. https://doi-org.proxy.library.uu.nl/10.1037/0096-3445.136.4.623
16. Xia, L., Pont, S. C., & Heynderick, I. (2017). Separate and simultaneous adjustment of light qualities in a real scene. *i-Perception, 8*(1), Article 2041669516686089. https://doi-org.proxy.library.uu.nl/10.1177/2041669516686089
17. Xia, L., Pont, S.C. & Heynderickx, I. (2014). The visual light field in real scenes. *i-Perception, 5*, 613-629. https://doi-org.proxy.library.uu.nl/10.1364/AO.48.005386
18. Xia, L., Pont, S., & Heynderickx, I. (2016). Light diffuseness metric Part 1: Theory. *Lighting Research & Technology*, *49*(4), 411–427. https://doi.org/10.1177/1477153516631391
19. Xia, L., Pont, S., & Heynderickx, I. (2016). Light diffuseness metric Part 2: Describing, measuring and visualizing the light flow and diffuseness in three-dimensional spaces. *Lighting Research & Technology*, *49*(4), 1–18. http://doi.org/10.1177/1477153516631392

# APPENDIX 1

## Instructions

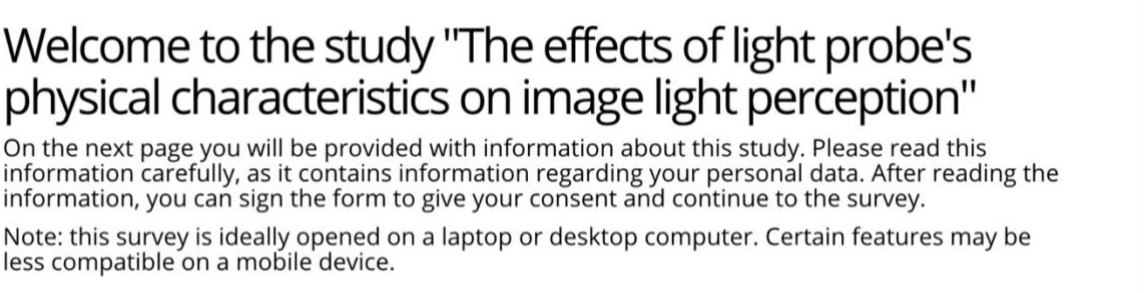
On each screen you will be shown an image with a sphere. After 3 seconds two buttons will appear in the screen.

Use the buttons to indicate **if the sphere's illumination fits the scene**.

You will have two 1minute breaks along the experiment. The maximum allowed time to take the experiment is 3:30h.

But, let's do a practice run first!

## Information letter


Welcome to the study "The effects of light probe's physical characteristics on image light perception"

On the next page you will be provided with information about this study. Please read this information carefully, as it contains information regarding your personal data. After reading the information, you can sign the form to give your consent and continue to the survey.

Note: this survey is ideally opened on a laptop or desktop computer. Certain features may be less compatible on a mobile device.

Aim of the study

The aim of this study is to get insight into humans' inference of light in images using spherical objects called light probes. By collecting data on this subject, we aim to answer questions about how surface structure, material and shape of light probes affect (correct) light assessments of an object in space.
This is a student research.

Data collection and storage

Personal data will not be collected. Participants will be asked to provide an email address to send them a link to the study. Data will remain confidential and will be anonymized before being stored using YoDa Storage. Only the researchers will have access to the full dataset and only anonymized data will be used in scientific publication.
Your data will be stored for at least 10 years. Anonymized versions of the data could be published in scientific literature. In addition, any data collected may be used for follow-up research or research with another purpose.

Content of the study

This task will take no longer than 60 minutes.

**In this study you will be asked to stare at a series of images with a sphere photographed in them and decide if the lighting of the sphere fits the image by selecting a button.**
For taking part in our study you will receive 1 PPU (participation credits, 0.25 every 15 minutes) or financial compensation of 8€ per hour (2€ for 15 minutes).
Taking part in our student research is voluntary and may be terminated at any moment. Termination will not have any consequences and may be done without providing a reason for doing so. Data that is collected up to the point of termination may be used for research.

## Contact information

If you have questions or remarks regarding this student research, please contact **Sònia Fanlo Garcia**, at s.fanlogarcia@students.uu.nl.
If you would rather address your remarks to a person independent of this research, please contact Susan Te Pas, at s.tepas@uu.nl.
Formal complaints can be directed to klachtenfunctionaris-fetcsocwet@uu.nl.

This information letter was last reviewed on December 14th 2020.

# APPENDIX 2

## Stimuli

### Scene 1

Black_1.1 Golf_1.1 White_1.1

Black_1.3 Golf_1.3 White_1.3

Black_1.6 Golf_1.6 White_1.6

Black_1.8 Golf_1.8 White_1.8

**Scene 2**

Black_2.2

Golf_2.2

White_2.2

Black_2.4

Golf_2.4

White_2.4

Black_2.5

Golf_2.5

White_2.5

Black_2.7

Golf_2.7

White_2.7

**Scene 3**

Black_3.3

Golf_3.3

White_3.3

Black_3.5

Golf_3.5

White_3.5

Black_3.9

Golf_3.9

White_3.9

Black_3.10

Golf_3.10

White_3.10

## Scene 4

Black_4.1

Golf_4.1

White_4.1

Black_4.4

Golf_4.4

White_4.4

Black_4.10

Golf_4.10

White_4.10

Black_4.13

Golf_4.13

White_4.13

**Scene 5**

Black_5.2 Golf_5.2 White_5.2

Black_5.3 Golf_5.3 White_5.3

Black_5.4 Golf_5.4 White_5.4

Black_5.5 Golf_5.5 White_5.5

## Scene 6

Black_6.6

Golf_6.6

White_6.6

Black_6.7

Golf_6.7

White_6.7

Black_6.9

Golf_6.8

White_6.9

Black_6.12

Golf_6.12

White_6.12

## Scene 7

Black_7.6

Golf_7.6

White_7.6

Black_7.7

Golf_7.7

White_7.7

Black_7.8

Golf_7.8

White_7.8

Black_7.14

Golf_7.14

White_7.14

## Scene 8

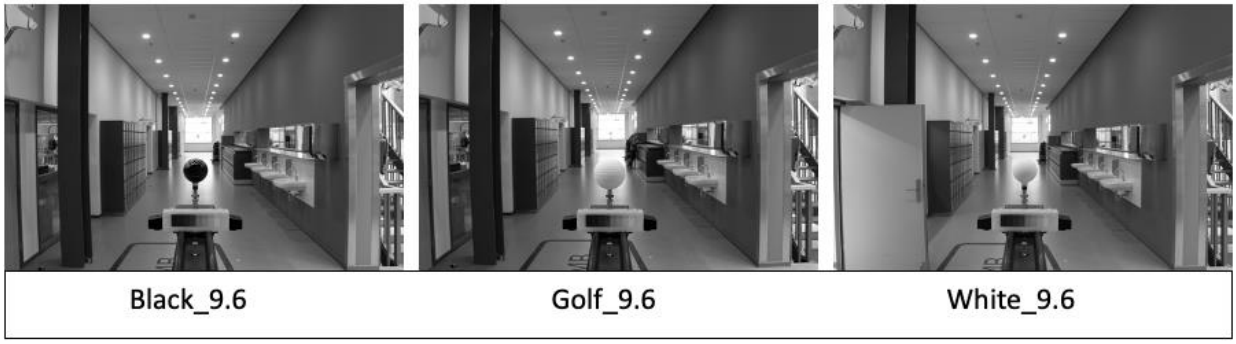
| Black_8.5 | Golf_8.5 | White_8.5 |
|---|---|---|

| Black_8.8 | Golf_8.8 | White_8.8 |
|---|---|---|

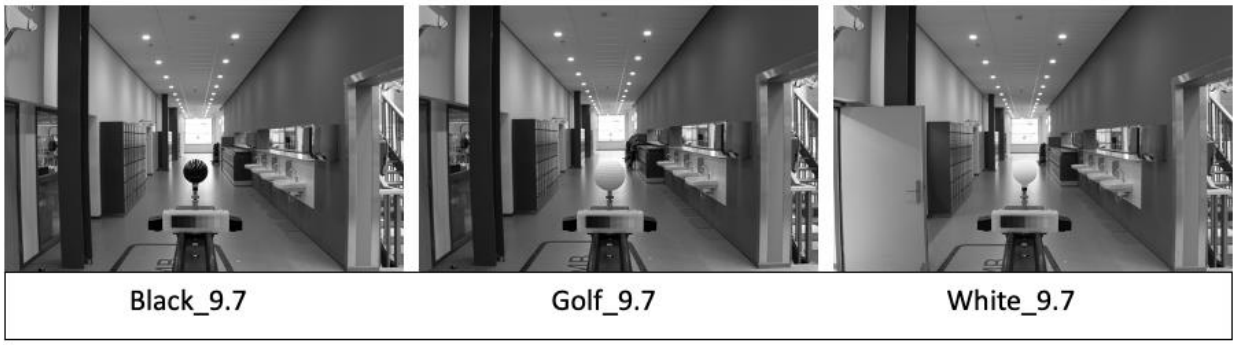
| Black_8.9 | Golf_8.9 | White_8.9 |
|---|---|---|

| Black_8.14 | Golf_8.14 | White_8.14 |
|---|---|---|

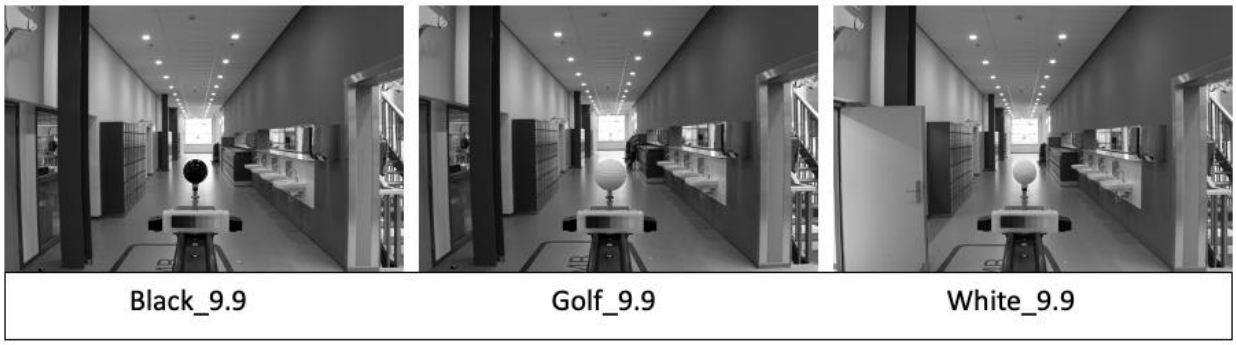
## Scene 9



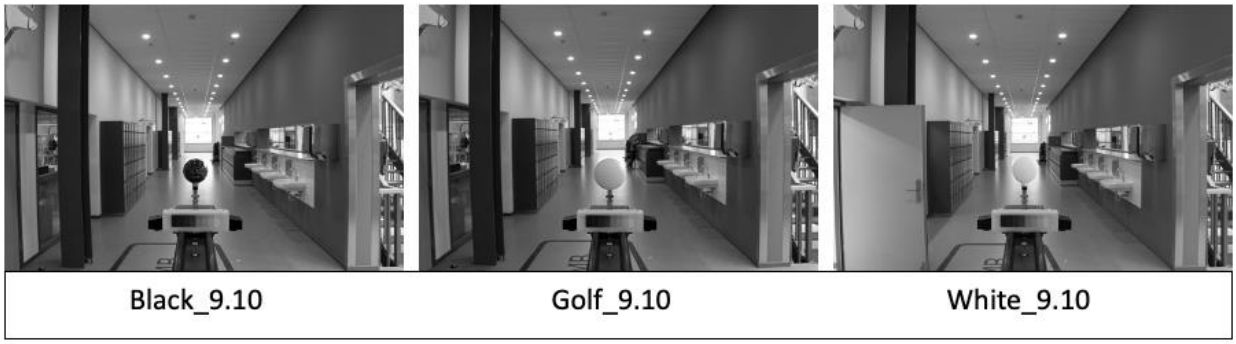
| Black_9.6 | Golf_9.6 | White_9.6 |
|---|---|---|



| Black_9.7 | Golf_9.7 | White_9.7 |
|---|---|---|



| Black_9.9 | Golf_9.9 | White_9.9 |
|---|---|---|



| Black_9.10 | Golf_9.10 | White_9.10 |
|---|---|---|

## Scene 10


Black_10.4


Golf_10.4


White_10.4


Black_10.7


Golf_10.7


White_10.7


Black_10.10


Golf_10.10


White_10.10


Black_10.13


Golf_10.13


White_10.13

**Scene 12**

Black_12.2 Golf_12.2 White_12.2

Black_12.7 Golf_12.7 White_12.7

Black_12.12 Golf_12.12 White_12.12

Black_12.14 Golf_12.14 White_12.14

**Scene 13**

Black_13.5

Golf_13.5

White_13.5

Black_13.10

Golf_13.10

White_13.10

Black_13.12

Golf_13.12

White_13.12

Black_13.13

Golf_13.13

White_13.13

**Scene 14**

| Black_14.2 | Golf_14.2 | White_14.2 |
|---|---|---|

| Black_14.5 | Golf_14.5 | White_14.5 |
|---|---|---|

| Black_14.12 | Golf_14.12 | White_14.12 |
|---|---|---|

| Black_14.14 | Golf_14.14 | White_14.14 |
|---|---|---|

# APPENDIX 3

## Coding

### Black probes

| Probetype | Scene | Probe | Difficulty difference between coders | Difficulty average |
|---|---|---|---|---|
| black | **1** | **1** | | 0 |
| black | **1** | **3** | 0 | 1 |
| black | **1** | **6** | 1 | 1,5 |
| black | **1** | **8** | 0 | 1 |
| black | **2** | **2** | | 0 |
| black | **2** | **4** | 1 | 1,5 |
| black | **2** | **5** | 1 | 1,5 |
| black | **2** | **7** | 0 | 1 |
| black | **3** | **3** | | 0 |
| black | **3** | **5** | 0 | 4 |
| black | **3** | **9** | 0 | 2 |
| black | **3** | **10** | 0 | 1 |
| black | **4** | **1** | 0 | 1 |
| black | **4** | **4** | | 0 |
| black | **4** | **10** | 1 | 2,5 |
| black | **4** | **13** | 1 | 1,5 |
| black | **5** | **2** | 0 | 1 |
| black | **5** | **3** | 0 | 2 |
| black | **5** | **4** | 0 | 1 |
| black | **5** | **5** | | 0 |
| black | **6** | **6** | | 0 |
| black | **6** | **7** | 1 | 3,5 |
| black | **6** | **9** | 0 | 1 |
| black | **6** | **12** | 0 | 1 |
| black | **7** | **6** | 1 | 3,5 |
| black | **7** | **7** | | 0 |
| black | **7** | **8** | 1 | 2,5 |
| black | **7** | **14** | 0 | 1 |
| black | **8** | **5** | 1 | 1,5 |
| black | **8** | **8** | | 0 |
| black | **8** | **9** | 1 | 2,5 |
| black | **8** | **14** | 0 | 1 |
| black | **9** | **6** | 0 | 1 |
| black | **9** | **7** | 0 | 1 |
| black | **9** | **9** | | 0 |
| black | **9** | **10** | 1 | 1,5 |
| black | **10** | **4** | 0 | 4 |
| black | **10** | **7** | 0 | 1 |
| black | **10** | **10** | | 0 |
| black | **10** | **13** | 0 | 1 |
| black | **12** | **2** | 0 | 1 |
| black | **12** | **7** | 0 | 1 |
| black | **12** | **12** | | 0 |
| black | **12** | **14** | 0 | 4 |
| black | **13** | **5** | 0 | 4 |
| black | **13** | **10** | 0 | 1 |
| black | **13** | **12** | 0 | 2 |
| black | **13** | **13** | | 0 |
| black | **14** | **2** | 0 | 1 |
| black | **14** | **5** | 0 | 1 |
| black | **14** | **12** | 0 | 4 |
| black | **14** | **14** | | 0 |

## Golf probes

| Probetype | Scene | Probe | Difficulty difference between coders | Difficulty average |
|---|---|---|---|---|
| golf | **1** | **1** | | 0 |
| golf | **1** | **3** | 0 | 3 |
| golf | **1** | **6** | 1 | 3,5 |
| golf | **1** | **8** | 0 | 4 |
| golf | **2** | **2** | | 0 |
| golf | **2** | **4** | 0 | 4 |
| golf | **2** | **5** | 0 | 1 |
| golf | **2** | **7** | 0 | 1 |
| golf | **3** | **3** | | 0 |
| golf | **3** | **5** | 1 | 3,5 |
| golf | **3** | **9** | 2 | 2 |
| golf | **3** | **10** | 0 | 1 |
| golf | **4** | **1** | 1 | 1,5 |
| golf | **4** | **4** | | 0 |
| golf | **4** | **10** | 1 | 2,5 |
| golf | **4** | **13** | 0 | 1 |
| golf | **5** | **2** | 0 | 1 |
| golf | **5** | **3** | 1 | 3,5 |
| golf | **5** | **4** | 0 | 1 |
| golf | **5** | **5** | | 0 |
| golf | **6** | **6** | | 0 |
| golf | **6** | **7** | 0 | 4 |
| golf | **6** | **8** | 1 | 3,5 |
| golf | **6** | **12** | 0 | 2 |
| golf | **7** | **6** | 0 | 4 |
| golf | **7** | **7** | | 0 |
| golf | **7** | **8** | 1 | 3,5 |
| golf | **7** | **14** | 0 | 1 |
| golf | **8** | **5** | 0 | 2 |
| golf | **8** | **8** | | 0 |
| golf | **8** | **9** | 0 | 4 |
| golf | **8** | **14** | 0 | 1 |
| golf | **9** | **6** | 0 | 3 |
| golf | **9** | **7** | 1 | 3,5 |
| golf | **9** | **9** | | 0 |
| golf | **9** | **10** | 0 | 1 |
| golf | **10** | **4** | 1 | 3,5 |
| golf | **10** | **7** | 0 | 3 |
| golf | **10** | **10** | | 0 |
| golf | **10** | **13** | 0 | 1 |
| golf | **12** | **2** | 0 | 2 |
| golf | **12** | **7** | 1 | 1,5 |
| golf | **12** | **12** | | 0 |
| golf | **12** | **14** | 1 | 3 |
| golf | **13** | **5** | 0 | 4 |
| golf | **13** | **10** | 0 | 1 |
| golf | **13** | **12** | 0 | 2 |
| golf | **13** | **13** | | 0 |
| golf | **14** | **2** | 0 | 2 |
| golf | **14** | **5** | 0 | 1 |
| golf | **14** | **12** | 0 | 3 |
| golf | **14** | **14** | | 0 |

## White probes

| Probetype | Scene | Probe | Difficulty difference between coders | Difficulty average |
|---|---|---|---|---|
| white | **1** | **1** | | 0 |
| white | **1** | **3** | 0 | 4 |
| white | **1** | **6** | 0 | 2 |
| white | **1** | **8** | 1 | 2,5 |
| white | **2** | **2** | | 0 |
| white | **2** | **4** | 2 | 2 |
| white | **2** | **5** | 0 | 1 |
| white | **2** | **7** | 1 | 2,5 |
| white | **3** | **3** | | 0 |
| white | **3** | **6** | 0 | 4 |
| white | **3** | **9** | 1 | 3,5 |
| white | **3** | **10** | 1 | 1,5 |
| white | **4** | **1** | 1 | 2,5 |
| white | **4** | **4** | | 0 |
| white | **4** | **10** | 1 | 2,5 |
| white | **4** | **13** | 0 | 1 |
| white | **5** | **2** | 0 | 1 |
| white | **5** | **3** | 0 | 4 |
| white | **5** | **4** | 0 | 2 |
| white | **5** | **5** | | 0 |
| white | **6** | **6** | | 0 |
| white | **6** | **7** | 0 | 2 |
| white | **6** | **9** | 0 | 4 |
| white | **6** | **12** | 0 | 2 |
| white | **7** | **6** | 1 | 2,5 |
| white | **7** | **7** | | 0 |
| white | **7** | **8** | 1 | 3,5 |
| white | **7** | **14** | 0 | 1 |
| white | **8** | **5** | 1 | 1,5 |
| white | **8** | **8** | | 0 |
| white | **8** | **9** | 0 | 4 |
| white | **8** | **14** | 0 | 1 |
| white | **9** | **6** | 1 | 2,5 |
| white | **9** | **7** | 0 | 2 |
| white | **9** | **9** | | 0 |
| white | **9** | **10** | 0 | 1 |
| white | **10** | **4** | 0 | 4 |
| white | **10** | **7** | 1 | 2,5 |
| white | **10** | **10** | | 0 |
| white | **10** | **13** | 0 | 1 |
| white | **12** | **2** | 1 | 1,5 |
| white | **12** | **7** | 0 | 2 |
| white | **12** | **12** | | 0 |
| white | **12** | **14** | 1 | 1,5 |
| white | **13** | **5** | 2 | 3 |
| white | **13** | **10** | 0 | 1 |
| white | **13** | **12** | 0 | 2 |
| white | **13** | **13** | | 0 |
| white | **14** | **2** | 2 | 2 |
| white | **14** | **5** | 0 | 1 |
| white | **14** | **12** | 1 | 2,5 |
| white | **14** | **14** | | 0 |

# APPENDIX 4

## Grouping

| Group 1 | Group 2 | Group 3 | Group 4 |
|---|---|---|---|
| Outside scenes | Inside scenes with big window | Inside scene in open space | Inside scene in corridors |
| Scene 1 | Scene 3 | Scene 6 | Scene 12 |
| Scene 2 | Scene 4 | Scene 7 | Scene 13 |
| | Scene 5 | Scene 8 | Scene 14 |
| | Scene 10 | Scene 9 | |

# APPENDIX 5

## Table: Variables

| Variable | Output | Explanation |
|---|---|---|
| Participant | p101 to p120 | 20 participants, coded for anonymity |
| ProbeType | Black=1, Golf=2, White=3 | Which probe was used |
| NumScene | 1-14 except 11 | Refers to the number scene of the image |
| NumProbe | 1-14 except 11 | Refers to the number of scene where the probe was originally photographed |
| ImageName | ProbeType_NumScene_NumProbe | Identify the image |
| ANSWER | FITS, DOES NOT FIT | Correct response |
| Correct | 0 or 1 | Binary variable to check if what participants answered matches ANSWER |
| Coding | 0-4 in 0.5 increments | Four-step difficulty scale how hard it is to adequately judge the image |
| Grouping | 1-4 | Captures what space properties are characteristic of a scene |
| MSRimage | 0-1 in 0.05 increments | Average proportion of correct responses each image had (SUM of Correct/Total participants) |
| MSRparticipant | 0-1 | Sensitivity of participants. Average number of correct responses/total responses |

# APPENDIX 6

## Image comparison by ProbeType

| Image | Mauchly sig | Sphericity | Greenhouse-Geisser | F Within Subjects | pvalue | df | Significance |
|---|---|---|---|---|---|---|---|
| 1.1 | 0,618 | yes | | 5,444 | 0,008 | 2 | Sig |
| 1.3 | 0,006 | no | 0,699 | 1,669 | 0,210 | 1,398 | Not sig |
| 1.6 | 0,00 | no | 0,632 | 0,000 | 1,000 | 1,263 | Not sig |
| 1.8 | 0,906 | yes | | 3,199 | 0,052 | 2 | Not sig |
| 2.2 | 0,817 | yes | | 0,297 | 0,740 | 2 | Not sig |
| 2.4 | 0,523 | yes | | 0,571 | 0,559 | 2 | Not sig |
| 2.5 | 0,007 | no | 0,7 | 0,487 | 0,554 | 1,401 | Not sig |
| 2.7 | 0,537 | yes | | 5,824 | 0,006 | 2 | Sig |
| 3.3 | 0,055 | yes | | 0,388 | 0,632 | 2 | Not sig |
| 3.6 | 0,701 | yes | | 2,869 | 0,070 | 2 | Not sig |
| 3.9 | 0,673 | yes | | 1,956 | 0,155 | 2 | Not sig |
| 3.10 | 0,482 | yes | | 0,322 | 0,727 | 2 | Not sig |
| 4.1 | 0,059 | yes | | 2,229 | 0,122 | 2 | Not sig |
| 4.4 | 0,93 | yes | | 0,222 | 0,800 | 2 | Not sig |
| 4.10 | 0,009 | no | 0,711 | 1,193 | 0,304 | 1,422 | Not sig |
| 4.13 | 0,855 | yes | | 1,088 | 0,347 | 2 | Not sig |
| 5.2 | 0,046 | no | 0,776 | 0,363 | 0,645 | 1,551 | Not sig |
| 5.3 | 0,059 | yes | | 2,229 | 0,112 | 2 | Not sig |
| 5.4 | 0,354 | yes | | 5,204 | 0,010 | 2 | Sig |
| 5.5 | 0,523 | yes | | 0,571 | 0,570 | 2 | Not sig |
| 6.6 | 0,826 | yes | | 0,079 | 0,924 | 2 | Not sig |
| 6.7 | 0,353 | yes | | 0,087 | 0,917 | 2 | Not sig |
| 6.8 | 0,139 | yes | | 5,195 | 0,010 | 2 | Sig |
| 6.12 | 0,32 | yes | | 7,550 | 0,002 | 2 | Sig |
| 7.6 | 0,791 | yes | | 0,769 | 0,471 | 2 | Not sig |
| 7.7 | 0,691 | yes | | 4,750 | 0,014 | 2 | Sig |
| 7.8 | 1,000 | yes | | 14,154 | 0,000 | 2 | Sig |
| 7.14 | 0,003 | no | 0,681 | 0,487 | 0,549 | 1,363 | Not sig |
| 8.5 | 0,763 | yes | | 0,297 | 0,745 | 2 | Not sig |
| 8.8 | 0,643 | yes | | 1,781 | 0,182 | 2 | Not sig |
| 8.9 | 0,948 | yes | | 7,308 | 0,002 | 2 | Sig |
| 8.14 | 0,117 | yes | | 2,771 | 0,075 | 2 | Not sig |
| 9.6 | 0,317 | yes | | 6,110 | 0,005 | 2 | Sig |
| 9.7 | 0,225 | yes | | 4,147 | 0,023 | 2 | Sig |
| 9.9 | 0,482 | yes | | 0,322 | 0,727 | 2 | Not sig |
| 9.10 | 0,46 | yes | | 0,925 | 0,405 | 2 | Not sig |
| 10.4 | 0,763 | yes | | 0,297 | 0,745 | 2 | Not sig |
| 10.7 | 0,541 | yes | | 0,689 | 0,508 | 2 | Not sig |
| 10.10 | 0,047 | no | 0,777 | 2,320 | 0,126 | 1,553 | Not sig |
| 10.13 | 0,476 | yes | | 2,452 | 0,100 | 2 | Not sig |
| 12.2 | 0,508 | yes | | 1,602 | 0,215 | 2 | Not sig |
| 12.7 | 0,378 | yes | | 7,795 | 0,001 | 2 | Sig |
| 12.12 | 0,389 | yes | | 0,068 | 0,934 | 2 | Not sig |
| 12.14 | 0,482 | yes | | 0,322 | 0,727 | 2 | Not sig |
| 13.5 | 0,192 | yes | | 3.953 | 0,028 | 2 | Sig |
| 13.10 | 0,007 | no | 0,700 | 0,487 | 0,554 | 1,401 | Not sig |
| 13.12 | 0,172 | yes | | 2,509 | 0,095 | 2 | Not sig |
| 13.13 | 0,029 | no | 0,755 | 0,919 | 0,385 | 1,510 | Not sig |
| 14.2 | 0,483 | yes | | 4,849 | 0,013 | 2 | Sig |
| 14.5 | 0,542 | yes | | 0,689 | 0,508 | 2 | Not sig |
| 14.12 | 0,102 | yes | | 2,365 | 0,108 | 2 | Not sig |
| 14.14 | 0,074 | yes | | 1,096 | 0,344 | 2 | Not sig |

# Pairwise Comparisons for images with significant main effect ANOVA

## Image 1.1

### Descriptive Statistics

| | Mean | Std. Deviation | N |
|---|---|---|---|
| MSR1 | ,70 | ,470 | 20 |
| MSR2 | ,45 | ,510 | 20 |
| MSR3 | ,85 | ,366 | 20 |

### Pairwise Comparisons

Measure: MSRimage1.1

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,250 | ,123 | ,169 | -,073 | ,573 |
| | 3 | -,150 | ,109 | ,559 | -,437 | ,137 |
| 2 | 1 | -,250 | ,123 | ,169 | -,573 | ,073 |
| | 3 | -,400* | ,134 | ,023 | -,751 | -,049 |
| 3 | 1 | ,150 | ,109 | ,559 | -,137 | ,437 |
| | 2 | ,400* | ,134 | ,023 | ,049 | ,751 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 2.7

### Estimates

Measure: image2.7

| ProbeType | Mean | Std. Error | 95% Confidence Interval Lower Bound | Upper Bound |
|---|---|---|---|---|
| 1 | ,750 | ,099 | ,542 | ,958 |
| 2 | ,550 | ,114 | ,311 | ,789 |
| 3 | ,300 | ,105 | ,080 | ,520 |

## Pairwise Comparisons

Measure: image2.7

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,200 | ,138 | ,488 | -,161 | ,561 |
| | 3 | ,450* | ,114 | ,003 | ,150 | ,750 |
| 2 | 1 | -,200 | ,138 | ,488 | -,561 | ,161 |
| | 3 | ,250 | ,143 | ,288 | -,125 | ,625 |
| 3 | 1 | -,450* | ,114 | ,003 | -,750 | -,150 |
| | 2 | -,250 | ,143 | ,288 | -,625 | ,125 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

### Image 5.4

## Estimates

Measure: image5.4

| ProbeType | Mean | Std. Error | 95% Confidence Interval Lower Bound | Upper Bound |
|---|---|---|---|---|
| 1 | ,650 | ,109 | ,421 | ,879 |
| 2 | ,950 | ,050 | ,845 | 1,055 |
| 3 | ,600 | ,112 | ,365 | ,835 |

## Pairwise Comparisons

Measure: image5.4

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | -,300* | ,105 | ,030 | -,576 | -,024 |
| | 3 | ,050 | ,135 | 1,000 | -,305 | ,405 |
| 2 | 1 | ,300* | ,105 | ,030 | ,024 | ,576 |
| | 3 | ,350* | ,109 | ,014 | ,063 | ,637 |
| 3 | 1 | -,050 | ,135 | 1,000 | -,405 | ,305 |
| | 2 | -,350* | ,109 | ,014 | -,637 | -,063 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 6.8

### Estimates

Measure: image6.8

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,700 | ,105 | ,480 | ,920 |
| 2 | ,250 | ,099 | ,042 | ,458 |
| 3 | ,300 | ,105 | ,080 | ,520 |

### Pairwise Comparisons

Measure: image6.8

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,450* | ,170 | ,047 | ,004 | ,896 |
| | 3 | ,400 | ,169 | ,085 | -,043 | ,843 |
| 2 | 1 | -,450* | ,170 | ,047 | -,896 | -,004 |
| | 3 | -,050 | ,114 | 1,000 | -,350 | ,250 |
| 3 | 1 | -,400 | ,169 | ,085 | -,843 | ,043 |
| | 2 | ,050 | ,114 | 1,000 | -,250 | ,350 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 6.12

### Estimates

Measure: image6.12

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,900 | ,069 | ,756 | 1,044 |
| 2 | ,450 | ,114 | ,211 | ,689 |
| 3 | ,400 | ,112 | ,165 | ,635 |

## Pairwise Comparisons

Measure: image6.12

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,450* | ,153 | ,026 | ,047 | ,853 |
| | 3 | ,500* | ,115 | ,001 | ,199 | ,801 |
| 2 | 1 | -,450* | ,153 | ,026 | -,853 | -,047 |
| | 3 | ,050 | ,153 | 1,000 | -,353 | ,453 |
| 3 | 1 | -,500* | ,115 | ,001 | -,801 | -,199 |
| | 2 | -,050 | ,153 | 1,000 | -,453 | ,353 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

### Image 7.7

## Descriptive Statistics

| | Mean | Std. Deviation | N |
|---|---|---|---|
| @1 | ,85 | ,366 | 20 |
| @2 | ,45 | ,510 | 20 |
| @3 | ,75 | ,444 | 20 |

## Pairwise Comparisons

Measure: image7.7

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,400* | ,134 | ,023 | ,049 | ,751 |
| | 3 | ,100 | ,124 | 1,000 | -,224 | ,424 |
| 2 | 1 | -,400* | ,134 | ,023 | -,751 | -,049 |
| | 3 | -,300 | ,147 | ,166 | -,686 | ,086 |
| 3 | 1 | -,100 | ,124 | 1,000 | -,424 | ,224 |
| | 2 | ,300 | ,147 | ,166 | -,086 | ,686 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 7.8

### Estimates

Measure: MSRimage7.8

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,800 | ,092 | ,608 | ,992 |
| 2 | ,300 | ,105 | ,080 | ,520 |
| 3 | ,250 | ,099 | ,042 | ,458 |

### Pairwise Comparisons

Measure: MSRimage7.8

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,500* | ,115 | ,001 | ,199 | ,801 |
| | 3 | ,550* | ,114 | ,000 | ,250 | ,850 |
| 2 | 1 | -,500* | ,115 | ,001 | -,801 | -,199 |
| | 3 | ,050 | ,114 | 1,000 | -,250 | ,350 |
| 3 | 1 | -,550* | ,114 | ,000 | -,850 | -,250 |
| | 2 | -,050 | ,114 | 1,000 | -,350 | ,250 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 8.9

### Estimates

Measure: image8.9

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,750 | ,099 | ,542 | ,958 |
| 2 | ,250 | ,099 | ,042 | ,458 |
| 3 | ,250 | ,099 | ,042 | ,458 |

## Pairwise Comparisons

Measure: image8.9

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,500* | ,154 | ,013 | ,096 | ,904 |
| | 3 | ,500* | ,154 | ,013 | ,096 | ,904 |
| 2 | 1 | -,500* | ,154 | ,013 | -,904 | -,096 |
| | 3 | ,000 | ,145 | 1,000 | -,381 | ,381 |
| 3 | 1 | -,500* | ,154 | ,013 | -,904 | -,096 |
| | 2 | ,000 | ,145 | 1,000 | -,381 | ,381 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

### Image 9.6

## Estimates

Measure: image9.6

| ProbeType | Mean | Std. Error | 95% Confidence Interval Lower Bound | Upper Bound |
|---|---|---|---|---|
| 1 | ,650 | ,109 | ,421 | ,879 |
| 2 | ,250 | ,099 | ,042 | ,458 |
| 3 | ,200 | ,092 | ,008 | ,392 |

## Pairwise Comparisons

Measure: image9.6

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] Lower Bound | Upper Bound |
|---|---|---|---|---|---|---|
| 1 | 2 | ,400* | ,152 | ,050 | ,001 | ,799 |
| | 3 | ,450* | ,153 | ,026 | ,047 | ,853 |
| 2 | 1 | -,400* | ,152 | ,050 | -,799 | -,001 |
| | 3 | ,050 | ,114 | 1,000 | -,250 | ,350 |
| 3 | 1 | -,450* | ,153 | ,026 | -,853 | -,047 |
| | 2 | -,050 | ,114 | 1,000 | -,350 | ,250 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 9.7

### Estimates

Measure: image9.7

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,650 | ,109 | ,421 | ,879 |
| 2 | ,350 | ,109 | ,121 | ,579 |
| 3 | ,300 | ,105 | ,080 | ,520 |

### Pairwise Comparisons

Measure: image9.7

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,300 | ,128 | ,089 | -,035 | ,635 |
| | 3 | ,350* | ,109 | ,014 | ,063 | ,637 |
| 2 | 1 | -,300 | ,128 | ,089 | -,635 | ,035 |
| | 3 | ,050 | ,153 | 1,000 | -,353 | ,453 |
| 3 | 1 | -,350* | ,109 | ,014 | -,637 | -,063 |
| | 2 | -,050 | ,153 | 1,000 | -,453 | ,353 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

## Image 12.7

### Estimates

Measure: image12.7

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,700 | ,105 | ,480 | ,920 |
| 2 | ,700 | ,105 | ,480 | ,920 |
| 3 | ,300 | ,105 | ,080 | ,520 |

## Pairwise Comparisons

Measure: image12.7

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,000 | ,103 | 1,000 | -,269 | ,269 |
| | 3 | ,400* | ,134 | ,023 | ,049 | ,751 |
| 2 | 1 | ,000 | ,103 | 1,000 | -,269 | ,269 |
| | 3 | ,400* | ,112 | ,006 | ,105 | ,695 |
| 3 | 1 | -,400* | ,134 | ,023 | -,751 | -,049 |
| | 2 | -,400* | ,112 | ,006 | -,695 | -,105 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

### Image 13.5

## Estimates

Measure: image13.5

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,350 | ,109 | ,121 | ,579 |
| 2 | ,050 | ,050 | -,055 | ,155 |
| 3 | ,100 | ,069 | -,044 | ,244 |

## Pairwise Comparisons

Measure: image13.5

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[a] | 95% Confidence Interval for Difference[a] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,300 | ,128 | ,089 | -,035 | ,635 |
| | 3 | ,250 | ,123 | ,169 | -,073 | ,573 |
| 2 | 1 | -,300 | ,128 | ,089 | -,635 | ,035 |
| | 3 | -,050 | ,088 | 1,000 | -,281 | ,181 |
| 3 | 1 | -,250 | ,123 | ,169 | -,573 | ,073 |
| | 2 | ,050 | ,088 | 1,000 | -,181 | ,281 |

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

## Image 14.2

### Estimates

Measure: MSRimage14.2

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,600 | ,112 | ,365 | ,835 |
| 2 | ,400 | ,112 | ,165 | ,635 |
| 3 | ,150 | ,082 | -,021 | ,321 |

### Pairwise Comparisons

Measure: MSRimage14.2

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,200 | ,156 | ,642 | -,208 | ,608 |
| | 3 | ,450* | ,153 | ,026 | ,047 | ,853 |
| 2 | 1 | -,200 | ,156 | ,642 | -,608 | ,208 |
| | 3 | ,250 | ,123 | ,169 | -,073 | ,573 |
| 3 | 1 | -,450* | ,153 | ,026 | -,853 | -,047 |
| | 2 | -,250 | ,123 | ,169 | -,573 | ,073 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

# APPENDIX 7

## Coding Regression Output by Probe

### 1. BLACK images

Curve Fit

**Model Description**

| | | |
|---|---|---|
| Model Name | | MOD_3 |
| Dependent Variable | 1 | MSRblack |
| Equation | 1 | Linear |
| Independent Variable | | CodingBlack |
| Constant | | Included |
| Variable Whose Values Label Observations in Plots | | Unspecified |

**Case Processing Summary**

| | N |
|---|---|
| Total Cases | 52 |
| Excluded Cases[a] | 0 |
| Forecasted Cases | 0 |
| Newly Created Cases | 0 |

a. Cases with a missing value in any variable are excluded from the analysis.

**Variable Processing Summary**

| | | Variables | |
|---|---|---|---|
| | | Dependent | Independent |
| | | MSRblack | CodingBlack |
| Number of Positive Values | | 52 | 39 |
| Number of Zeros | | 0 | 13 |
| Number of Negative Values | | 0 | 0 |
| Number of Missing Values | User-Missing | 0 | 0 |
| | System-Missing | 0 | 0 |

MSRblack

Linear

**Model Summary**

| R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|
| .690 | .476 | .465 | .117 |

The independent variable is CodingBlack.

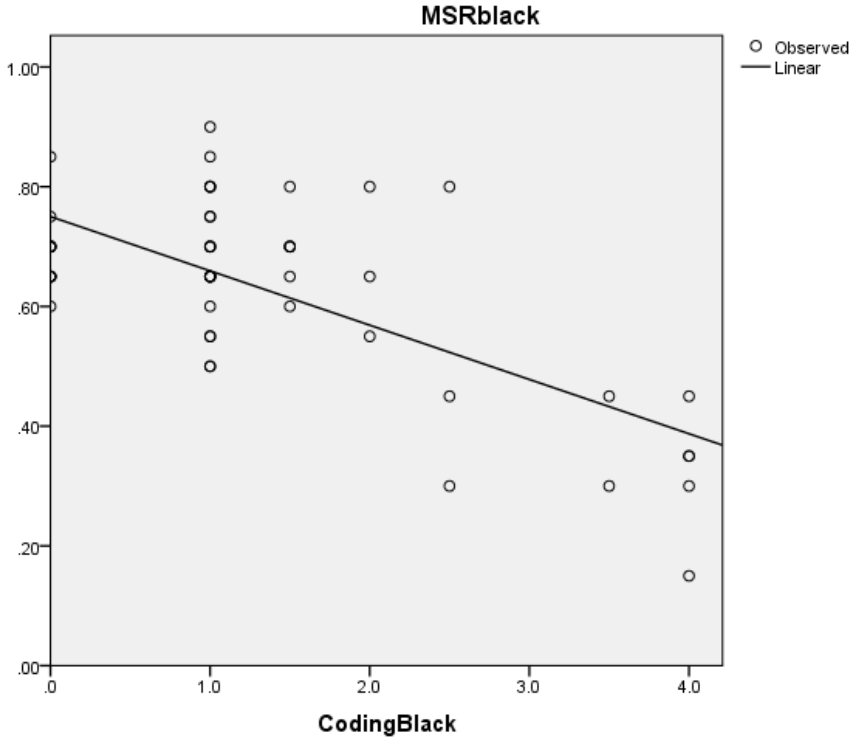
**ANOVA**

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | .624 | 1 | .624 | 45.357 | .000 |
| Residual | .688 | 50 | .014 | | |
| Total | 1.312 | 51 | | | |

The independent variable is CodingBlack.

**Coefficients**

| | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
|---|---|---|---|---|---|
| | B | Std. Error | Beta | | |
| CodingBlack | -.091 | .013 | -.690 | -6.735 | .000 |
| (Constant) | .750 | .024 | | 30.920 | .000 |



## 2. GOLF images

Curve Fit

**Model Description**

| | | |
|---|---|---|
| Model Name | | MOD_4 |
| Dependent Variable | 1 | MSRgolf |
| Equation | 1 | Linear |
| Independent Variable | | CodingGolf |
| Constant | | Included |
| Variable Whose Values Label Observations in Plots | | Unspecified |

**Case Processing Summary**

| | N |
|---|---|
| Total Cases | 52 |
| Excluded Cases[a] | 0 |
| Forecasted Cases | 0 |
| Newly Created Cases | 0 |

a. Cases with a missing value in any variable are excluded from the analysis.

**Variable Processing Summary**

| | | Variables | |
|---|---|---|---|
| | | Dependent | Independent |
| | | MSRgolf | CodingGolf |
| Number of Positive Values | | 52 | 39 |
| Number of Zeros | | 0 | 13 |
| Number of Negative Values | | 0 | 0 |
| Number of Missing Values | User-Missing | 0 | 0 |
| | System-Missing | 0 | 0 |

MSRgolf

Linear

**Model Summary**

| R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|
| .546 | .298 | .284 | .180 |

The independent variable is CodingGolf.

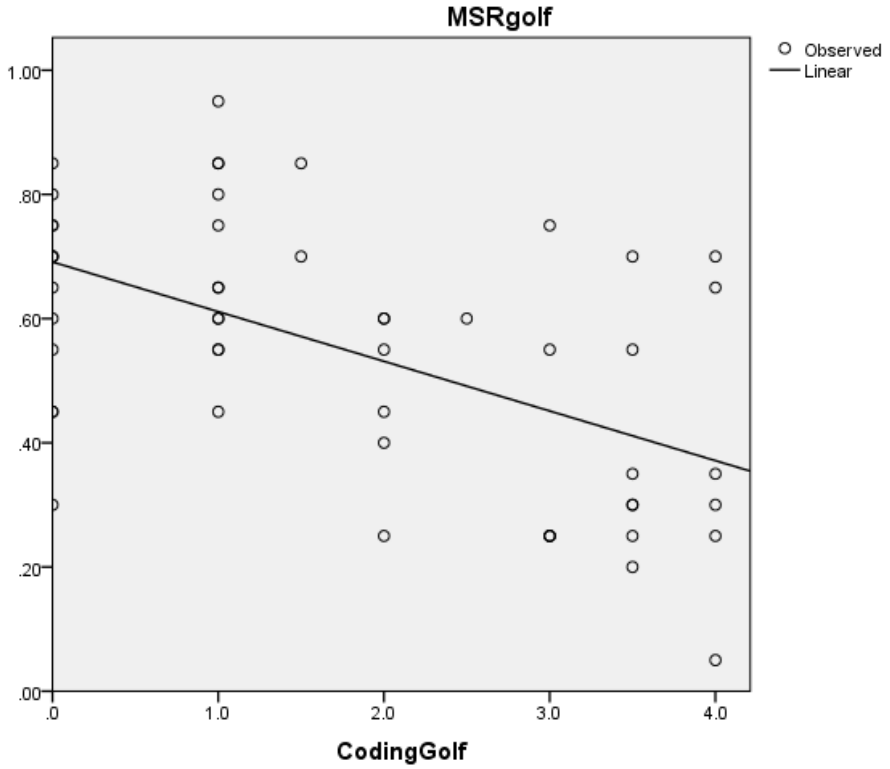
**ANOVA**

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | .685 | 1 | .685 | 21.231 | .000 |
| Residual | 1.614 | 50 | .032 | | |
| Total | 2.300 | 51 | | | |

The independent variable is CodingGolf.

**Coefficients**

| | Unstandardized Coefficients | | Standardized Coefficients | | |
|---|---|---|---|---|---|
| | B | Std. Error | Beta | t | Sig. |
| CodingGolf | -.080 | .017 | -.546 | -4.608 | .000 |
| (Constant) | .691 | .040 | | 17.362 | .000 |



## 3. WHITE images

Curve Fit

**Model Description**

| | | |
|---|---|---|
| Model Name | | MOD_5 |
| Dependent Variable | 1 | MSRwhite |
| Equation | 1 | Linear |
| Independent Variable | | CodingWhite |
| Constant | | Included |
| Variable Whose Values Label Observations in Plots | | Unspecified |

**Case Processing Summary**

| | N |
|---|---|
| Total Cases | 52 |
| Excluded Cases[a] | 0 |
| Forecasted Cases | 0 |
| Newly Created Cases | 0 |

a. Cases with a missing value in any variable are excluded from the analysis.

**Variable Processing Summary**

| | | Variables | |
|---|---|---|---|
| | | Dependent | Independent |
| | | MSRwhite | CodingWhite |
| Number of Positive Values | | 52 | 39 |
| Number of Zeros | | 0 | 13 |
| Number of Negative Values | | 0 | 0 |
| Number of Missing Values | User-Missing | 0 | 0 |
| | System-Missing | 0 | 0 |

MSRwhite

Linear

**Model Summary**

| R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|
| .661 | .437 | .426 | .151 |

The independent variable is CodingWhite.

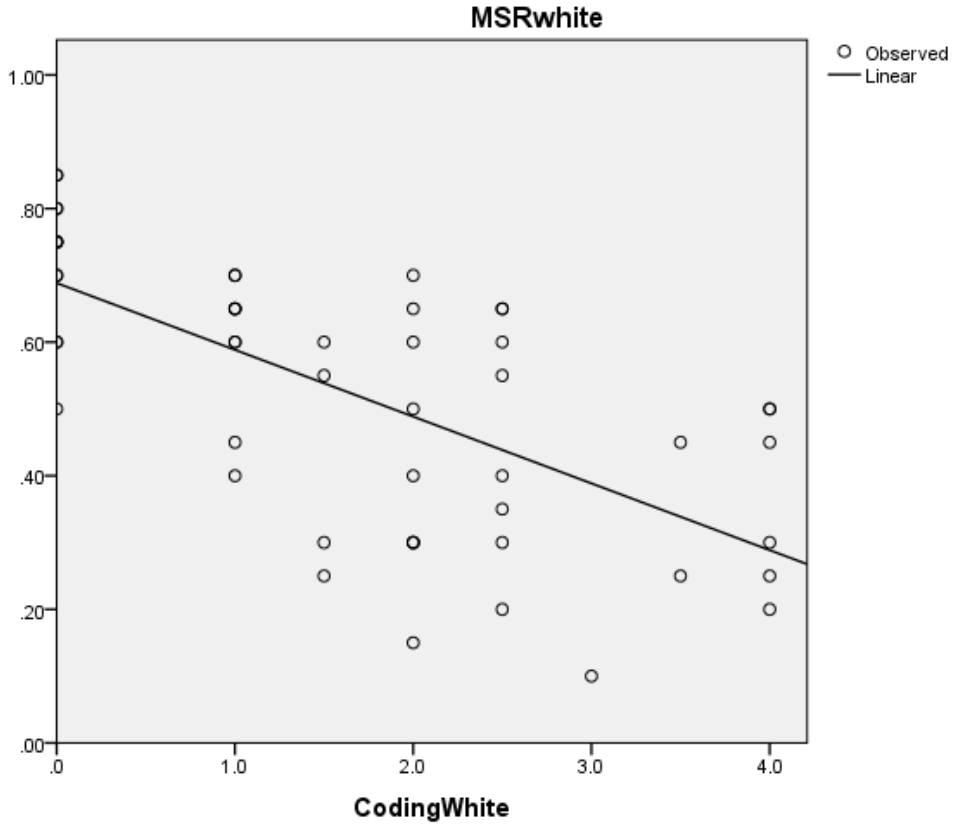
**ANOVA**

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Regression | .879 | 1 | .879 | 38.773 | .000 |
| Residual | 1.133 | 50 | .023 | | |
| Total | 2.012 | 51 | | | |

The independent variable is CodingWhite.

**Coefficients**

| | Unstandardized Coefficients | | Standardized Coefficients | | |
|---|---|---|---|---|---|
| | B | Std. Error | Beta | t | Sig. |
| CodingWhite | -.100 | .016 | -.661 | -6.227 | .000 |
| (Constant) | .688 | .034 | | 20.237 | .000 |



# APPENDIX 8

## Two-way ANOVA Output with Groups and ProbeType

```
GET DATA
  /TYPE=XLSX
  /FILE='\\Client\H$\Desktop\MSRprobetype.xlsx'
  /SHEET=name 'Two-way ANOVA data setup'
  /CELLRANGE=FULL
  /READNAMES=ON
  /LEADINGSPACES IGNORE=YES
  /DATATYPEMIN PERCENTAGE=95.0
  /HIDDEN IGNORE=YES.
EXECUTE.
DATASET NAME DataSet1 WINDOW=FRONT.
UNIANOVA MSRimage BY ProbeType Group
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /POSTHOC=ProbeType Group(TUKEY BONFERRONI)
  /PLOT=PROFILE(Group*ProbeType) TYPE=LINE ERRORBAR=NO MEANREFERENCE=NO YAXIS=AUTO
  /EMMEANS=TABLES(ProbeType) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(Group) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(ProbeType*Group)
  /PRINT ETASQ DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /DESIGN=ProbeType Group ProbeType*Group.
```

### Univariate Analysis of Variance

| Between-Subjects Factors | | N |
|---|---|---|
| ProbeType | 1 | 52 |
| | 2 | 52 |
| | 3 | 52 |
| Group | 1 | 24 |
| | 2 | 48 |
| | 3 | 48 |
| | 4 | 36 |

**Descriptive Statistics**

Dependent Variable: MSRimage

| ProbeType | Group | Mean | Std. Deviation | N |
|---|---|---|---|---|
| 1 | 1 | ,7250 | ,06547 | 8 |
| | 2 | ,6156 | ,16805 | 16 |
| | 3 | ,6406 | ,17342 | 16 |
| | 4 | ,5667 | ,16002 | 12 |
| | Total | ,6288 | ,16038 | 52 |
| 2 | 1 | ,6000 | ,08452 | 8 |
| | 2 | ,6156 | ,23785 | 16 |
| | 3 | ,4656 | ,18949 | 16 |
| | 4 | ,5333 | ,24433 | 12 |
| | Total | ,5481 | ,21235 | 52 |
| 3 | 1 | ,5875 | ,17879 | 8 |
| | 2 | ,5812 | ,14361 | 16 |
| | 3 | ,4563 | ,19822 | 16 |
| | 4 | ,4833 | ,25436 | 12 |
| | Total | ,5212 | ,19861 | 52 |
| Total | 1 | ,6375 | ,13126 | 24 |
| | 2 | ,6042 | ,18417 | 48 |
| | 3 | ,5208 | ,20234 | 48 |
| | 4 | ,5278 | ,21988 | 36 |
| | Total | ,5660 | ,19593 | 156 |

**Tests of Between-Subjects Effects**

Dependent Variable: MSRimage

| Source | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Corrected Model | ,836[a] | 11 | ,076 | 2,139 | ,021 | ,140 |
| Intercept | 47,208 | 1 | 47,208 | 1329,195 | ,000 | ,902 |
| ProbeType | ,316 | 2 | ,158 | 4,443 | ,013 | ,058 |
| Group | ,343 | 3 | ,114 | 3,220 | ,025 | ,063 |
| ProbeType * Group | ,166 | 6 | ,028 | ,778 | ,589 | ,031 |
| Error | 5,114 | 144 | ,036 | | | |
| Total | 55,930 | 156 | | | | |
| Corrected Total | 5,950 | 155 | | | | |

a. R Squared = ,140 (Adjusted R Squared = ,075)

## Estimated Marginal Means

## 1. ProbeType

**Estimates**

Dependent Variable: MSRimage

| ProbeType | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,637 | ,027 | ,583 | ,691 |
| 2 | ,554 | ,027 | ,500 | ,607 |
| 3 | ,527 | ,027 | ,473 | ,581 |

**Pairwise Comparisons**

Dependent Variable: MSRimage

| (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig.[b] | 95% Confidence Interval for Difference[b] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,083 | ,038 | ,096 | -,010 | ,177 |
| | 3 | ,110* | ,038 | ,015 | ,017 | ,203 |
| 2 | 1 | -,083 | ,038 | ,096 | -,177 | ,010 |
| | 3 | ,027 | ,038 | 1,000 | -,067 | ,120 |
| 3 | 1 | -,110* | ,038 | ,015 | -,203 | -,017 |
| | 2 | -,027 | ,038 | 1,000 | -,120 | ,067 |

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Bonferroni.

**Univariate Tests**

Dependent Variable: MSRimage

| | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Contrast | ,316 | 2 | ,158 | 4,443 | ,013 | ,058 |
| Error | 5,114 | 144 | ,036 | | | |

The F tests the effect of ProbeType. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

## 2. Group

**Estimates**

Dependent Variable: MSRimage

| Group | Mean | Std. Error | 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 1 | ,638 | ,038 | ,561 | ,714 |
| 2 | ,604 | ,027 | ,550 | ,658 |
| 3 | ,521 | ,027 | ,467 | ,575 |
| 4 | ,528 | ,031 | ,466 | ,590 |

**Pairwise Comparisons**

Dependent Variable: MSRimage

| (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig.[a] | 95% Confidence Interval for Difference[a] | |
|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound |
| 1 | 2 | ,033 | ,047 | 1,000 | -,093 | ,159 |
| | 3 | ,117 | ,047 | ,087 | -,009 | ,243 |
| | 4 | ,110 | ,050 | ,172 | -,023 | ,243 |
| 2 | 1 | -,033 | ,047 | 1,000 | -,159 | ,093 |
| | 3 | ,083 | ,038 | ,192 | -,020 | ,186 |
| | 4 | ,076 | ,042 | ,408 | -,035 | ,188 |
| 3 | 1 | -,117 | ,047 | ,087 | -,243 | ,009 |
| | 2 | -,083 | ,038 | ,192 | -,186 | ,020 |
| | 4 | -,007 | ,042 | 1,000 | -,118 | ,104 |
| 4 | 1 | -,110 | ,050 | ,172 | -,243 | ,023 |
| | 2 | -,076 | ,042 | ,408 | -,188 | ,035 |
| | 3 | ,007 | ,042 | 1,000 | -,104 | ,118 |

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

**Univariate Tests**

Dependent Variable: MSRimage

| | Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|
| Contrast | ,343 | 3 | ,114 | 3,220 | ,025 | ,063 |
| Error | 5,114 | 144 | ,036 | | | |

The F tests the effect of Group. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

# Post Hoc Tests

## Group

**Multiple Comparisons**

Dependent Variable: MSRimage

| | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Tukey HSD | 1 | 2 | ,0333 | ,04711 | ,894 | -,0891 | ,1558 |
| | | 3 | ,1167 | ,04711 | ,068 | -,0058 | ,2391 |
| | | 4 | ,1097 | ,04966 | ,126 | -,0194 | ,2388 |
| | 2 | 1 | -,0333 | ,04711 | ,894 | -,1558 | ,0891 |
| | | 3 | ,0833 | ,03847 | ,138 | -,0167 | ,1833 |
| | | 4 | ,0764 | ,04155 | ,260 | -,0316 | ,1844 |
| | 3 | 1 | -,1167 | ,04711 | ,068 | -,2391 | ,0058 |
| | | 2 | -,0833 | ,03847 | ,138 | -,1833 | ,0167 |
| | | 4 | -,0069 | ,04155 | ,998 | -,1149 | ,1011 |
| | 4 | 1 | -,1097 | ,04966 | ,126 | -,2388 | ,0194 |
| | | 2 | -,0764 | ,04155 | ,260 | -,1844 | ,0316 |
| | | 3 | ,0069 | ,04155 | ,998 | -,1011 | ,1149 |
| Bonferroni | 1 | 2 | ,0333 | ,04711 | 1,000 | -,0927 | ,1594 |
| | | 3 | ,1167 | ,04711 | ,087 | -,0094 | ,2427 |
| | | 4 | ,1097 | ,04966 | ,172 | -,0231 | ,2426 |
| | 2 | 1 | -,0333 | ,04711 | 1,000 | -,1594 | ,0927 |
| | | 3 | ,0833 | ,03847 | ,192 | -,0196 | ,1862 |
| | | 4 | ,0764 | ,04155 | ,408 | -,0348 | ,1875 |
| | 3 | 1 | -,1167 | ,04711 | ,087 | -,2427 | ,0094 |
| | | 2 | -,0833 | ,03847 | ,192 | -,1862 | ,0196 |
| | | 4 | -,0069 | ,04155 | 1,000 | -,1181 | ,1042 |
| | 4 | 1 | -,1097 | ,04966 | ,172 | -,2426 | ,0231 |
| | | 2 | -,0764 | ,04155 | ,408 | -,1875 | ,0348 |
| | | 3 | ,0069 | ,04155 | 1,000 | -,1042 | ,1181 |

.

## Homogeneous Subsets

**MSRimage**

| | Group | N | Subset 1 | Subset 2 |
|---|---|---|---|---|
| Tukey HSD[a,b,c] | 3 | 48 | ,5208 | |
| | 4 | 36 | ,5278 | ,5278 |
| | 2 | 48 | ,6042 | ,6042 |
| | 1 | 24 | | ,6375 |
| | Sig. | | ,243 | ,069 |

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = ,036.

a. Uses Harmonic Mean Sample Size = 36,000.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c. Alpha = ,05.

## ProbeType

**Multiple Comparisons**

Dependent Variable: MSRimage

| | (I) ProbeType | (J) ProbeType | Mean Difference (I-J) | Std. Error | Sig. | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| | | | | | | Lower Bound | Upper Bound |
| Tukey HSD | 1 | 2 | ,0808 | ,03696 | ,077 | -,0068 | ,1683 |
| | | 3 | ,1077* | ,03696 | ,011 | ,0202 | ,1952 |
| | 2 | 1 | -,0808 | ,03696 | ,077 | -,1683 | ,0068 |
| | | 3 | ,0269 | ,03696 | ,747 | -,0606 | ,1145 |
| | 3 | 1 | -,1077* | ,03696 | ,011 | -,1952 | -,0202 |
| | | 2 | -,0269 | ,03696 | ,747 | -,1145 | ,0606 |
| Bonferroni | 1 | 2 | ,0808 | ,03696 | ,091 | -,0088 | ,1703 |
| | | 3 | ,1077* | ,03696 | ,012 | ,0182 | ,1972 |
| | 2 | 1 | -,0808 | ,03696 | ,091 | -,1703 | ,0088 |
| | | 3 | ,0269 | ,03696 | 1,000 | -,0626 | ,1164 |
| | 3 | 1 | -,1077* | ,03696 | ,012 | -,1972 | -,0182 |
| | | 2 | -,0269 | ,03696 | 1,000 | -,1164 | ,0626 |

Based on observed means.
The error term is Mean Square(Error) = ,036.

*. The mean difference is significant at the ,05 level.

## Homogeneous Subsets

**MSRimage**

| | ProbeType | N | Subset 1 | Subset 2 |
|---|---|---|---|---|
| Tukey HSD[a,b] | 3 | 52 | ,5212 | |
| | 2 | 52 | ,5481 | ,5481 |
| | 1 | 52 | | ,6288 |
| | Sig. | | ,747 | ,077 |

Means for groups in homogeneous subsets are displayed.
Based on observed means.
The error term is Mean Square(Error) = ,036.

a. Uses Harmonic Mean Sample Size = 52,000.

b. Alpha = ,05.