Technische Universität Berlin
Fakultät IV - Elektrotechnik und Informatik
Institut für Technische Informatik und Mikroelektronik
Dept. Computational Psychology

Bachelor Thesis

Measuring the perceptual scales of brightness assimilation effect

---

submitted by
Navdha Jain
in fulfillment of the requirements for the degree
"Bachelor of Science" (B. Sc.)
in the field of Computer Science

First examiner: Prof. Dr. Marianne Maertens
Second examiner: Prof. Dr. Felix Wichmann
Supervisor: Dr. Joris Vincent

29. April 2024

SELBSTÄNDIGKEITSERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den 29. April 2024

Navdha

Navdha Jain

ABSTRACT

This paper investigates human brightness perception using maximum likelihood conjoint measurement (MLCM) to estimate the perceived brightness of assimilation displays specifically bullseye and checkerboard displays. The target luminance and target placement (either white or black background) are varied to determine these scales across the entire luminance range. Two sets of scales corresponding to the two different contexts in which the targets are placed, are drawn and the strength of the effect can also be determined with the help of these scales. The scales of checkerboard images show stronger or equal assimilation compared to the bullseye scales. Less inter-observer variability is observed for checkerboard scales as the majority shows assimilation. When the observers show both assimilation and contrast, it is more common to show assimilation in the intermediate luminance targets and contrast on the extremes. We compare these assimilation scales to previously estimated perceptual scales of White's effect (Vincent, Maertens, and Aguilar, in press), noting greater variability among observers and stimuli in assimilation scales. However, limitations prevent definitive conclusions about the mechanisms underlying surround context effects, emphasizing the need for consistent experimental conditions.

ZUSAMMENFASSUNG

Diese Arbeit untersucht die menschliche Helligkeitswahrnehmung mithilfe von Maximum Likelihood Conjoint Measurement (MLCM), um die wahrgenommene Helligkeit von Assimilation-Displays wie Bullseye- und Checkerboard-Displays zu schätzen. Die Target-Luminanz und die Target-Platzierung (entweder weißer oder schwarzer Hintergrund) werden variiert, um diese Skalen über den gesamten Luminanzbereich zu bestimmen. Zwei Gruppen von Skalen, die den beiden unterschiedlichen Kontexten entsprechen, in denen die Targets platziert sind, werden gezeichnet, und die Stärke des Effekts kann ebenfalls mit Hilfe dieser Skalen bestimmt werden. Die Skalen der Checkboard-Bilder zeigen eine stärkere oder gleiche Assimilation im Vergleich zu den Bullseye-Skalen. Bei Checkerboard-Skalen wird eine geringere Variabilität zwischen den Beobachtern beobachtet, da die Mehrheit Assimilation zeigt. Wenn die Beobachter sowohl Assimilation als auch Kontrast zeigen, ist die Assimilation bei den Targets mit mittlerer Luminanz und der Kontrast bei den Extremen häufi-

ger. Wir vergleichen diese Assimilations-Skalen mit zuvor geschätzten perzeptuellen Skalen des White-Effekts (Vincent et al., in press) und stellen eine größere Variabilität zwischen Beobachtern und Stimuli in den Assimilations-Skalen fest. Die Einschränkungen verhindern jedoch definitive Schlussfolgerungen über die Mechanismen, die den Surround-Kontext-Effekten zugrunde liegen, was die Notwendigkeit konsistenter experimenteller Bedingungen unterstreicht.

ACKNOWLEDGMENTS

I would like to begin with expressing my gratitude to my supervisor Dr. Joris Vincent for his valuable feedback, comments and time throughout my journey of writing my thesis. I really appreciate how he would push me to improve certain things while also supporting me and giving me encouragement. I would like to thank Prof. Dr. Marianne Maertens for her support and feedback during the experimentation period and for giving me the opportunity to write my thesis in her department. I would also like to thank Prof. Dr. Felix Wichmann for agreeing to be my second examiner. Finally, I would like to thank my family who encouraged me and supported me through this period. I would like to thank all my friends who helped me remain focused and the participants who agreed to take out their time and made this thesis possible.

CONTENTS

1 Introduction 1
2 Method 9
    2.1 Participants 9
    2.2 Stimuli 9
    2.3 Apparatus 12
    2.4 Design and Procedure 12
    2.5 Scale estimation 13
3 Results 16
4 Discussion 23
References 27

1

INTRODUCTION

Human perception of the surrounding world is grounded in measurable physical quantities, yet often diverges from them. This distinction between the physical and perceptual quantities is studied in the field of vision.

The three important physical quantities, illuminance, reflectance, and luminance are shown in Figure 1.1. The incident light, also known as the illuminance, is defined as the light falling on the surface of an object. It can vary as the light falling on an object depends on factors like the environment, the lighting conditions, shadows, etc. Reflectance is the proportion of incident light that bounces off the surface of an object. It is a physical property of an object and it remains unvaried. Luminance is the total light that falls on our eye; it is the product of the light falling on the object (illuminance) and the light reflected by the object (reflectance) (Adelson, 2000).

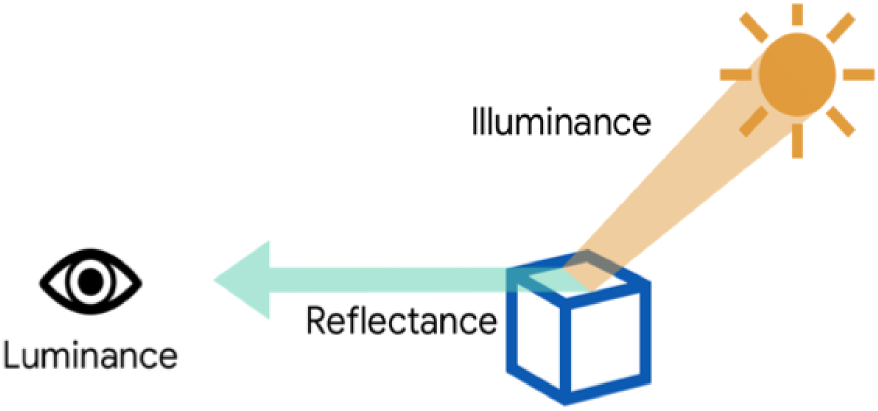


Figure 1.1: Illuminance (I) is the light falling on an object; Reflectance (R) is a property of a surface that quantifies how much incident light is reflected by that surface; Luminance (L) is the amount light that comes to the eye from a surface. The relationship between these three physical quantities is given as L = R x I.

While illuminance, reflectance and luminance are physical quantities that can be measured by devices, they cannot be directly perceived by us. Brightness is how we perceive the intensity of light or objects visually; it is the perceptual correlate of luminance. Natural daylight (illuminance) changes both in colour and in intensity (by more than three orders of magnitude), thus for the same object (having fixed reflectance), the luminance changes a lot as well. Despite changes in illuminance (and as a result in luminance), our perception of the object's reflectance (lightness) remains consistent, illustrating that our perception goes beyond simply relying on luminance (Kingdom, 2011).

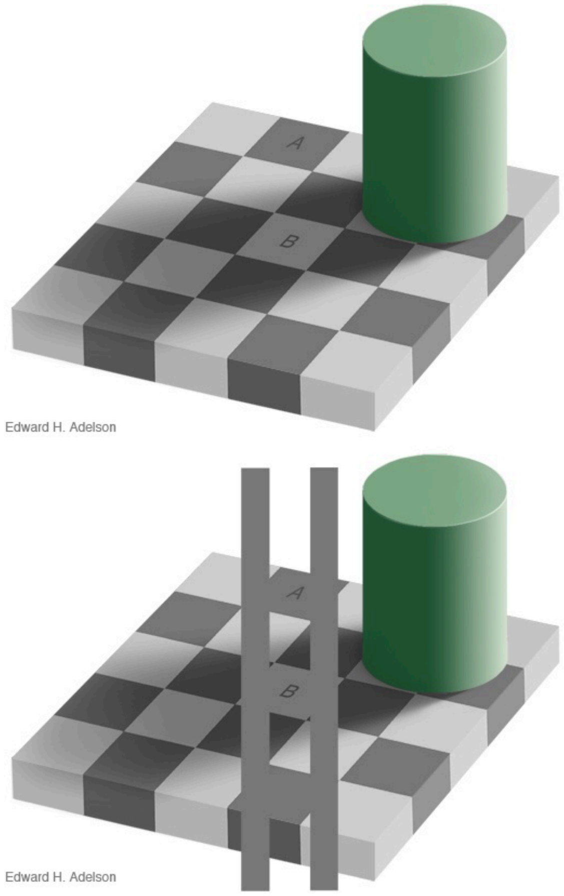
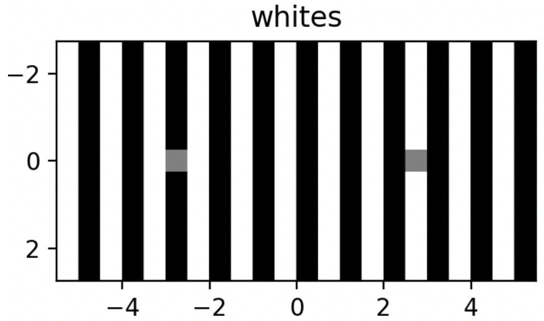


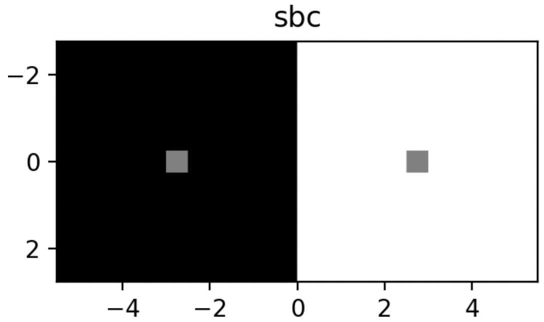
Figure 1.2: Adelson's Check-Shadow display (upper figure) - tiles A and B have the same physical luminance (lower figure) but the difference in perceived illuminance causes the difference in perceived reflectance.

Apart from changes in illuminance, surround context also affects the perceived lightness of a surface as seen in Adelson's Check-Shadow display (Figure 1.2). Adelson's Check-Shadow display depicts a checkerboard with a green cylinder casting a shadow caused by a single light source. The tiles A and B on the checkerboard are physically the same, meaning that they are the same shade of gray. Since the tiles are identical, the luminance is the same. If our visual system calculated perceived reflectance (lightness) only with the help of physical luminance, both the tiles would appear equally light to us. As seen in the upper Figure 1.2, this is not the case as tile B appears significantly lighter than tile A. Since it is a 2D image on a screen, there is no reflectance or illuminance but there is still perceived illuminance and perceived reflectance. The perceived illuminance on each tile is different because tile B is in perceived shadow while the other tile is in direct perceived illumination. To compensate for this difference in perceived illuminance, the perceived reflectance, i.e. lightness of both tiles must differ in order to keep their product, i.e. physical luminance constant.

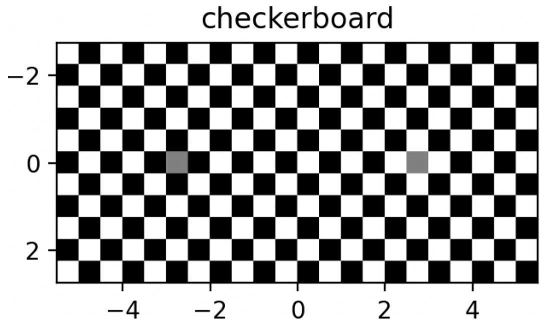
Even though Adelson's Check-Shadow display (Figure 1.2) is a 2D image on a screen, it allows for 3D interpretation. Some surround context effects are much simpler and easy to manipulate because they are just 2D images with no 3D interpretation. Figure 1.3 shows three surround context effects: (a) White's Effect (White, 1979), (b) the classical simultaneous brightness contrast (Adelson, 2000) and (c) assimilation displays (Bindman and Chubb, 2004; De Valois and De Valois, 1988; Gilchrist et al., 1999). Since there is no perceived illumination/shadow in these displays, lightness (perceived reflectance) and brightness (perceived luminance) mean the same. These displays highlight the difference between physical and perceptual descriptions by demonstrating that the perceived brightness of the target patches is not solely determined by their physical luminance.



(a) White's effect: a gray target patch on the darker collinear bar appears brighter than an identical target patch on the lighter collinear bar.



(b) Simultaneous brightness contrast: gray target patch bordered by darker surround context appears brighter than an identical gray target patch bordered by lighter surround context.



(c) Assimilation display: gray target patch surrounded by lighter immediate surround context appears brighter than an identical gray target surrounded by darker immediate surround context.

Figure 1.3: Three surround context effects: (a) White's Effect, (b) the classical simultaneous brightness contrast and (c) Checkerboard display (example of assimilation display)

The two gray target patches in these effects have the same physical luminance. However, most people perceive one target patch brighter than the other. If a gray target patch surrounded by a darker surround context appears brighter, then it is an example of a simultaneous brightness contrast (Figure 1.3(b)). The contrast effect emphasises how the surround context of the target patch can make the patch 'stand out' from its surroundings. Assimilation displays (Figure 1.3(c)), on the other hand, cause an opposite effect. Assimilation effect demonstrates how squares surrounding the target patch can make the patch assimilate or 'blend' with its surroundings. While contrast is an increase in the perceived differences between the target patch and the surround context, assimilation is a reduction (de Weert and van Kruysbergen, 1997). We see these opposing effects even though the four edges of the target patches in both assimilation and contrast are surrounded by either black or white squares, meaning that they share same local contrast and consistent local borders. We can therefore see how altering the surround context relative to the target patches can lead to substantial changes in our perception of the target patches' brightness.

Even when assimilation and contrast appear as opposite effects, they could be derived from the same mechanism (Helson, 1963 as reviewed in Barkan, Spitzer, and Einav, 2008). An example supporting this suspicion is the White's effect (Figure 1.3(a)). White's effect can be explained both as the targets assimilating with the flanking bars, or as contrast with the collinear bars. The gray target patch appears brighter when surrounded by brighter flanking bars, which is an assimilation effect. But this target patch is also placed on a darker collinear bar and when interpreted this way, it appearing brighter is a contrast effect. Depending on how it is interpreted, White's effect can be explained as an instance of assimilation effect or contrast effect (Kingdom, 2011). This suggests that 'contrast' and 'assimilation' are simply terms used to describe the direction of surround context effects, rather than indications of different underlying mechanisms (Kingdom, 2011). Many theories have tried to figure these mechanisms but according to Kingdom (2011), no theory has fully flowered yet. The question if contrast and assimilation effects are generated from the same mechanism in our brains remains unresolved. If we can better characterise the relationship between target luminance, and target brightness, in assimilation and contrast displays, we may have more detailed data to distinguish them.

Matching has been the most widely used experimental paradigm to study brightness perception. The task of matching requires the observer to adjust the physical luminance of a test stimulus, known as probe until it matches the target stimulus (reference) perceptually. This method is useful in comparing the luminance of two patches in relation to each other but it fails to directly measure a brightness value from a physical luminance value. Furthermore, since matching takes more time and effort, it is always limited to a small range of luminance values. Vincent et al. (in press) argued that scaling methods, in particular Maximum Likelihood Conjoint Measurement (MLCM) promise reliable estimation of the relationship between luminance and perceived brightness (perceptual scales), under some assumptions. MLCM can uncover the nonlinearities inherent in the relationship between perceived brightness and luminance (Vincent et al., in press). It also allows us to investigate the full luminance range, including the extreme luminance values which are never accounted for in matching experiments. It is therefore interesting to move away from a certain tested range of luminance and see if there is any difference in results for the less tested luminance values.

MLCM can be applied to measure perceptual scales for surround context effects such as those in Figure 1.3. Vincent et al. (in press) already measured perceptual scales for White's Effect using MLCM (Figure 1.4). First they generated stimulus configurations by altering both luminance values and the surround context. Participants in the experiment were presented with pairs of stimulus configurations and were asked to judge which of the two appeared brighter (Maloney and Knoblauch, 2020). Observers' brightness judgements were collected through repeated trials as choice probabilities. Finally, MLCM estimated the complete perceptual scale by maximising the likelihood of a set of scale values based on the choice probabilities. A detailed review of MLCM can be found in Maloney and Knoblauch (2020). The scales of White's effect (Figure 1.4) that were constructed in Vincent et al. (in review) show that the target collinear with the the black bar was consistently perceived as having the same or higher perceived brightness compared to the target collinear with the white bar. Across all observers, the direction of the effect remained consistent (in direction of White's effect) but the magnitude of the effect, i.e. the difference in perceived brightness varied. The effect was the strongest for the intermediate luminance targets and it decreased towards the extremes for many observers. MLCM could therefore help unveil these variations across observers and luminance values for White's effect.

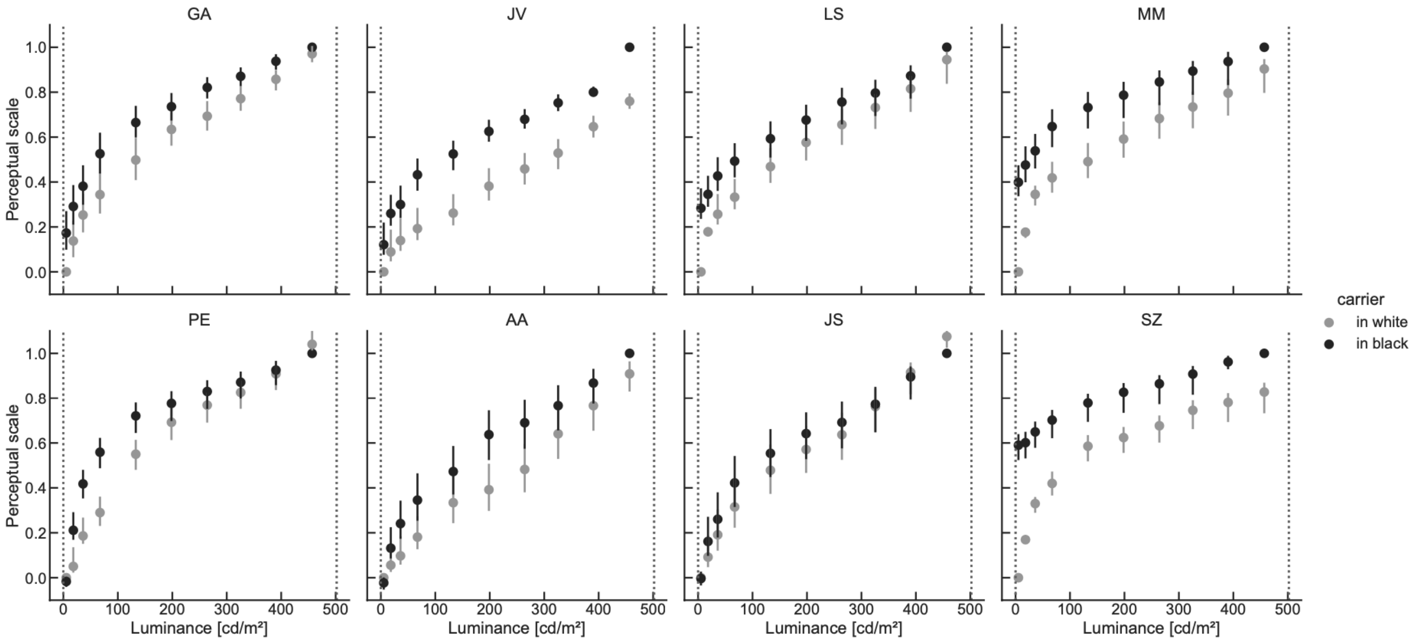


Figure 1.4: Perceptual scales of White's Effect measured using MLCM. Each panel depicts the perceptual scales from an observer (y-axis) as a function of target luminance (x-axis) and target placement (color). Both scales correspond to the respective target placement. The vertical distance between them reflects the strength of the illusion. The target collinear with the black bar was consistently perceived as having the same or higher brightness compared to the target collinear with the white bar. The magnitude of the effect varied individually. Dashed vertical lines indicate the luminance of the black and white carrier gratings in White's stimulus (taken from Vincent et al., in press)

The primary goal of this thesis will be to measure perceptual scales of assimilation displays, specifically examining checkerboard images (De Valois and De Valois, 1988) and bullseye displays (Bindman and Chubb, 2004) using the MLCM method with the similar design as Vincent et al. (in press). I intend to compare the perceptual scales derived from both the checkerboard and bullseye displays, aiming to identify any intriguing patterns or disparities between them. I will also be examining the scales across the entire luminance range to uncover patterns at the extreme luminance values. It has been shown that in White's effect both brightness contrast and assimilation mechanism play a role (Blakeslee and McCourt, 2004). Therefore, it might be interesting to look at any possible correlations between the perceptual scales of the assimilation displays and those of White's effect measured in the paper by Vincent et al. (in press). If they differ significantly, it might be useful to consider that both these effects cannot be explained with the help of the same mechanism. Conversely, if both scales show similarities, there might be a possibility that both effects can be derived from the same mechanism.

The measurement of perceptual scales using MLCM not only allows us to look into the relationship between luminance and brightness, but also facilitates comparisons of diverse perceptions on a standardised scale. As a result, we take a step closer to understanding the complex processes involved in brightness perception, thereby being able to look into the black box of human visual system.

2

METHOD

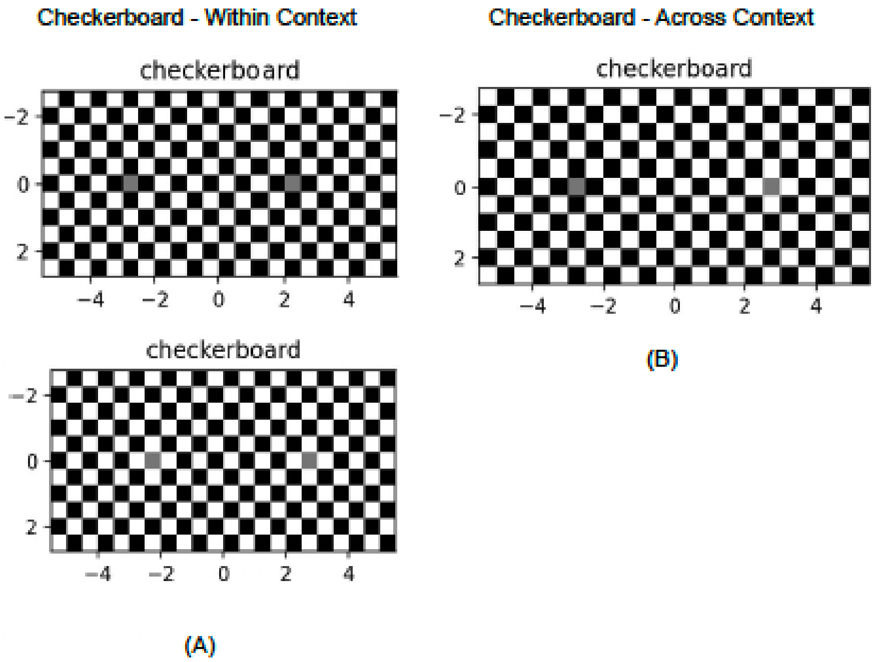
2.1 PARTICIPANTS

3 expert observers (the 1 author and 2 affiliates of the Computational Psychology department at TU Berlin) and 6 volunteers from the author's personal environment naive to the purpose of the experiment took part in the experiment. All observers had normal or corrected to normal visual ability. Naive observers were reimbursed for participation. To conduct this experiment, obtaining approval from the ethics committee was necessary. Additionally, every participant had to sign a declaration of consent before partaking in the experiment.

2.2 STIMULI

Two stimuli were used to produce assimilation effect in the experiment — checkerboard contrast illusion (De Valois and De Valois, 1988; Gilchrist et al., 1999) and bullseye images (Bindman and Chubb, 2004). Both the checkerboard and bullseye stimuli were generated with the help of the python package 'stimupy' (Schmittwilken, Maertens, and Vincent, 2023). Bullseye display had two targets placed in the centre which were surrounded by 5 alternating black and white square frames. If the first frame was black, then the second would always be white since they alternate. The checkerboard image had a 2-D checkerboard consisting of black and white squares which were all the same size as the targets. The checkerboard had 11 rows and 22 columns. On the top and bottom left, there was a white square, and on the top and bottom right, there was a black square. The target patches on both stimuli were 0.5 x 0.5 degrees and the luminance of the background on which the stimuli were presented was 75 cd/m^2.

Two stimulus dimensions were manipulated in the experiments: the target luminance and the target placement or context. The luminance of each target patch could take one of 10 possible values: 2.5, 27.25, 52, 76.75, 101.5, 126, 150.75, 175.5, 200.25, 225 cd/m^2. There were 2 ways to place a target: either around a black surround context or a white surround context. As a result, there were 20 possible target types. The target placement is referred to 'within' context (Figure 2.1 (A) and Figure 2.2 (C)) when both targets are completely abutted by either white or by black squares in checkerboard images. In bullseye displays both the targets are surrounded by bands that alternate outward either from black to white or white to black. 'Across' context (Figure 2.1 (B) and Figure 2.2 (D)) for checkerboard is when one target is being abutted by black squares while the other by white and for bullseye when one target is surrounded by bands that alternate outward from black to white while the bands surrounding the other target alternate outward from white to black.



(A)

(B)

Figure 2.1: (A) within context checkerboard contrast display : the surround context of the target is kept the same (top is black and below is white), but the luminance of the target is changed (B) across context checkerboard contrast display : One target is surrounded by the white local surround context while the other is surrounded by black. Here the target luminance can be changed or kept the same.

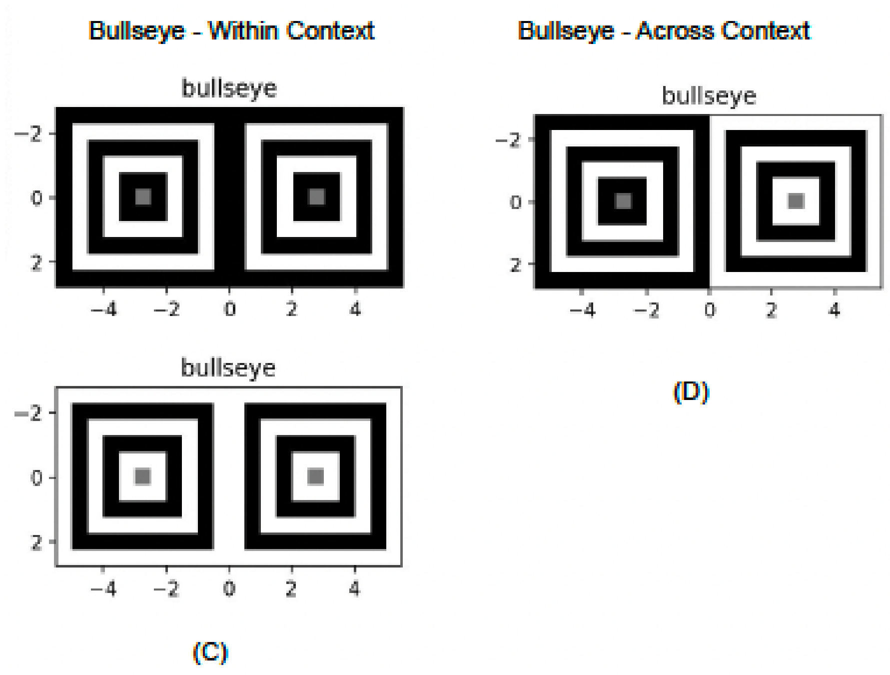
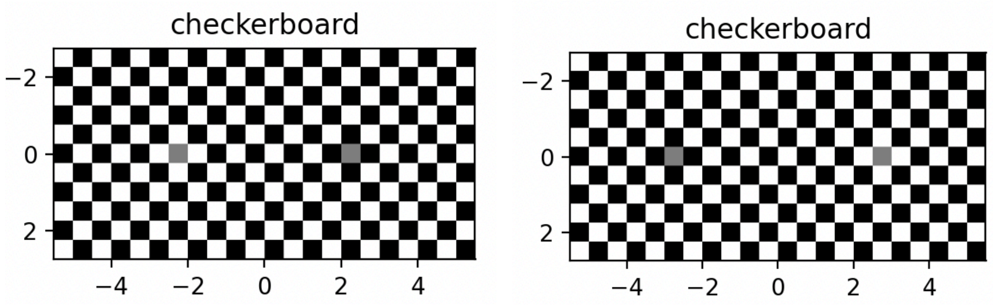


Figure 2.2: (C) within context bullseye image display : the target luminance is changed but the surround context is the same (top is black and bottom is white) (D) across context bullseye image display : one target patch is bordered by black followed by white whereas the other is bordered by white first followed by black. Here the target luminance can be changed or kept the same.

The checkerboard remains fixed in every trial and therefore, the target is placed in one of the two columns depending on the target context (Figure 2.3). The row in which the targets are placed remains constant, the middle of the checkerboard, the 6th row in our stimulus.



(a) The checkerboard stimulus configuration showing the left target placed in white context on the 7th column from the left and right target within the black context on the 7th column from the right.

(b) The checkerboard stimulus configuration showing the left target placed in black context on the 6th column from the left and right target within the white context on the 6th column from the right.

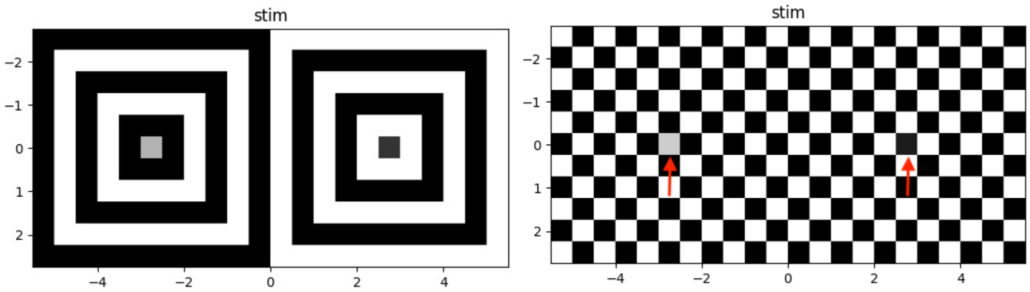
Figure 2.3: Change in the location of the target placements in checkerboard displays with the change in target context.

2.3 APPARATUS

The sessions with the participants were conducted in the Computational Psychology department lab of TU Berlin. The lab is equipped with a monitor, an adjustable chin and forehead rest and a computer. The monitor was kept on a desk at a fixed distance of 80cm from the chin and forehead rest. The participant could sit on an adjustable chair, adjust their chin and forehead on the rest and start the experiment. All of this equipment was enclosed behind a thick black curtain ensuring that no light except the light emitted by the monitor was evaluated by the participants. The monitor used to present the stimuli to the participants was a VIEWPixx/3D LCD with a resolution of 1920 x 1080 pixels and a refresh rate of 120 Hz. This 24-inch monitor with 16-bit resolution was able to display luminance values ranging from 0 to 250 cd/m^2. The inputs from the participants were taken with the help of RESPONSEPixx response box with 5 buttons. For the participants in this experiment only the left, right and centre buttons were relevant. The experiment software was programmed in Python, following the standardised template provided by the research group for lab experiments. The Python library 'stimupy' and 'hrl', which were developed and provided by the research group of Computational Psychology, were utilised in this experiment.

2.4 DESIGN AND PROCEDURE

Different stimulus configurations resulted from all possible pairwise combinations of the 10 luminance values placed in 2 contexts. For each stimulus configuration, the left and the right targets were placed either within context or across context. A stimulus configuration with targets having identical luminance value and identical context were not shown. All possible paired combinations of the 20 target types resulted in a total of 190 different configurations. Trials were organised in blocks and one block consisted of all of these 190 stimuli configurations. The trials were planned to be repeated 15 times resulting in a total of 2850 trials for bullseye and similarly another 2850 trials for the checkerboard displays. In total 5700 trials were planned per observer. On each trial the participant indicated which of the two target patches appeared brighter by pressing the left or the right button (2AFC) on the button-box. There was no time limit for individual trials but a total of 3 sessions, roughly an hour each were planned to divide up the trials equally. Each session contained 1900 trials resulting from 5 repetitions of 190 bullseye configurations and 5 repetitions of 190 checkerboard configurations. The stimulus configurations in checkerboard displays were harder than bullseye as it was not always obvious to locate the target (see Figure 2.4(b)). As a result, measures like introducing horizontal markers to signal the column in which the target is, were implemented. In order to separate the different trials, a quick 200 ms gray blank interval screen was shown in between the response and the next stimulus configuration. Despite these efforts, the checkerboard trials were significantly longer and not all 15 repetitions for checkerboard stimulus configurations could be done for all participants (see discussions for further explanation and possible impact of this limitation).



(a) The bullseye stimulus configuration with the left target luminance 200.25 cd/m^2 and right target luminance of 27.25 cd/m^2. Since the target is always located at the centre, the participant doesn't have to look for it.

(b) The checkerboard stimulus configuration with the left target luminance of 200.25 cd/m^2 and right target luminance of 27.25 cd/m^2. The right target has a very low luminance which makes it harder to locate.

Figure 2.4: Difficulty in locating the targets in checkerboard displays

2.5 SCALE ESTIMATION

Maximum likelihood conjoint measurement (Ho, Landy, and Maloney, 2008; Knoblauch and Maloney, 2012) uses statistical models to map the physical luminance values of target patches within a given context to perceived brightness. These models help psychophysicists to account for the complex interactions between the different variables (here target luminance and target placement) and estimate perceptual scales (here for brightness) (Maloney and Knoblauch, 2020). The statistical models in MLCM take the stochasticity of observers' judgements into account.

On each trial, the observer computes two perceptual values $\psi_{c_1}(s_1)$ and $\psi c_2(s_2)$ prompted by luminance s_1 and s_2 in contexts c_1 and c_2. The resulting estimates of perceived brightness amount to the decision variable δ, which is the simple difference between the two perpetual magnitudes:

$$\delta = \psi_{c_2}(s_2) - \psi_{c_1}(s_1) + \epsilon$$

where the gaussian random variable $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is a judgement error. It captures the possible irregularities in the responses of the observers.

If $\delta > 0$, the observer opted for the second stimulus $\psi c_2(s_2)$ as the brighter one (Maloney and Knoblauch, 2020; Vincent et al., in press).

The model in MLCM can either be (a) the additive model, or (b) the saturated model. A significance test (likelihood ratio test) helps determine which model is suitable for the collected data (Knoblauch and Maloney, 2012). The additive model tests whether perceptual judgements depend on both stimulus dimensions in an additive way. In this case, the perceptual scales have the same shape and only differ by a vertical offset (Fitousi, 2021; Vincent et al., in press). The perceptual scales are calculated independently of each other using the saturated model and therefore the shape of the scales can look very different. For both the checkerboard and bullseye displays, it turned out that the saturated model fits best for most participants' data. Only in the case of 2 participants, AJ and GA, we saw the additive model being a better fit for their checkerboard data (see results).

We used the MLCM implementation available for the R programming language (Knoblauch and Maloney, 2012; R Core Team, 2021). This implementation uses a generalised linear model (GLM) to estimate the scale parameters. Through bootstrapping procedures, confidence intervals were determined, and the goodness-of-fit of the model was evaluated. For this purpose, new responses were generated from the already estimated values. Subsequently, the scales were estimated using the simulated responses. Residuals were then calculated for each bootstrap dataset. The residuals obtained from the experiment were subsequently compared with the distribution of the simulated residuals. If the experimentally determined distribution of residuals fell within the simulated distribution, the assumption of the model was correct (Aguilar and Maertens, 2020, 2022; Knoblauch and Maloney, 2012). By default MLCM provides scales that are anchored at zero for any stimulus. In this thesis, the target with lowest luminance and within the black context is the zero anchor. The perceptual scales are normalised by dividing all their values by their maximum for each participant individually so that the maximum scale value is 1. This is done in order to compare scales across participants. The maximum value of the scales reflects the estimated noise present in the observers' judgments, where a higher maximum value reflects lower observers' noise. Knoblauch and Maloney (2012) provide detailed information about these statistical procedures.

3

RESULTS

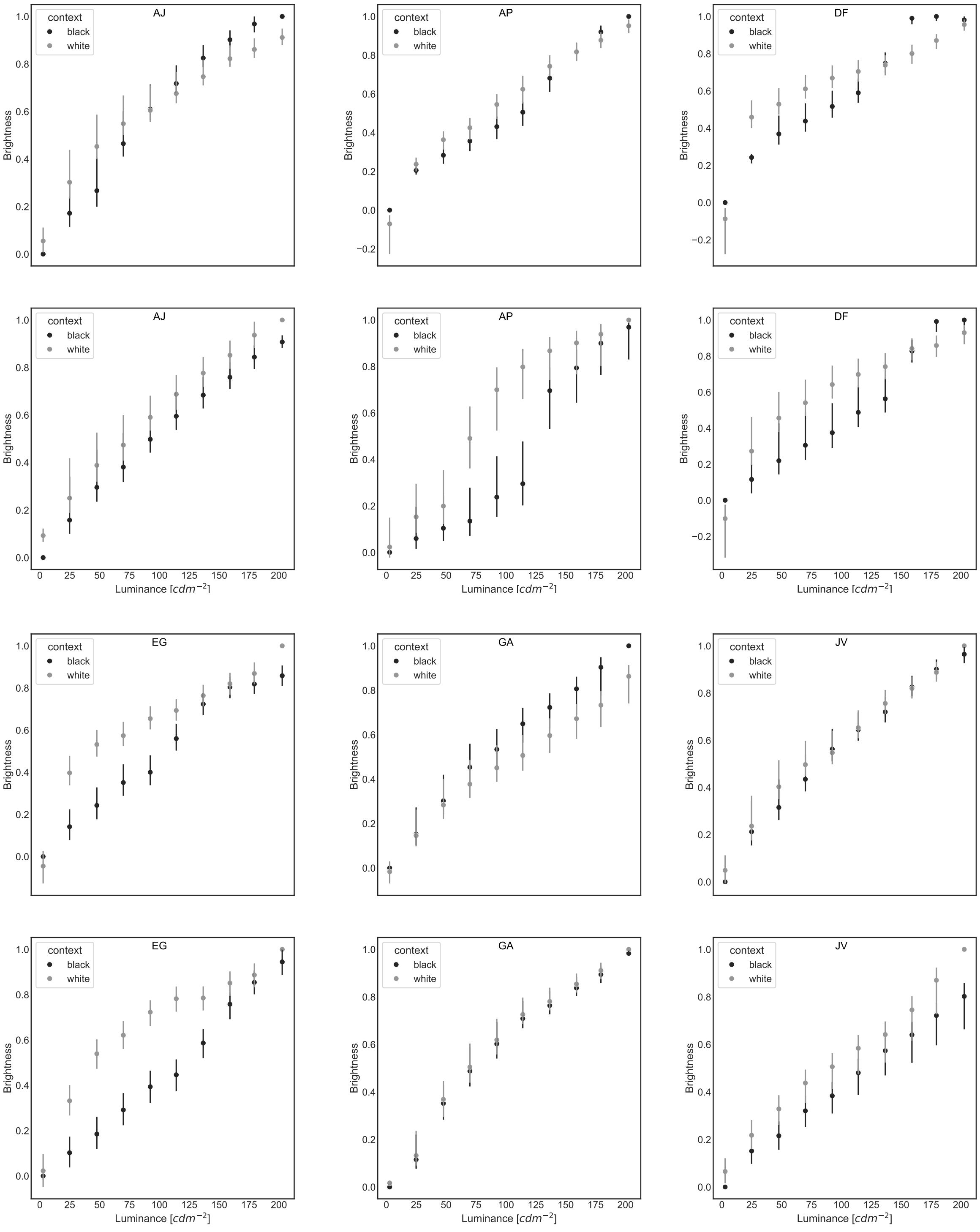


Figure 3.1: Perceptual scales of assimilation effects in bullseye-high-freq (upper panel for each observer) and checkerboard (lower panel for each observer). The names of each observer can be seen on the top center of each panel. Each panel depicts the perceptual scales from an observer as a function of target luminance (x-axis) and target placement (color). GA, JV and NJ are the informed participants, others are naive. Error bars depict 95% confidence intervals.

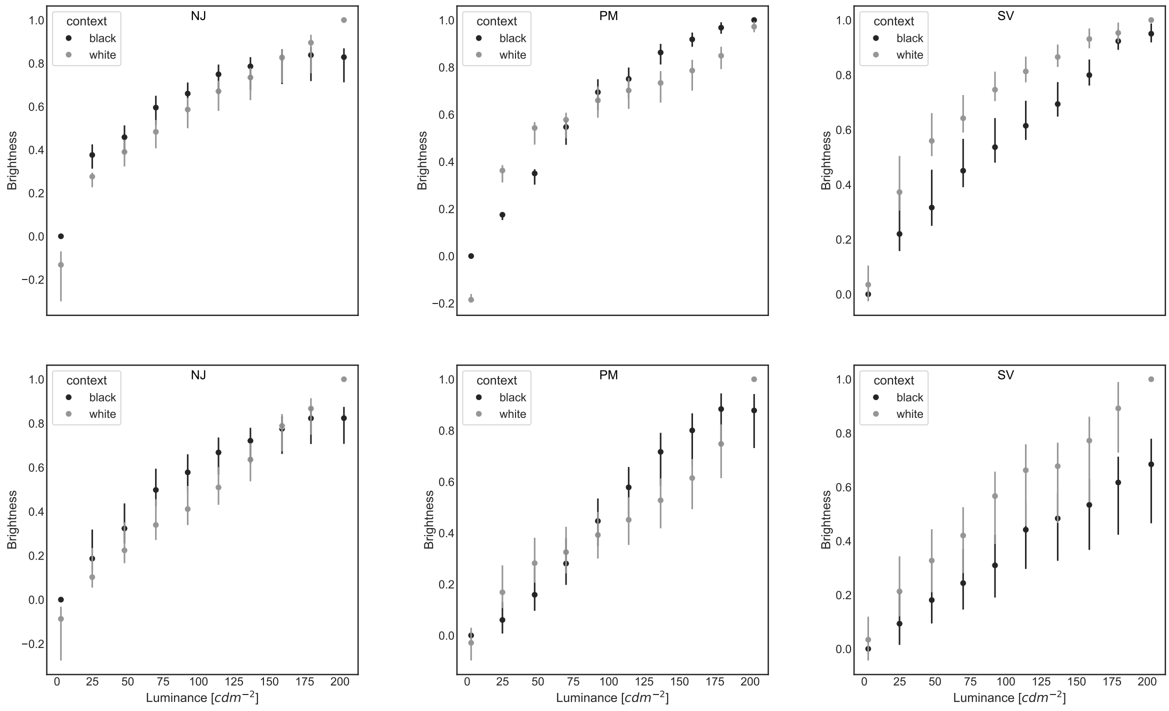


Figure 3.1: Perceptual scales of assimilation effects in bullseye-high-freq (upper panel for each observer) and checkerboard (lower panel for each observer). The names of each observer can be seen on the top center of each panel. Each panel depicts the perceptual scales from an observer as a function of target luminance (x-axis) and target placement (color). GA, JV and NJ are the informed participants, others are naive. Error bars depict 95% confidence intervals (cont.)

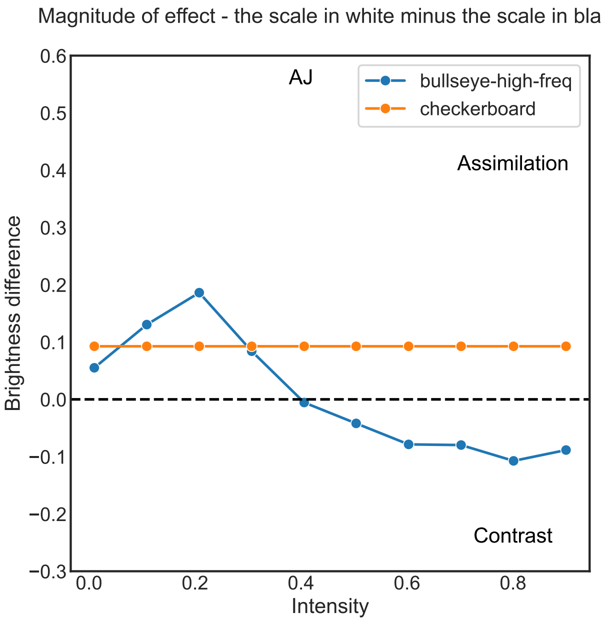
The perceptual scales of White's effect (see Introduction, Figure 1.4) showed variability in the magnitude of the effect across observers, but the direction of the effect throughout the entire luminance range and across all observers was consistent. Similar to the scales of White's effect, the scales measured for assimilation effects (Figure 3.1) in both stimuli are monotonically increasing. But there is definitely more inter-observer and inter-stimulus variability seen in perpetual scales of assimilation effects.

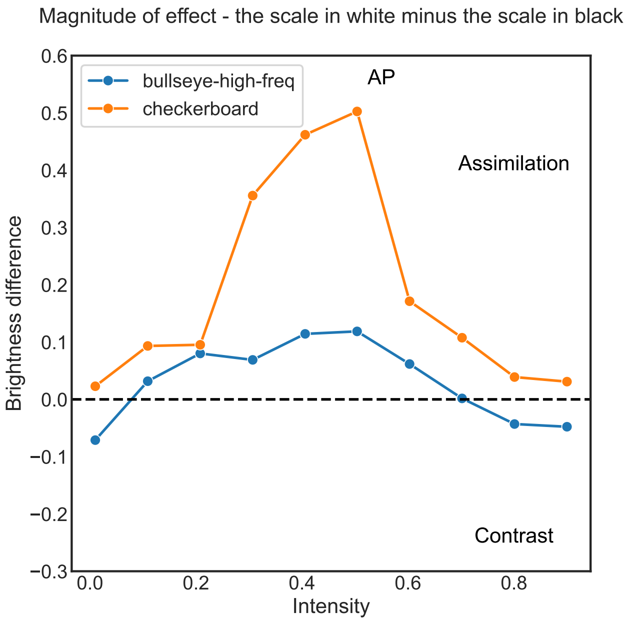
In checkerboard displays, consistent with the direction of assimilation, the target within a white context was consistently perceived brighter than the target in black context (the scales in white are higher than the scales in black for checkerboard displays in Figure 3.1) by majority of the participants (AJ, AP, EG, GA, JV, and SV). The remaining 3 participants DF, NJ and PM do not show an effect in the assimilation direction throughout all luminance values. DF shows an opposite effect in the contrast direction at both extremes and a rather strong assimilation at the intermediate luminance targets. The participants NJ and PM show a more consistent contrast rather than assimilation, which is only shown in some luminance values.

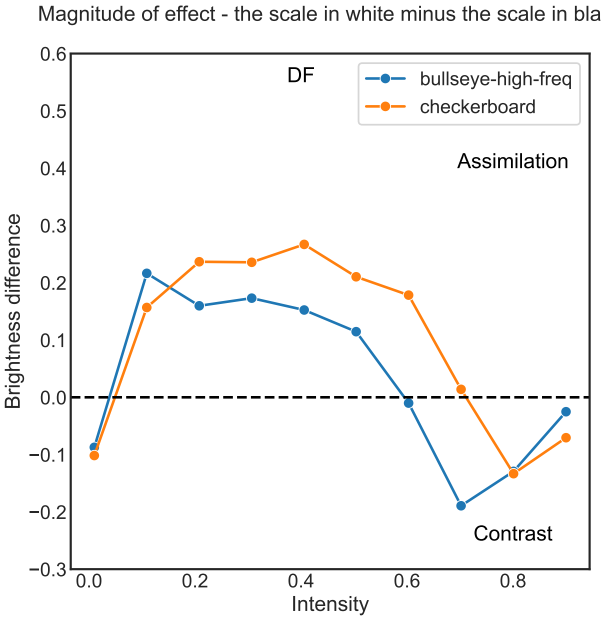
The bullseye scales exhibit greater variability, with only 3 out of 9 participants (EG, JV and SV) showing assimilation across all luminance values. A consistent but very weak effect in the contrast direction is shown by GA. The rest of the participants are divided between showing assimilation for intermediate luminance targets or showing it for one of the extreme luminance range. AP, DF and PM belong to the former group while AJ and NJ belong to the latter. Assimilation for the lower luminance values is shown by AJ while assimilation only for the higher luminance values is shown by NJ.

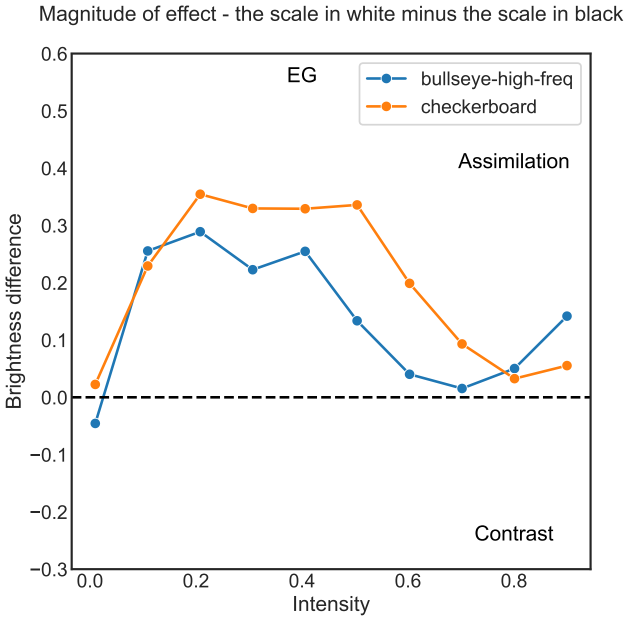
Across the entire luminance range, not only are there changes in the magnitude of effects (also seen in scales of White's effect, Figure 1.4), but also a change in the direction of effects (can be seen in Figure 3.1 as crossovers in the scales or in Figure 3.2 as changes in magnitude from positive to negative or vice versa) shown by some participants. This change of direction from assimilation to contrast or vice versa is predominantly observed at the extreme luminance values (AP, JV for bullseye displays and DF, NJ for both stimuli). Crossovers at the more intermediate luminance values are shown by PM and AJ. Participants EG, SV and GA show either an effect in contrast or in assimilation direction throughout the luminance range for a given assimilation display, therefore no change in direction of effects was observed for these participants.

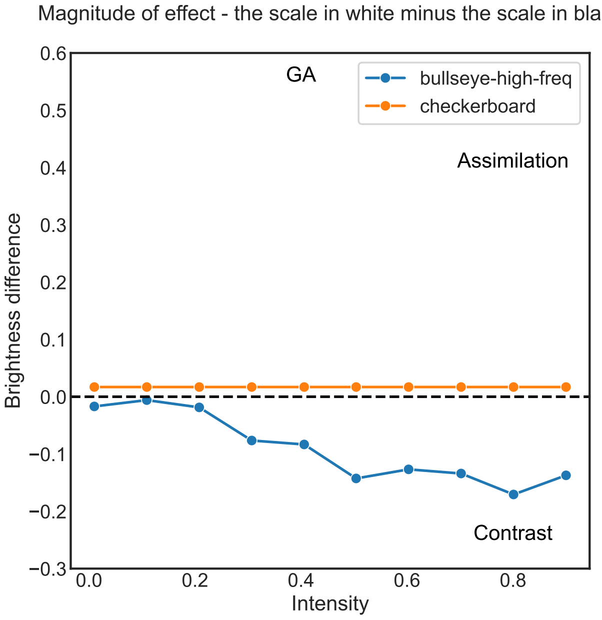
The magnitude of an effect can be calculated by subtracting the scale value in black context from the scale value in white context. If the difference is zero, it means that there was no effect in any direction. If the difference is positive, the effect was shown in the assimilation direction and finally if the difference is negative, it is in the contrast direction. This magnitude varies a lot across observers, luminance intensities and the assimilation displays (see Figure 3.2). Strong assimilation for intermediate luminance targets in checkerboard displays are shown by the participants AP, DF, EG, SV. These participants also show a sudden increase in their brightness perception of the target in white going from the first luminance value to the second luminance value. Sudden changes like these are also observed for other participants like PM's checkerboard scales but here, it is accompanied by the change in direction — contrast to assimilation in the highest luminance target. Similar to the checkerboard scales, the bullseye scale values in white seem to increase suddenly going from the first to the second luminance value for the participants AP, DF, EG, NJ and PM.

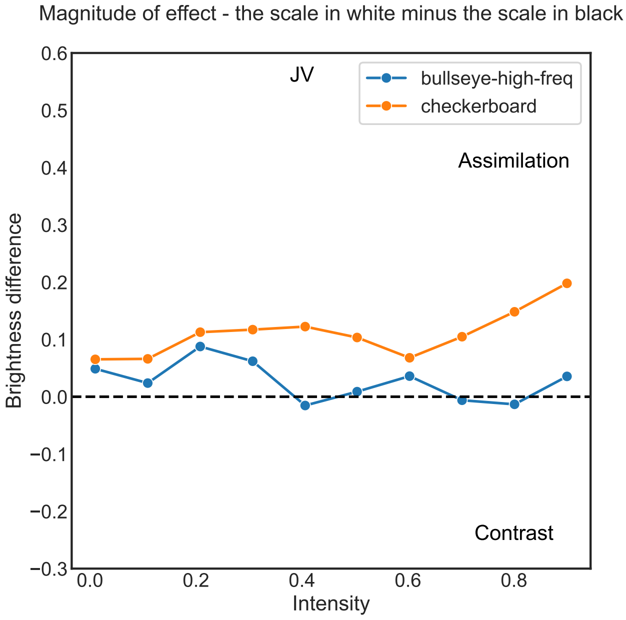

Magnitude of effect - the scale in white minus the scale in black
AJ
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8


Magnitude of effect - the scale in white minus the scale in black
AP
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8


Magnitude of effect - the scale in white minus the scale in black
DF
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8


Magnitude of effect - the scale in white minus the scale in black
EG
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8


Magnitude of effect - the scale in white minus the scale in black
GA
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8


Magnitude of effect - the scale in white minus the scale in black
JV
bullseye-high-freq
checkerboard
Assimilation
Contrast
Brightness difference
Intensity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.2
0.4
0.6
0.8

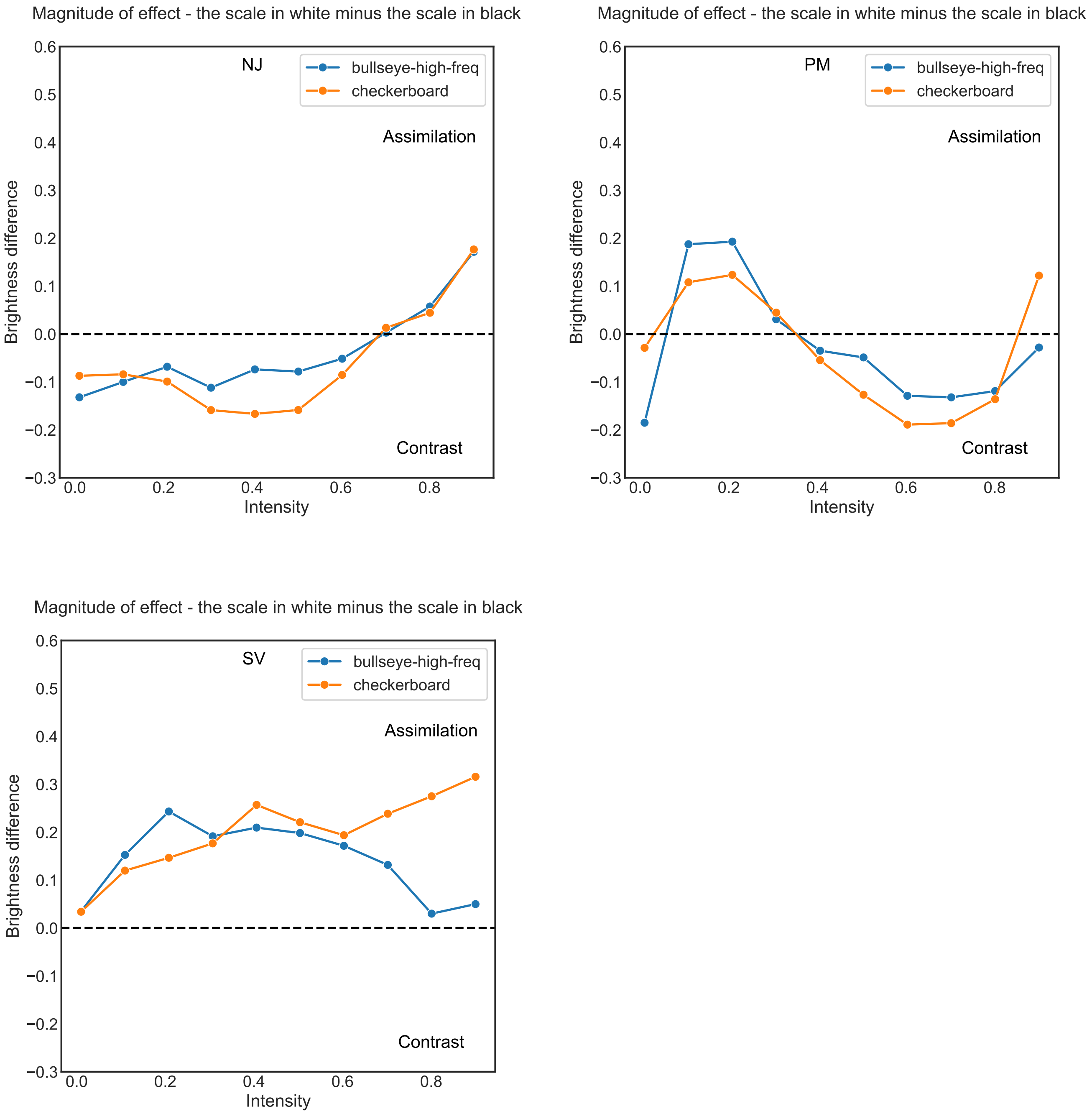


Figure 3.2: The difference between the scales in black context from the white context which gives the magnitude/strength of the effect. If the magnitude of effect is zero (dashed line in each panel), there is no effect. If the magnitude is positive, the effect is in assimilation direction whereas magnitude in negative means that the effect is in contrast direction. The bigger the difference, the stronger is the effect. In each panel is the magnitude of effect for each observer for both the bullseye (blue) and checkerboard (orange) scales. The x-axis shows the luminance intensity values and the y-axis shows the brightness difference. The checkerboard scales of AJ (first row top left) and GA (third row bottom left) are a flat line because those scales were estimated with the additive model.

When comparing checkerboard and bullseye displays for the same observer, notable similarities in the shape of the perceptual scales are seen for most observers. It can be noticed that the perceptual scales of checkerboard displays consistently exhibit effects in the assimilation direction that are either stronger or equal to those observed in bullseye displays across all participants. This can also be seen in Figure 3.2, where for most participants in any given luminance value, the checkerboard has a higher positive difference, indicating a stronger effect in assimilation direction. Participant AJ shows assimilation in the lower target luminance of bullseye displays, followed by contrast at higher luminance values. However, in checkerboard displays, this participant consistently shows only assimilation across the entire luminance range. Similar patterns are noted in participants AP and JV, where effects in the assimilation direction predominate for lower luminance targets in bullseye displays, transitioning to effects in the contrast direction for higher luminance targets, albeit with minor deviations from this trend. However, in checkerboard displays, these participants uniformly show assimilation across all luminance values, with an overall increase in the vertical distance between the black and light gray scale, indicating increased assimilation strength. Other participants, such as EG and SV, consistently show only assimilation in both bullseye and checkerboard displays; however, the overall strength of the effect is greater in checkerboard displays. For the participant GA the scales in bullseye show a consistent effect in contrast direction which reverses in checkerboard, where they show weak assimilation across all targets. The only exception to this observation is the participant NJ showing contrast in the lower luminance range, followed by assimilation for higher luminance values. There is no indication of increased assimilation in checkerboard for this participant as the scales of both stimuli look very similar.

The effects in assimilation direction are perceived stronger in checkerboard than bullseye across all participants but there is still a lot of variability. In cases where the participant shows effects in both assimilation and contrast direction for different luminance targets, it is more common to see assimilation direction in the intermediate or in the lower luminance targets. The luminance value at which the transition from assimilation to contrast, or vice versa, happens, varies among different observers. NJ being an exception to most participants shows relatively strong contrast throughout most luminance values except for the high luminance targets where assimilation is shown.

4

DISCUSSION

We were able to successfully measure the perceptual scales of assimilation displays, compare these scales with each other and across different participants for interesting patterns, look at the more untested extreme luminance values for inconsistencies and finally, compare the scales of assimilation displays with the scales of White's effect (measured in Vincent et al., in press).

The perceptual scales of bullseye and checkerboard exhibited considerable variability across different observers and target luminance values. The perceptual scales in both stimuli were monotonically increasing. The checkerboard scales consistently exhibited effects in assimilation direction that were either stronger or equal to those observed in bullseye displays across all participants. A few participants (EG, JV and SV) showed assimilation across all luminance targets in both stimuli. For most participants however, the direction of the effects would also change across the luminance values. In these cases, assimilation was mostly shown in the intermediate or in the lower luminance targets. The participants GA (only bullseye scales), NJ, and PM showed contrast for the majority of the luminance range. Although certain patterns emerged within the scales of individuals and between checkerboard and bullseye scales, the substantial variability makes it challenging to draw definitive conclusions from the results. We compared the scales of assimilation displays with the scales of White's effect. There was more robustness observed in the scales of White's effect as all participants showed effects in the same direction (depending on the interpretation, assimilation or contrast). The only difference across all participants was the vertical difference between the scales in black and white context for a given luminance value, meaning that the strength of the effect varied. As the results seen for White's effect are much more consistent and conclusive, it seems to be robuster than assimilation displays. This difference in robustness between them could only be due to the difference in the physical stimuli itself. The target in White's effect does not have the same local contrast meaning that all the 4 edges of the target patch are not surrounded by the same physical contrast. Somehow this difference in the local borders might be causing the effect to strengthen causing the scales to look more reliable.

From the results it is clear that, both displays do show assimilation effects but not exclusively, not throughout the entire range of target luminance or not at all. In their matching experiments with bullseye displays, Bindman and Chubb (2004) tested the strength of assimilation effects across three luminance targets. Specifically, the intermediate target luminance was equal to the background luminance, while the remaining two targets varied slightly, one being higher and the other lower. It was observed that assimilation was strongest with low luminance targets, no effects were seen for the intermediate luminance targets and reverse effects, i.e. effects in contrast direction were observed for targets with higher luminance. It is clear that even when only 3 luminance values relatively close to each other were tested, consistent effects in assimilation direction were not noted. It is also seen that making changes to the stimulus, like changing the target luminance can cause the direction of the effect to reverse.

The perceptual scales of the checkerboard stimulus showed more consistency as most participants showed assimilation. The strength of assimilation was still very different depending on the observer but it was predominantly stronger at the intermediate luminance targets. Due to the fact that most observers showed assimilation in checkerboard displays throughout the luminance range while not showing similar patterns for bullseye displays, it is very much possible that both displays produce effects in the direction of assimilation but might not be from the same mechanism. Therefore there is a chance that there might not just be one type of assimilation, but multiple, arising from different mechanisms. Consequently, it is then also possible that effects such as White's effect might be explained through one assimilation direction but not the other. These are just some assumptions we can draw from comparing the scales of both assimilation displays but further experiments have to be conducted, preferably without the limitations mentioned below, to say something more concrete about these displays and assimilation in general.

The comparison of the two assimilation displays might have not lead to the most congruent results. During the data collection process, it was clear that the difficulty and time taken to select the brighter target was different for the checkerboard and bullseye images. In the bullseye stimulus configuration, the target is very easy to find and the task is just restricted to selecting the brighter target. In checkerboard stimulus configurations on the other hand, the task is not just limited to finding the brightest target but first to find the location of the target in certain cases. It becomes exceptionally hard to find a bright target (close to 0.9) surrounded by black surround context or a dark target (close to 0.1) surrounded by white surround context because then it looks like the rest of the checkerboard (see Figure 2.4). Even with the aid of horizontal markers which indicated the columns in which both targets were, it was really hard for the participants to look at both targets at the same time and compare their brightness. As a result depending on the participant, we only managed to get the checkerboard data from 5-12 (out of 15) trials in total.

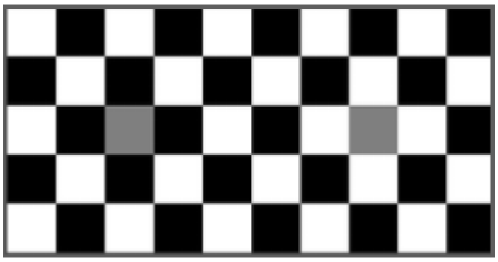


Figure 4.1: Original checkerboard stimulus with 5 rows and 10 columns target size of 1 x 1 degrees.

The original checkerboard (Figure 4.1) only had 5 rows and 10 columns with the target size 2 x 2 degrees. The reason for changing the size of the checkerboard was the lack of assimilation effects seen by the author in the first few pilot sessions. Heat maps were generated for the initial pilot session data and the original checkerboard stimulus barely showed assimilation effects for the author — the changed checkerboard showed better results and therefore, we decided to move forward with it. Since the pilot sessions were done on non-naive participants, the difference in time and effort for both stimuli was not recognised.

MLCM could be successfully used to estimate the perceptual scales of assimilation effects over a larger luminance range. Not every participant showed consistent assimilation through the entire range of luminance values. It was more common to see effects in both assimilation and contrast direction. Often contrast was shown at the extreme luminance values rather than the intermediate, where the observers predominantly showed assimilation. This was a pattern which could only be uncovered because a method like MLCM was used to measure the perceptual scales. It was possible to look outside of the normally tested luminance range for inconsistencies at the extreme luminance values. We were also able to compare the scales of White's effect estimated in Vincent et al. (in press) with the scales of bullseye and checkerboard displays estimated in this thesis. The scales for White's effect were far more consistent across different observers and therefore we can say that the White's stimulus is much more robust than the assimilation displays. Since checkerboard scales showed stronger assimilation and was more consistent across all observers compared to bullseye scales, we explored the possibility of different displays producing effects in assimilation direction but coming from different mechanisms. Though there were certain limitations in the design of the experiment, the results suggest that assimilation is complex and there is definitely not just one assimilation. It is not possible to fully answer if all the surround context effects arise from the same mechanism or not. This can only be answered if the participants, experiment tasks and procedures, and the stimulus design (like the target size, the luminance range) across all surround context effects remain constant and then those perceptual scales are compared.

REFERENCES

Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., p. 339-351). Cambridge, MA: MIT Press.

Aguilar, G., and Maertens, M. (2020). Toward reliable measurements of perceptual scales in multiple contexts. *Journal of Vision*, *20*(4), 19. Retrieved from `https://doi.org/10.1167/jov.20.4.19` doi: doi:10.1167/jov.20.4.19

Aguilar, G., and Maertens, M. (2022). Conjoint measurement of perceived transparency and perceived contrast in variegated checkerboards. *Journal of Vision*, *22*(2), 2. Retrieved from `https://doi.org/10.1167/jov.22.2.2` doi: doi:10.1167/jov.22.2.2

Barkan, Y., Spitzer, H., and Einav, S. (2008). Brightness contrast–contrast induction model predicts assimilation and inverted assimilation effects. *Journal of Vision*, *8*(7), 27. Retrieved from `http://journalofvision.org/8/7/27/` doi: doi:10.1167/8.7.27

Bindman, D., and Chubb, C. (2004). Brightness assimilation in bullseye displays. *Vision Research*, *44*(3), 309–319. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0042698903004309` doi: doi:10.1016/S0042-6989(03)00430-9

Blakeslee, B., and McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research*, *44*, 2483–2503.

De Valois, R. L., and De Valois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

de Weert, C. M., and van Kruysbergen, N. A. (1997). Assimilation: central and peripheral effects. *Perception*, *26*(10), 1217–1224. doi: doi:10.1068/p261217

Fitousi, D. (2021). How facial aging affects perceived gender: Insights from maximum likelihood conjoint measurement. *Journal of Vision*, *21*(12), 12. Retrieved from `https://doi.org/10.1167/jov.21.12.12` doi: doi:10.1167/jov.21.12.12

Gilchrist, A. L., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., ... Economou, E. (1999). A new theory of lightness perception. *Psychological Review*, *106*(4), 795–834.

Helson, H. (1963). Studies of anomalous contrast and assimilation. *Journal of the Optical Society of America*, *53*, 179–184. ([PubMed])

Ho, Y. X., Landy, M. S., and Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science*, *19*, 196–204.

Kingdom, F. A. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, *51*(7), 652-673. Retrieved from `https://doi.org/10.1016/j.visres.2010.09.012` doi: doi:10.1016/j.visres.2010.09.012

Knoblauch, K., and Maloney, L. T. (2012). *Modeling psychophysical data in r*. Springer Science and Business Media.

Maloney, L. T., and Knoblauch, K. (2020). Measuring and modeling visual appearance. *Annual Review of Vision Science*, *6*, 13.1–13.19. Retrieved from `https://doi.org/10.1146/annurev-vision-030320-041152` doi: doi:10.1146/annurev-vision-030320-041152

R Core Team. (2021). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/` ([Computer software])

Schmittwilken, L., Maertens, M., and Vincent, J. (2023). stimupy: A python package for creating stimuli in vision science. *Journal of Open Source Software*, *8*(86), 5321. doi: doi:10.21105/joss.05321

Vincent, J., Maertens, M., and Aguilar, G. (in press). What fechner could not do: Separating perceptual encoding and decoding with difference scaling. *Journal of Vision (In Press)*.

White, M. (1979). A new effect on perceived lightness. *Perception*, *8*, 413–416.