The Crispening Effect: An Artifact of a Method or a Feature of the Visual System

Thesis zur Erlangung des Grades
Bachelor of Science (B. Sc.)
im Studiengang Informatik

Carolin Brunn
carolin.brunn@campus.tu-berlin.de

Computational Psychology
Institut für Technische Informatik und Mikroelektronik
Fakultät Elektrotechnik und Informatik
Technische Universität Berlin

Gutachter:
Marianne Maertens
Felix Wichmann

eingereicht am: 28. Oktober 2020

Hiermit erkläre ich an Eides statt, dass die vorliegende, dieser Erklärung angefügte Arbeit selbstständig und nur unter Zuhilfenahme der im Literaturverzeichnis genannten Quellen und Hilfsmittel angefertigt wurde. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind kenntlich gemacht. Ich reiche die Arbeit erstmals als Prüfungsleistung ein.

Berlin, 28. Oktober 2020

. . . .

(Carolin Brunn)

Zusammenfassung

Trotz langjähriger Forschung wurden viele Prozesse, die in unserer visuellen Wahrnehmung eine Rolle spielen, noch nicht verstanden. Ein fundamentales Beispiel dafür ist die Helligkeitswahrnehmung, also die Wahrnehmung der Oberflächenreflektion. Situationen, in denen die Helligkeitswahrnehmung nicht auf Grundlage der physikalischen Luminanzwerte vorhergesagt werden kann, sind besonders interessant in der Forschung, da sie uns Besonderheiten des visuellen Systems aufzeigen können. Ein Beispiel für solche Phänomene ist der "Crispening Effect" (CE), bei dem die Wahrnehmung eines Objekts von seinem Hintergrund beeinflusst wird, je ähnlicher die Luminanz des Objekts und die des Hintergrunds werden.

Bisher wurde der CE hauptsächlich mit Methoden gemessen, die auf der Anpassung der Helligkeit von Objekten beruhen. Diese Methoden sind jedoch nicht in der Lage, die tatsächliche interne "Wahrnehmungsfunktion" einer Person zu messen, die den Zusammenhang zwischen Luminanz und Helligkeit darstellt. Daher soll in dieser Arbeit analysiert werden, ob der CE nur im Zusammenhang mit den genannten Methoden auftritt oder ob der Effekt eine Charakteristik des visuellen Systems ist.

Für die Beantwortung dieser Frage analysiere ich existierende Daten, die von Aguilar und Maertens (2020) mit statistischen Methoden (MLDS, MLCM) gemessen wurden. Beide Methoden messen die Wahrnehmungsfunktion und basieren auf dem Vergleich der Helligkeit statt deren Anpassung.

Tatsächlich ist der CE in den Skalen der einen Methode (MLDS) sichtbar, was impliziert, dass der Effekt eine Charakteristik des visuellen Systems ist. Mögliche Gründe, warum der CE in den Skalen der anderen Methode (MLCM) nicht auftritt, erkläre ich später. Außerdem stelle ich mögliche Indikatoren für den CE vor, wie z.B. den Funktionsanstieg.

Abgesehen von der Datenanalyse simuliere ich ein Experiment, um herauszufinden, welchen Einfluss der Abstand der Datenpunkte auf die Wahrnehmungsfunktion hat. Hierfür nutze ich verschiedene interne Funktionen für die fiktiven Versuchsteilnehmer. In der Evaluation der Simulation analysiere ich, wie gut die internen Funktion gemessen wurden und wie gut die Wahrnemungsfunktionen, die auf verschiedenen internen Funktionen basieren, unterschieden werden können.

Die Schlussfolgerungen, die ich aus der Simulation ziehe, könnten eine Erklärung dafür bieten, warum der CE nicht in den Wahrnehmungsfunktionen aller Versuchsteilnehmer in der Analyse auftritt. Da die Ergebnisse dieser Arbeit darauf hinweisen, dass der CE eine Charakteristik des visuellen Systems ist, sollte in der Zukunft versucht werden, den Grund und die genauen Eigenschaften des Effekts zu ergründen.

Abstract

Notwithstanding decades of research, many perceptual processes involved in humans' visual perception are yet to be fully understood. Understanding the perception of surface reflectance, also called lightness perception, is fundamental for this research. Of particular interest are visual phenomena in which perceived lightness cannot be predicted from the physical luminance signal. As such, the Crispening Effect (CE) describes a phenomenon in which the perceived lightness of an object is influenced by its surroundings. The difference of two samples is enhanced when the samples' luminance approaches the background luminance.

Until now, CE has mostly been studied with matching methods using adjustment tasks. One weakness of adjustment tasks is that they only reflect but do not recover an observer's perceptual scales, which describe the relationship between luminance and perceived lightness. This thesis therefore investigates whether CE might simply be an artifact that results from the methodological weaknesses of adjustment tasks, or whether CE is an actual feature of visual processing.

For this, I reanalyze perceptual scales obtained by Aguilar and Maertens (2020) with two statistical scaling methods, Maximum Likelihood Difference Scaling (MLDS) and Maximum Likelihood Conjoint Measurement (MLCM). Both methods directly measure perceptual scales of the observers and are based on difference judgments instead of adjustment tasks.

In my reanalysis, I observe CE in most perceptual scales obtained with MLDS but not with MLCM. The reproduction of CE with MLDS indicates that CE is indeed a feature of the visual system. Potential reasons for the lack of CE in MLCM scales are discussed. As possible indicators for CE, I analyze the slope and the discriminability of the scales.

Additionally, I investigate the influence of differently distributed luminance values on perceptual scales and on the occurrence of CE. For this, I simulate data in an asymmetric matching experiment and an MLDS experiment with hypothetical observers based on different perceptual scales. Some distributions recover the internal functions accurately while others do not. The simulations indicate that data points close to the background luminance are needed for the detection of CE. Furthermore, I analyze the discriminability of the simulated scales to have a notion of expected parameters for further work on CE.

Following from the results of the simulation, the spacing of the luminance values in Aguilar and Maertens (2020) might cause the absence of CE in the reanalysis of some observers' scales. Since the results of this thesis indicate that the CE is a processing characteristic of the visual system, future research should focus on characterizing the nature and the cause of the effect.

Acknowledgment

First I want to thank my advisor Guillermo Aguilar for all the patience and encouragement. I am grateful for the numerous talks to clarify my questions and the valuable advice that I got from him.

I also want to thank Marianne Maertens, who trusted me to start in her department without any previous knowledge of psychophysics. I am extremely grateful for all the opportunities I got thanks to her. I am thankful for the trust she has given me and all the knowledge I gained in her department. While working with her I learned new ways of thinking and got to know new points of view.

I would also like to thank my colleagues Minjung Kim and Lynn Schmittwilken who shared their knowledge with me and gave me a lot of valuable feedback. Additionally, I want to thank Felix Wichmann and Bernhard Lang for their advice on my thesis.

Finally, I would like to thank my family, friends, and colleagues, without whom I would have never come so far.

Contents

Chapter 1

Introduction

Our perception of the appearance of objects depends primarily on the illumination of a scene by a light source and on an object's reflectance. A light source emits a certain amount of light defined as illuminance. An object, also called stimulus, reflects this light, which causes a signal, luminance, to reach the eye. Figure 1.1 shows this concept. Illuminance, reflectance and luminance are labeled with I, R, and L, respectively. Luminance depends on the illumination and the amount of reflectance. Additionally, luminance might be influenced by transparencies that are located between the object and the observer, labeled as T.

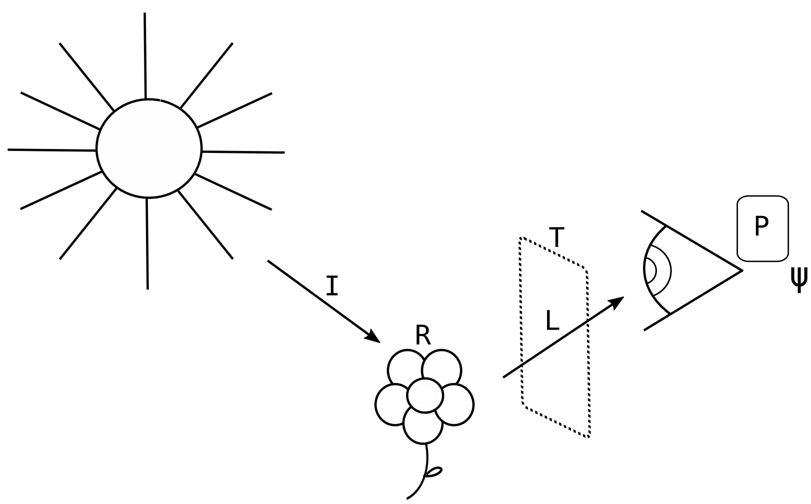

Figure 1.1: Relationship between illuminance (I), reflectance (R), luminance (L), and perceived lightness ψ. Luminance is a combination of illuminance, reflectance, and potentially intervening media such as transparency, T. The P represents the unknown underlying process in the visual system originating the perception of lightness, ψ. Normally, the perceived lightness ψ highly correlates with the reflectance of an object, independent of the illuminance in the scene. The drawing was inspired by Aguilar and Maertens, 2020.

When the luminance signal reaches the eye, it stimulates photoreceptors. This sensory signal is further processed, and a perceptual response is evoked in the form of perceived lightness, which can be thought of as perceived surface reflectance. In this thesis, the terms lightness and perceived lightness are used interchangeably. In Figure 1.1, perceived lightness is labeled as ψ. The exact neuronal mechanisms underlying these perceptual processes are still not fully understood. These unknown processes are shown as P in Figure 1.1.

1.1 Lightness Constancy

The luminance, the physical signal that reaches the eye, depends on both the illuminance reaching the object and the object's reflectance. Perceived lightness does not correlate with the illumination; but it does so with the reflectance of an object. Therefore, we perceive the whole range between black and white also in darker viewing conditions such as shaded areas. One of the open questions in vision research is how the visual system maps these different luminance ranges to the internal lightness scale. Physically, the luminance value of a white object in the shadow might be equal to a gray object in sunlight – nevertheless, it might still be perceived as white. This concept is called lightness constancy.

Checker Shadow Illusion - An Example of Lightness Constancy

Figure 1.2 provides an example, which is known as the checker shadow illusion (Rogers, 2017), to illustrate lightness constancy in a better way. The squares A and B have the same physical luminance value. However, we perceive Square B as lighter than Square A because in a real-world object the dark check has lower reflectance than the light check. Despite decades of research, it is still unknown how exactly the luminance signal is processed by the visual system and how the respective perceived lightness is produced.

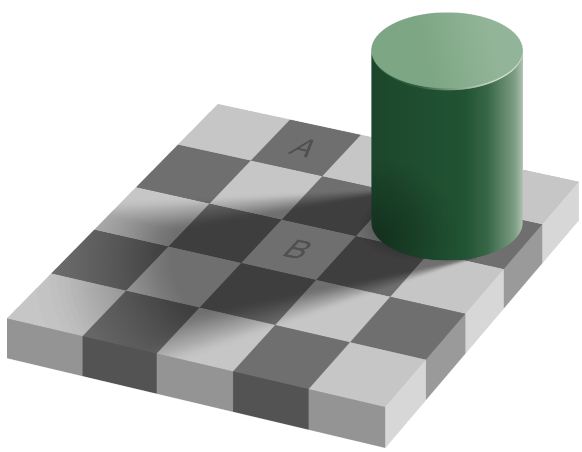

Figure 1.2: In the checker shadow illusion the squares A and B have the same luminance value. Nevertheless, B is perceived as lighter than A. The figure was taken from Thomson and Macpherson, 2017.

The checker shadow illusion shows that the lightness of an object does not only depend on its illuminance but also on its surrounding or context. The same luminance value can be perceived differently, depending on the context it is embedded in. The luminance value is perceived lighter in the shadow than in the outside of the shadow. This shows us that the surrounding context is important. Such phenomena in which perceived lightness cannot be predicted from the physical luminance signal are of particular interest as they impart knowledge of special characteristics of perceptual processes.

1.2 The Crispening Effect

The Crispening Effect (CE) is a known phenomenon where the lightness of a luminance patch is influenced by its background (bg) (e.g., Takasaki, 1966; Whittle, 1992). Figure 1.3 illustrates the phenomenon. Panel 1 shows eight luminance patches placed on a homogeneous background. The patches A to D are decrements – i.e., they have a luminance value lower than the background. Conversely, the patches E to H are increments – i.e., they have a luminance value higher than the background. The luminance values of the decrements and increments are equally spaced in luminance space. The difference between each two consecutive decrements (increments) is the same as the difference between the decrement D (increment E) and the background. In order to illustrate the arrangement of luminance values in a better way, the black line in Panel 2 of Figure 1.3 shows the luminance profile of the scene.

Although the luminance values are equally spaced, the perceived lightness of the patches is not. The difference between the decrement D (increment E) and the background is perceived bigger than the difference between A and B (G and H) by most of the observers. The dashed blue line in Panel 2 illustrates the profile of perceived lightness. Crispening Effect is this change in perceived lightness when the target is close to the background's luminance. The mechanisms causing CE are not fully understood (Kane, 2017; Kane & Bertalmio, 2019).

In order to study lightness phenomena, we should reliably measure lightness and quantify the relationship between luminance and lightness. Several methods have been proposed to elucidate the relationship between luminance and lightness. In the following, I present those which are relevant in the context of this thesis.

1.3 Psychophysical Methods for Measuring Lightness

In order to investigate the relationship between luminance and perceived lightness we, as experimenters, first need to be able to quantify that relationship empirically. Different methods have been proposed to measure their relationship, each of which has advantages and disadvantages. This thesis focuses on two predominant approaches for measuring perceptual appearance: matching and scaling procedures (e.g., MLDS and MLCM) (Kingdom & Prins, 2016).

Scaling procedures measure perceptual scales, which describe an observer's internal perceptual process. A perceptual scale is the mapping from luminance to lightness. Different observers and viewing conditions – e.g., in the sunlight or in the shadows – produce different perceptual scales. With each of these scales, it is possible to predict the perceptual response

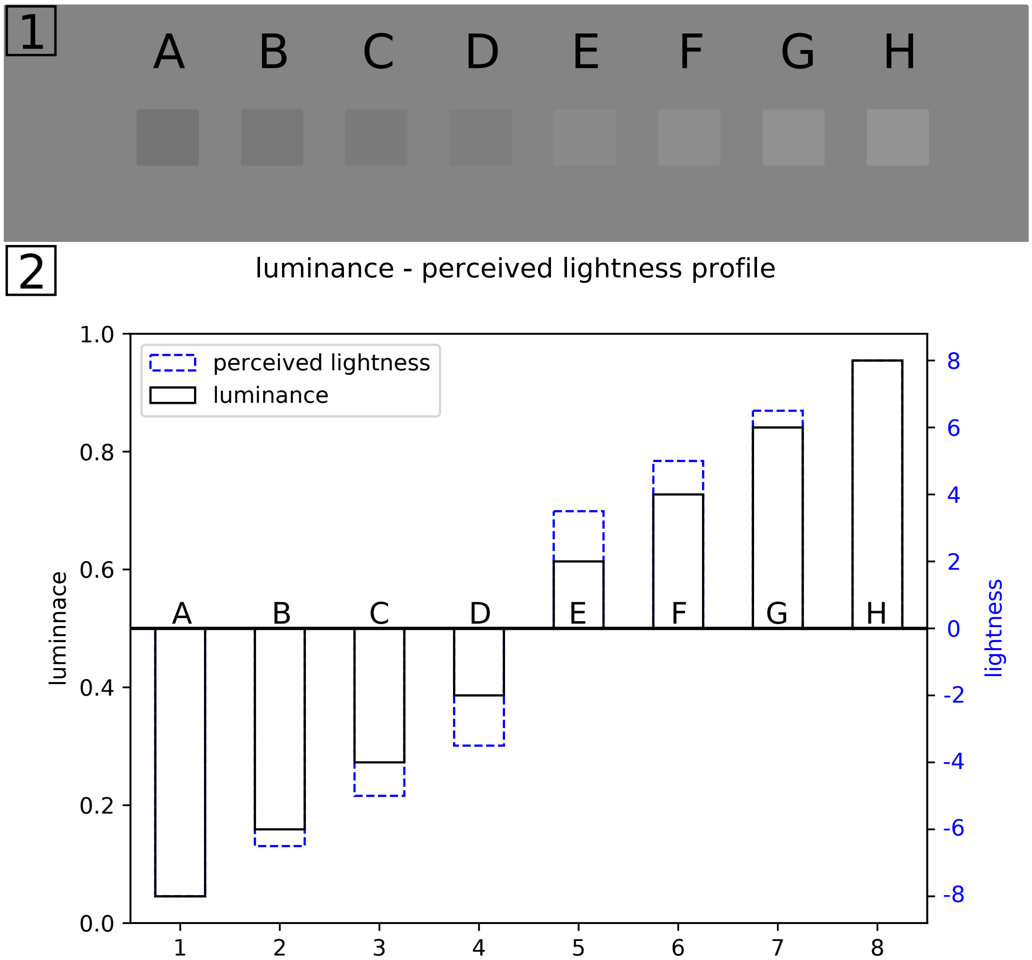

Figure 1.3: The Crispening Effect: Panel 1 displays 8 luminance patches on a homogeneous background – 4 decrements (A–D) and 4 increments (E–H). Panel 2 shows the corresponding luminance and the perceived lightness profile. The black line shows the luminance profile. The decrements and increments are equally spaced in the luminance space. The dashed blue line shows the perceived lightness profile. The perceived lightness of the patches is not equally spaced – e.g., the difference between D (E) and the background is perceived as bigger than the difference between A and B (G and H).

of the particular observer for a stimulus in a particular viewing condition. The matching data obtained in matching procedures only reflect the observer's perceptual scales. Matching data show a mapping from luminance to luminance, and therefore, they do not recover the perceptual scale directly. In the following, I describe the matching and scaling procedures and emphasize their differences.

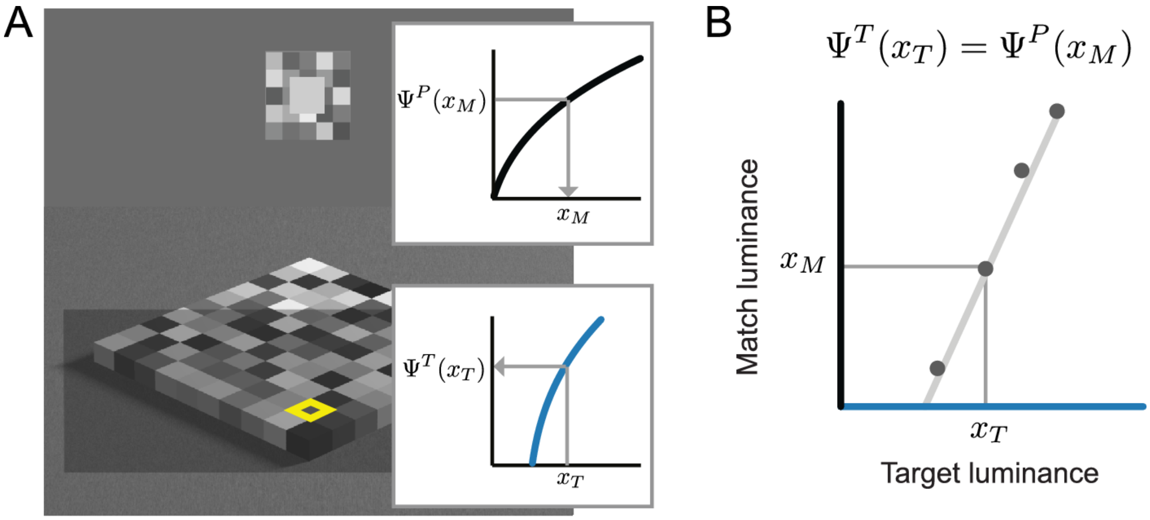

Figure 1.4: Panel A shows the matching procedure with the observer's underlying perceptual scales. The observer has to compare two targets. The upper central square is called match and the lower framed square is called target. The match needs to be adjusted in such a way that its luminance, x_M, is perceived as equal to the luminance of the target, x_T. x_M should evoke a perceptual response, $\psi^P(x_M)$, in such a way that $\psi^P(x_M) = \psi^T(x_T)$, where $\psi^T(x_T)$ is the perceptual response evoked by x_T. Panel B shows the resulting matching data. In order to display the matching data, the corresponding luminance values x_M and x_T are plotted against each other. This procedure is repeated with different values for x_T in order to obtain the data depicted in Panel B. The two graphs in Panel A represent the observer's hypothetical internal processes. However, they cannot be measured directly in matching experiments. The figure is taken from Aguilar and Maertens, 2020.

1.3.1 Matching Method

One of the most commonly used methods is Fechner's method of adjustment, also called matching method. Figure 1.4 conveys the example of a matching procedure and the underlying perceptual processes. In this type of method, the observer has to judge two spatially separated samples and should adjust the luminance value of one sample (*match*) in such a way that the match and the other sample (*target*) are perceived equally. This adjustment can be done, for instance, by choosing from a palette of gray scale values or by adjusting the lightness gradually with a modulator (e.g., Takasaki, 1966; Ekroll and Faul, 2009). In Figure 1.4A the match is the upper central square and the target is the framed square in the checkerboard. It is assumed that the observer internally compares the magnitude of the perceived lightness for the target and the match, each of which is based on a particular internal scale for that viewing condition. The two graphs next to the samples represent the hypothetical internal scales of an observer.

However, these scales cannot be measured with matching procedures. The graphs in Figure 1.4A serve as an illustration of the assumed processes that an observer performs in matching procedures. The luminance of the target, x_T, in Figure 1.4A evokes a certain perceptual response labeled as $\psi^T(x_T)$. The observer, then, adjusts the luminance of the match, x_M, in such a way that they perceive $\psi^P(x_M)$, the perceptual response of the match, as equal to $\psi^T(x_T)$. The procedure is repeated using different luminance values for the target. The samples may be presented in front of the same as well as different background conditions. The former design is called symmetric matching and the latter asymmetric matching. Usually, asymmetric matching is used, as it reveals more about the relation between different viewing conditions and perceptual constancy. The data of the matching experiment (matching data) are displayed by plotting the luminance values of the match against the luminance values of the target (as in Figure 1.4B). Based on these matching data, we are able to study lightness constancy. In a symmetric matching experiment, the data would be displayed as a line through origin. Conversely, if the samples are presented in different viewing conditions, the line might be shifted and it would be possible to study lightness constancy.

In order to model an observer's perception, a perceptual scale would be needed (Kingdom & Prins, 2016). However, matching procedures only obtain the relationship between two luminance values (match and target) instead of a perceptual scale. The matching data shows the relation between two equally perceived luminance values instead of the relation between luminance and perceived lightness. The matching data still reflects the relation between lightness and luminance, as the adjustment process is internally based on perceived lightness. However, two experiments with completely different viewing conditions due to an interposed transparency could lead to the same matching data.

1.3.2 Scaling Methods

As mentioned above, perceptual scales relate luminance and the corresponding perceived lightness. These scales differ among observers and also vary among viewing conditions – e.g., sunlight or shadow – for a single observer. With each of these scales, it is possible to predict an observer's perception in a particular viewing condition.

Scaling procedures are an approach to measure lightness and assess such perceptual scales. Several experimental methods have been developed to assess these scales. Whittle (1992), for instance, used equisection scaling. In his data, he also reported the appearance of CE. Despite using a scaling method, I observe that his task was also based on adjustment procedures. Therefore, the focus in Chapter 3 and 4 is on two more recent methods called Maximum Likelihood Difference Scaling (MLDS), developed by Maloney and Yang (2003) and Maximum Likelihood Conjoint Measurement (MLCM), developed by Knoblauch and Maloney (2012) as their tasks are based on difference judgments.

In the following, I first present how Whittle (1992) implemented equisection scaling and his results regarding CE. Then, I briefly explain MLDS and MLCM to emphasize the difference between equisection scaling and these two methods.

Equisection Scaling

In equisection scaling, several targets have to be adjusted between two fixed outer limits. The targets in between shall form an equidistant series in lightness space. Figure 1.5 provides an example of the setup of the equisection scaling experiment that was conducted by Whittle

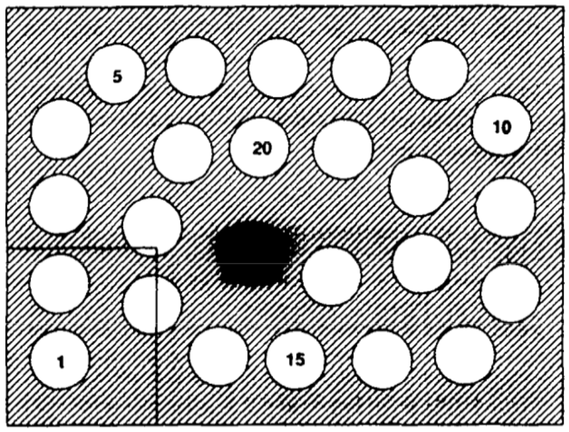

Figure 1.5: Set up for the equisection scaling experiment conducted by Whittle. The central patch was black and the first patch (bottom left) was white. In the experiment, the observers had to adjust the remaining white patches to be equidistant in lightness space. The figure is taken from Whittle, 1992.

(1992). In Whittle's experiment, Patch number 1 was fixed to white and the central patch was fixed to black. The remaining 24 patches had to be adjusted in order to form a series comprising equidistant gray patches.

Statistical Methods: MLDS & MLCM

MLDS and MLCM are two more recent scaling methods based on difference judgments introduced by L. Maloney and Yang (2003), Knoblauch and Maloney (2012). The goal of MLDS is to assess a psychophysical scale by setting up a "stochastic model [a generalized linear model] of difference judgments, which allow maximum likelihood estimation of the underlying perceptual scale" (Aguilar & Maertens, 2020). In MLDS, the method of triads or the method of quadruples is used, where three or four ordered stimuli, respectively, are presented. In this thesis, I focus on the method of triads where the observer has to compare the difference between the two outer stimuli with the stimulus in between. The observer has to judge which of the two outer stimuli is more different from the middle one. The exact procedure is described in Chapter 4.1.2. MLDS reliably estimates the perceptual scales within one viewing condition and promises to be more robust than other scaling methods (Aguilar & Maertens, 2020). All the three elements of one triad are always presented in the same viewing condition. Aguilar and Maertens (2020), among others, scrutinize problems that might arise from this implementation. The experiment of Aguilar and Maertens (2020) is described in more detail in Chapter 4.

MLCM has been developed as an extension of MLDS in order to model the effect of more than one stimulus dimension (Aguilar & Maertens, 2020). In this method, the perceptual scale of one context is set as a reference scale, and the perceptual scales measured in other contexts are assessed relative to the reference scale. In both methods, the judgments are based on the perceived difference between the stimuli and not on the adjustment of the stimuli as in matching or equisection scaling in Whittle (1992) (Wiebel, Aguilar, & Maertens, 2017).

1.4 Research Question & Motivation

In previous work, CE was measured using matching procedures or equisection scaling. Although equisection scaling creates a perceptual scale, the task of the method can be similar to matching procedures. In the experiment in Whittle (1992), the observers had to adjust 24 simultaneously presented patches. Therefore, both procedures that have observed CE are based on adjustment tasks. It remains open whether CE can be observed solely in the data obtained using methods based on adjustment tasks, or whether CE could also be found in the perceptual scales measured with other tasks. The former implies that it is only the expression of a particular way of measuring perceived lightness, whereas the latter indicates that the effect is an internal processing characteristic of the visual system.

This thesis aims to reanalyze perceptual scales measured with MLDS and MLCM with regard to the occurrence of CE. I reanalyze existing data collected in a former experiment by Aguilar and Maertens (2020) to investigate whether the effect could also be seen in these scales. Additionally, I analyze the slope of the perceptual scales and I develop a distance metric as potential indicators of CE. Apart from that, I simulate an asymmetric matching experiment and an MLDS experiment to investigate the influence of differently distributed luminance values on perceptual scales.

In this thesis, I first give a short overview of previous works on CE. In Chapter 3, I simulate an experiment in order to investigate the effect of differently spaced luminance values on perceptual scales. Afterward, I describe the experiment of Aguilar and Maertens (2020) in Chapter 4 and reanalyze these data. Finally, I discuss my results and suggest some experiments for future research.

Chapter 2

The Crispening Effect in Previous Work

In this chapter, I review the literature on CE. I focus on the different methods and tasks that were used in experiments in which CE occurred. Most importantly, I examine the experimental methods in order to determine whether CE is an artifact of a method or a feature of the visual system. Additionally, I briefly describe the conclusions that experimenters have drawn regarding the mechanisms potentially involved in CE.

The “Crispening Effect” was named by Takasaki, 1966. As described above, Figure 1.3 illustrates the phenomenon. CE is an example of a phenomenon where the surrounding influences an object’s perception. Although the physical luminance values are equally spaced, the perceived lightness difference between the decrements (increments) is not. This change in perceived lightness dependent on the background is called the Crispening Effect. In matching data (e.g. Takasaki, 1966) or in the perceptual scales obtained with equisection scaling (Whittle, 1992), CE is observed as non-linearity and a steepening of the slope around the background luminance (c.f. Figure 2.4). The effect has been reported in experiments using mainly two types of psychophysical methods – asymmetric matching and equisection scaling – which have briefly been described in Section 1.3.1 and Section 1.3.2, respectively. In the following, I review these findings separately.

2.1 Asymmetric Matching

As mentioned above, asymmetric matching is one of the predominant methods for measuring perceived lightness. CE occurred several times in the data recorded with this method. First, I describe the experiment of Takasaki (1966). Second, I refer to the previous work by Ekroll et al. who investigated CE in color space.

Figure 2.1: Simultaneous lightness contrast: The two squares in the middle have the same luminance value. However, the left sample is perceived as lighter than the right sample.

2.1.1 Takasaki's Work on CE

The first that referred to the phenomenon as CE was Takasaki (1966). Takasaki used the method of asymmetric matching. He presented two gray samples in front of two different homogeneous gray backgrounds – a so-called center-surround (CS) stimulus. One of the two samples was the fixed target, while the other was the match which the observer had to adjust in such a way that both samples were perceived equally in lightness space. Each observer did several adjustment series with different background pairs. When plotting the matching data, the luminance or reflectance values of the matching sample are plotted on the y-axis against those of the target sample on the x-axis. A straight line through origin, that has 45° to the axis, is expected for the samples which are embedded in the same surrounding. For different background conditions, this line is expected to be shifted due to the following effect.

A gray field in a darker surround is perceived as lighter than a field with the same luminance value in a lighter surround. Figure 2.1 shows an illustrative example of this differing perception, the simultaneous lightness contrast. The effect is described in previous works by, for instance, Bressan and Actis-Grosso, 2006; Soranzo, Lugrin, and Wilson, 2013. In a matching experiment, the observer has to assign a significantly higher luminance value to a square in front of a lighter background in order to perceive it equally light as the one embedded in the dark background (Takasaki, 1966). Owing to this, the line of the plotted matching data is shifted (c.f. Figure 2.2).

Figure 2.2 shows the matching data of Takasaki (1966). In order to indicate the uncertainty of the matches, Takasaki drew pairs of symbols (e.g., x, ○) – one referring to a sample slightly lighter and the other to a sample slightly darker than the actual match. The actual matching data is a line in the middle of these two symbols. As mentioned above, Takasaki reported non-linearity in the matching data of his experiment. Besides the expected shift of the line, a "hump" in the curve was also visible.

The author stated that "the hump must be caused by the relation between the reflectance of the samples and backgrounds" because its magnitude and position change when changing the background. He also developed the Equation 2.1 which fits, according to him, well with the data. Takasaki introduced three user-specific constants, C_i with $i \in \{1, 2, 3\}$, for adjusting the amount of induction and CE and the reciprocal sharpness of CE. Like this, the equation is

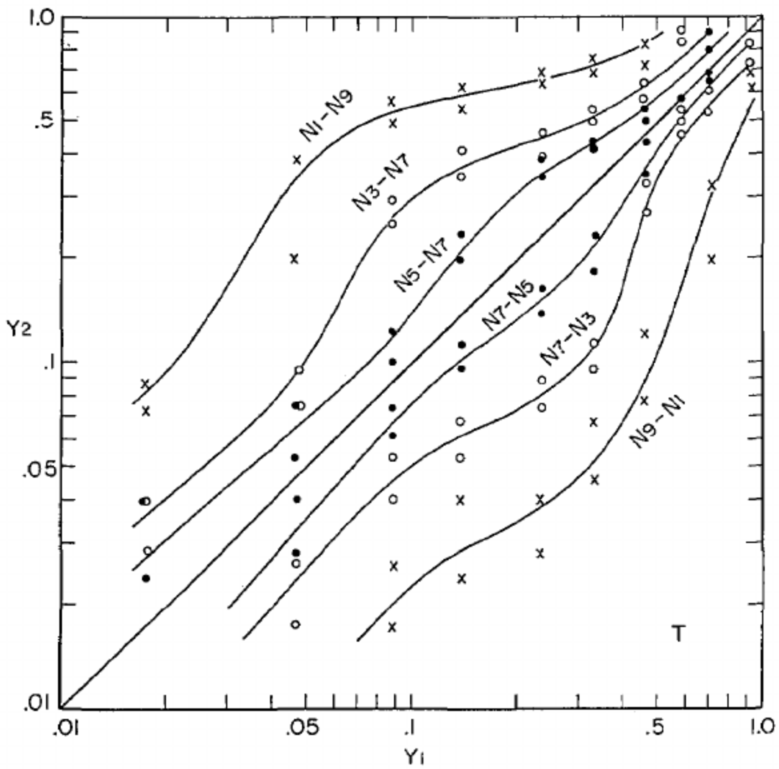

Figure 2.2: Results of Takasaki, 1966 showing CE. The x and y axes show the reflectance values of the test and match samples, respectively. For symmetric matching, a 45° line through origin is expected. For asymmetric matching between different background pairs, the line is expected to be shifted vertically. In addition to the shift, we can also observe a "hump" in the results. Takasaki (1966) describes this hump as the Crispening Effect. The figure is taken from Takasaki, 1966.

adjustable for each observer individually and the magnitude of CE can be changed.

$$L = V - C_1 V_b + f(V, V_b, \bar{V}_b) \tag{2.1}$$

$$\text{with} f(V, V_b, \bar{V}_b) = C_2 \bar{V}_b \left[\frac{(V - V_b)}{C_3} \right] e^{-\frac{|V - V_b|}{C_3}} \tag{2.2}$$

Due to the user-specific constants, the formula is different for each observer. This finding is also supported by the research presented in Section 2.1.2.

L is the lightness, while V and V_b are the Munsell values of the sample and the background, respectively. $\bar{V}_b$ is the average Munsell value of the two backgrounds. Thus, to calculate the lightness values for the two examined samples, L_1 (L_2), we have to insert V_1 (V_2) for V and V_{b1} (V_{b2}) for V_b. The exact derivation of the formula can be found in Takasaki (1966).

2.1.2 Ekroll et al.'s Work on CE

As in Takasaki (1966), Ekroll et al. used asymmetric matching in several experiments that observed CE (Ekroll, Faul, & Niederee, 2004; Ekroll & Faul, 2009; Ekroll, Faul, & Wendt, 2011). In contrast to Takasaki, these experiments were conducted in color space. They focused on understanding the underlying mechanisms of CE, the gamut expansion effect, and the simultaneous contrast effect. They also developed a model to describe the different effects. Additionally, they reported that CE appears mainly while center-surround (CS) stimuli are used with a homogeneous background; they also reported that the change to a variegated background actually prevents the effect. They argued that complex processing mechanisms may be involved in the perception of CS stimuli, despite the "simple" color appearance changes.

Similar to Takasaki (1966), Ekroll et al. used CS stimuli in their experiments. In the experiments, the observers adjusted two samples in such a way that their chromaticity is perceived equally. Apparently, the CS stimuli evoked the impression of a transparent medium in front of a homogeneous background. This transparency impression was the clearest for the stimuli where the contrast between target and surround was very low. According to Masin and Idone (1981) (after Ekroll et al., 2004), the likelihood of perceiving any transparency is inversely related to the contrast between the target and the surround. CS stimuli with a high level of contrast did not cause any impression of transparency. This implies that CS stimuli with a low level of contrast between the target and the background are processed differently than those with high contrast.

This line of thought is also supported by non-linear matching data of Smith and Pokorny (1996), thereby indicating that the test field was influenced by adaptation to the surround. According to Ekroll et al. (2004), the matching data of Smith and Pokorny (1996) can be interpreted in such a way that, even for a fixed surround, different perceptual mechanisms are active in the visual processing depending on the contrast's magnitude. As pointed out by Evans (1974) (after Ekroll et al., 2004), homogeneous surfaces are "close to being contrary to the laws of nature." Dense fog is one of the only occasions when we perceive homogeneous surfaces in our environment. This may explain the complex mechanisms involved in the perception of CS stimuli (Ekroll et al., 2004).

The transparency impression complicates or even impedes the creation of satisfying asymmetric matches. In this case, the observers can only try to realize a match as good as possible. Therefore, it is important to concentrate on the development of psychophysical methods designed for measuring internal perceptual processes (Ekroll & Faul, 2013). Additionally, it is important to scrutinize whether the method of matching also causes other phenomena – for instance, CE – or whether these phenomena are a feature of the visual system.

Apart from raising questions concerning the validity of the traditional methods for measuring color appearance, Ekroll et al. analyze and compare three effects, CE, the simultaneous contrast effect, and the gamut expansion effect regarding their underlying mechanism. The simultaneous contrast in the achromatic space is described above in Section 2.1.1 and is visualized in Figure 2.1. In a color space, the effect is very similar. The two samples have the same nominal gray; but, instead of a gray background, the backgrounds are colored. The two samples seem to be tinged with a color complementary to their respective background colors.

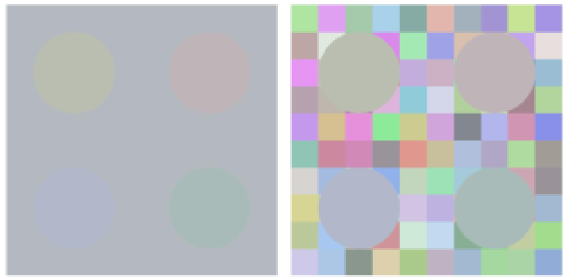

Figure 2.3: Gamut expansion effect: The four circles on the left and on the right sides are physically equal. However, the circles embedded in the homogeneous gray surround appear more saturated than the others. The figure is taken from Ekroll, Faul, and Wendt, 2011)

Figure 2.3 shows the gamut expansion effect. This effect was described in color space by Brown and MacLeod (1997). The effect describes the phenomenon that the samples embedded in a homogeneous gray background appear more saturated than those which are presented in front of a variegated background.

The gamut expansion is not linear. It is maximal for targets with a low level of contrast; it decreases toward zero, the more the difference between the target and the background increases (Ekroll & Faul, 2009). The data for the gamut expansion resembles the data for CE in achromatic stimuli, as the discriminability is maximal where the color of the targets is the closest to the surround. Following this line of thought, Ekroll et al. (2011) conducted experiments in order to confirm the hypothesis that the gamut expansion effect and CE share the same underlying mechanism. They investigated the effects' strengths across different observers in order to show the correlation between the gamut expansion effect and CE. If the two observable phenomena shared the same mechanism and the strength of this mechanism varied across the observers, the strengths of the observable phenomena would also be expected to vary across the observers. Indeed, Ekroll et al. (2011) confirm that the amount of CE correlates with the amount of the gamut expansion effect across the observers.

Furthermore, Ekroll and Faul (2009) argued that the simultaneous contrast shares the same underlying mechanism as CE and the gamut expansion. CE seems to be able to describe the general characteristics of the simultaneous contrast. In order to model the simultaneous contrast, they implemented two distinct mechanisms – a time-dependent component (von Kries adaptation von Kries, 1905) and an instantaneous spatial mechanism for which they used the CE term as in Equation 2.1. All free parameters of the CE term are set individually for each observer. They chose the CE term for the instantaneous mechanism because "phenomena characteristic of crispening such as gamut expansion effect or Meyer's effect are typically observable immediately upon inspection of a demonstration" (Ekroll & Faul, 2009).

2.2 Equisection Scaling

Ekroll et al. and Takasaki used the method of asymmetric matching in their experiments. Whittle (1992) also reported the occurrence of CE when using an equisection scaling procedure (c.f. Section 1.3.2). The observers had to adjust 24 patches in such a way that they built a range of luminance patches which are perceived as equidistantly in lightness space (c.f. Figure 1.5). The observers could select the luminance values from a palette of 256 grays. The experimenters explicitly warned the observers against the risk of using the distance of the regularly arranged gray values in the palette as a reference. Whittle used different background conditions – black, gray, and white – in his experiment.

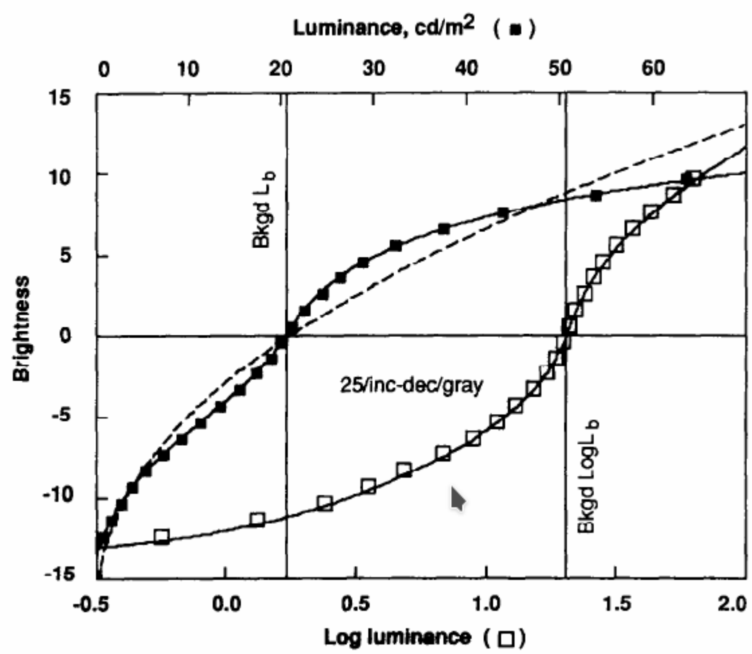

Figure 2.4: Results of Whittle, 1992. Perceived lightness (here called brightness) is plotted against luminance ($[cd/m^2]$, solid squares) and log luminance (open squares). CE can be seen as a steepening of the curve approximately at a luminance of 21 cd/m^2 (Bkgd L_b) and a log luminance of 1.3 (Bkgd $LogL_b$). The dashed line illustrates a hypothetical scale without CE according to the best fitting power law.

The scale for the gray background condition can be seen in Figure 2.4. The solid squares show the data for perceived lightness (here called brightness) against luminance ($[cd/m^2]$), and the empty squares show the same data against log luminance. In the figure, CE can be observed as a steepening of the curve around the background luminance, marked as "Bkgd L_b" and "Bkgd $LogL_b$". The dashed line represents a hypothetical scale without CE according to the best fitting power law $a + bL^n$, with $a = -17.3$, $b = 4.23$ and $n = 0.46$.

Figure 2.5 shows the scales for other background conditions. As in the previous figure,

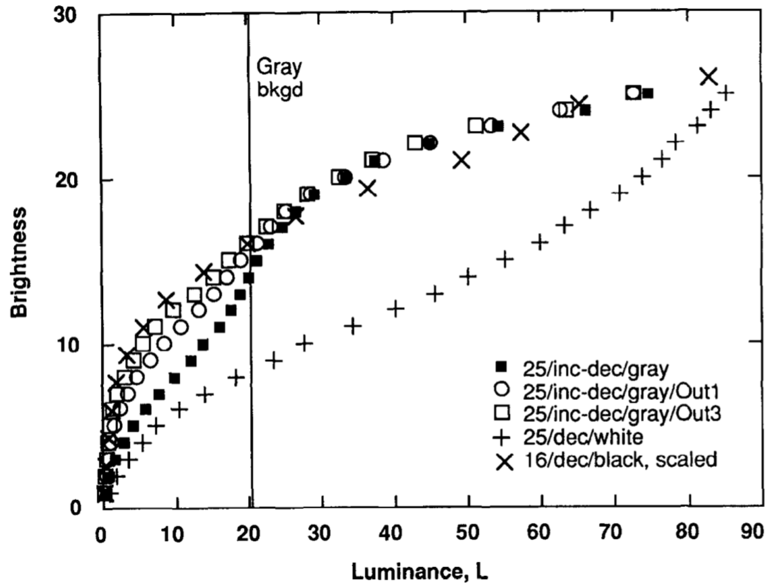

Figure 2.5: Results of Whittle, 1992. The perceived lightness (here called brightness) of different experimental background conditions is plotted against the luminance, $[cd/m^2]$. CE can be seen as a steepening of the curve at the respective background luminances – e.g., for gray: approximately 21 cd/m^2 (Bkgd L$_b$) and a log luminance of 1.3 (Bkgd LogL$_b$). The solid squares represent the data for the gray background, whereas the empty squares and empty circles represent an adaptation of this condition where a black frame was added around the patches. The x and + represent the data measured with the white and black backgrounds, respectively, as labeled in the graph.

the solid squares represent the gray background condition. The scales measured with black and white backgrounds are represented by x and +, respectively. For these two background conditions, we can also observe a steepening of the curve around the black and white background luminance – once at approximately 0.13 cd/m^2 (x) and at other time at 87 cd/m^2 (+). Additionally, Whittle (1992) found that adding a black outline around the targets in the gray surround reduces CE significantly. He added a black frame of one or three pixels around the patches. The empty circles and empty squares, respectively, show these data in Figure 2.5. As we can see, CE is no longer observable at 21 cd/m^2. However, the scales of this viewing condition resemble the scale of the black background condition. Therefore, one may argue that CE has just shifted due to the adjacent black background.

Similar to Takasaki (1966), Whittle developed a formula that mathematically described such data.

$$\text{Brightness} = a \log(1 + bW) \tag{2.3}$$

$$\text{with: } W \text{ for decrements} = \frac{(1-k)\Delta L}{l + L_d + k\Delta L} \tag{2.4}$$

$$W \text{ for increments} = \frac{\Delta L}{L_b + L_d} \tag{2.5}$$

L is the luminance, L_b is the background luminance, and ΔL represents the brightness. “k is the proportion of scattered light into a decrement”(Whittle, 1992). L_d is a small constant, set to 0.39, which allows to generalize W when $L_{min} = 0$. When inserting a = -7.07 (decrements) and a = 8.22 (increments) and b = 6.58 (both) in Equation 2.3, the continuous lines in Figure 2.4 are generated (Whittle, 1992).

2.3 Open Questions

The CE might occur only when methods are used which are based on adjustment tasks. Whittle (1992) uses equisection scaling, which can be called scaling method; however, the experimental task is based on adjustment procedures and therefore similar to matching experiments. The occurrence of CE might be a consequence of internal adjustment procedures. However, not all papers in which an adjustment task is used observe CE. Several reasons might be responsible for that. First, as Whittle (1992) reported, an outline around the patches reduces or eliminates CE. Another reason could be the reduction or even abolishment of CE when the targets are embedded in a variegated background (Ekroll et al., 2004; Ekroll & Faul, 2009; Faul, Ekroll, & Wendt, 2008). Furthermore, the individual differences between the observers seem to be significant. As described in Ekroll and Faul, 2009, some of the observers seem to be almost immune against the effect. Experiments investigating CE or the simultaneous contrast are typically performed with only a few observers. For this reason, it is difficult to estimate the individual variance (Ekroll & Faul, 2009).

Author, Year	Domain	Method	Task
Takasaki, 1966	Achromatic	Asymmetric matching	Adjustment
Whittle, 1992	Achromatic	Equisection scaling	Adjustment
Ekroll, 2004, 2009, 2011	Color	Asymmetric matching	Adjustment

Table 2.1: Comparison of previous works regarding their domain, the used methods, and the kind of imposed task.

In this chapter, I presented some representative previous works by Takasaki (1966), Whittle (1992), Ekroll et al. (2004), Ekroll and Faul (2009), Ekroll et al. (2011) which described CE and compared it with other effects such as the gamut expansion effect and the simultaneous contrast effect. Table 2.1 shows a summary of the methods and tasks of the previous works on CE. It is notable that all the experiments are based on adjustment tasks. Therefore, it is not evident whether CE occurs only in experiments based on adjustment tasks or not. As Whittle (1992) used equisection scaling and observed CE, one may argue that there is evidence to suggest that CE is not caused by the method of matching. However, his task was also based on adjustment procedures, as the observers had to adjust the luminance of the ordered stimuli and all the gray patches were presented simultaneously in one palette. Therefore, it is possible that the visual processes involved in Whittle's experiment are more similar to those being involved in asymmetric matching, despite referring to equisection scaling as a scaling method.

Chapter 3

Simulations

Before reanalyzing the experimental data of Aguilar and Maertens (2020), I investigate how different distributions of the luminance values (spacings) affect perceptual scales. In adjustment procedures, the observers perform the adjustment task in a small luminance range. In contrast, scaling procedures are normally realized with luminance values distributed across the whole luminance range. Additionally, in statistical scaling procedures it is not possible to use too many luminance values, as the amount of necessary trials, which the observers have to realize, increases rapidly with an increase in the number of luminance values. Section 4.1.2 describes an example of the calculation of necessary trials with 10 luminance values. Methods based on adjustment tasks can use more data points within a luminance range, because not all luminance values have to be compared with each other. Therefore, these methods need less trials than scaling procedures.

Scales with and without CE should be distinguishable, as CE shows as a steepening of the slope around the background luminance. However, it is not clear whether a coarse spacing of the data points influences the discriminability of scales with or without CE, as only a few data points are distributed across the whole luminance range. It might be that few data points cause that the steepening of the slope cannot recovered. In order to clarify whether the scales with and without CE are always distinguishable, I simulate an experiment with a hypothetical observer based on two different ground truth functions. For the ground truth function with CE, I use Whittle's formula (c.f. Formula 2.3). For the ground truth function without CE, I implement the best-fitting power law

$$\psi(L) = a + bL^n \text{ with } a \approx -1.93, b \approx 4.16, n \approx 0.45 \tag{3.1}$$

I obtain these values by fitting a function to the digitized data points of the dashed line from Whittle's data in Figure 2.4.

I simulate an asymmetric matching experiment and an MLDS experiment based on these ground truth functions. The asymmetric matching experiment serves as a validation of this simulation method. As CE occurs in the matching data of Takasaki (1966) and Ekroll et al., I expect CE to also occur in my simulated matching data. After the validation, I simulate an MLDS experiment as I will reanalyze data obtained with this method by Aguilar and Maertens (2020) in Chapter 4. Hence, I want to scrutinize the effect of different spacings on the occurrence of CE beforehand. I choose three different spacings for the simulation – with evenly distributed luminance values, with luminance values centered around the background (centbg), and with luminance values far away from the background luminance in luminance space. The last type is called coarse spacing. After the simulation, I implement two different metrics to evaluate the similarity of the MLDS scales to the ground truth functions and the discriminability of the scales measured with MLDS. Additionally, I analyze the slope of the MLDS scales to investigate whether the slope could serve as an indicator for CE.

All programs used in this simulation and the corresponding data and plots are uploaded in GitHub[1].

3.1 Simulation of an Asymmetric Matching Experiment

In the simulation of the asymmetric matching experiment, I simulate the match's perception based on the ground truth without CE and the target's perception based on the ground truth with CE. In case of a realistic experiment, this condition can be created by presenting the match within a variegated checkerboard and the target in front of a homogeneous background (e.g., Ekroll et al., 2004; Ekroll and Faul, 2009).

I use three background luminances: 50 cd/m^2 (light gray), 20.56 cd/m^2 (mid gray), and 5 cd/m^2 (dark gray). Equation 2.3) is based on an experiment by Whittle (1992) that only used the mid gray background. It is not certain whether inserting light and dark gray values for the background luminance in Equation 2.3 leads to real perceptual scales of observers. This should be taken into account when interpreting the simulated matching data.

As mentioned above, I modify the spacing of the data points apart from varying the background luminance. I use an evenly distributed, a coarse and a centered spacing.

3.1.1 Simulation Procedure

The ground truth functions serve as hypothetical mapping functions between luminance and perceived lightness, as illustrated, for instance, in Figure 1.4A. Figure 3.1 shows the simulation concept of an asymmetric matching task. I first calculate the perceived lightness value $\psi(l_t)$ for the corresponding target luminance, l_t, by inserting l_t into Formula 2.3. Next, I add normally distributed noise (with $\sigma = 0.04$) to the internal dimension $\psi(l_t)$. When using a noise with σ higher than 0.05, the standard deviation becomes too big to make a reliable statement about the matching data. In the next step, I set $\psi(l_m)$ (perceived lightness of the match) to be very close[2] to the "noisy" $\psi(l_t)$ in such a way that $\psi(l_m)$ = $\psi(l_t)$+noise. $\psi(l_t)$ was generated based on l_t. Reversely, the match's luminance, l_m, is now generated for the corresponding $\psi(l_m)$

[1] https://github.com/computational-psychology/brunn_ba2020

[2] In the implementation, I calculated 10 000 perceived lightness values corresponding to a set of 10 000 luminance values, which were evenly distributed across the luminance range. Consequently, $\psi(l_m)$ might not be equal to $\psi(l_t)$, but it just might be very close to the actual value.

based on the power function without CE. This corresponds to

$$l_m = \sqrt[n]{\frac{\psi(l_t) - a}{b}}$$

with a, b, and n as specified above in the introduction of Chapter 3[3]. I repeat this procedure with 10 different luminance values for the target. Each lightness estimation is repeated 100 times, and I take the average across the results of the different runs to calculate the matching data. I repeat the simulation for the three different spacings of the luminance values.

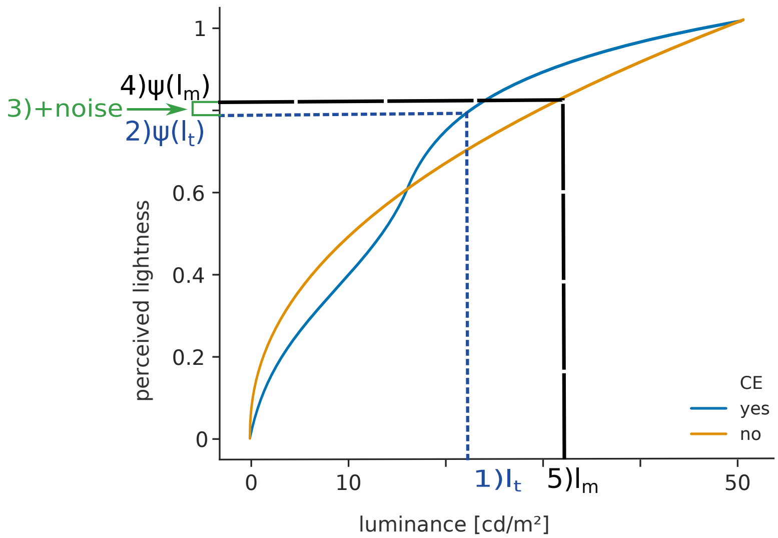

Figure 3.1: Concept of an asymmetric matching simulation. 1) A target luminance is set (l_t). 2) The corresponding perceived lightness $\psi(l_t)$ is estimated based on the function with CE. 3) Random noise is added to $\psi(l_t)$. 4) $\psi(l_m)$, the match's lightness, is set to this value in such a way that $\psi(l_m)$ = $\psi(l_t)$+noise. 5) The corresponding match luminance, l_m, for $\psi(l_m)$ is estimated based on the function without CE. The functions in this figure are not the ground truth functions with the parameters specified above, but they are adapted to illustrate the procedure in a better way.

3.1.2 Results of an Asymmetric Matching Experiment

Figure 3.2 shows the simulated matching data for different spacing schemes where the match's background was mid gray. The match luminance (y-axis) is plotted against the target luminance (x-axis). The diagonal line through origin refers to how the data would look in a symmetric matching experiment if both target and match were presented in the same context

[3] In the implementation, I used the luminance value at the corresponding index of $\psi(l_m)$ because I already calculated these values beforehand. Since I used 10 000 values in the range of $[0, 65]$, the value should be almost identical to a calculated value.

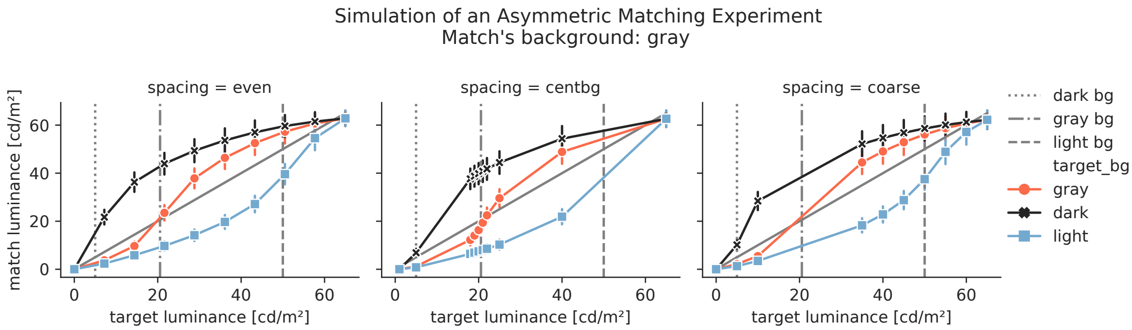

Figure 3.2: Asymmetric matching data of simulation with different spacings and a mid gray background of the match. The luminance values of the target's background are indicated with vertical gray lines. The black, red, and blue lines show the data for a dark, gray, and light target background, respectively. The diagonal gray line is the reference of the data in a symmetric matching experiment. The position of the non-linearity of the matching data varies according to the target's background luminance.

regarding the background luminance and arrangement. The black, red and blue line show the data for the target presented in front of light, mid gray, and dark backgrounds, respectively. The vertical lines at the markers show the standard deviation of the matching data over 100 runs.

As expected, non-linearity occurs in the asymmetric matching data. The position of the non-linearity varies depending on the target's background luminance. This was also expected because CE occurs as a steepening of the slope close to the background luminance. The target's perception was based on a ground truth function with CE, hence the position of the non-linearity depends on the target's background. The non-linearity is always observable, independent of the spacing of the data points.

The matching data is independent of the match's background luminance because the simulation of the match's perception is based on the best fitting power law without CE (c.f. Equation 3.1). This equation does not consider the background luminance, hence, the matching data should not change when adjusting the match's background luminance.

Figure 3.3 shows the matching data for the same experiment but where the match's background was dark gray. The plot design and legends are the same as in Figure 3.2. Is is noticeable, that the matching data looks like the data in Figure 3.2 and is not influenced by the different luminance of the match's background. For this reason and because CE is observed in the simulated matching data, the simulation method seems appropriate.

I also simulated the asymmetric matching experiment with three other spacings. However, the three spacings presented here show the clearest tendencies. Therefore, the other spacings are only presented as additional information in the appendix in Section .1.1. I also simulated matching data with different noise levels with $\sigma = 0.02$ and $\sigma = 0.08$. However, they do not differ significantly from the data presented in Figure 3.2 and Figure 3.3 apart from a higher standard deviation. Therefore, these data are not shown in this thesis but can be easily

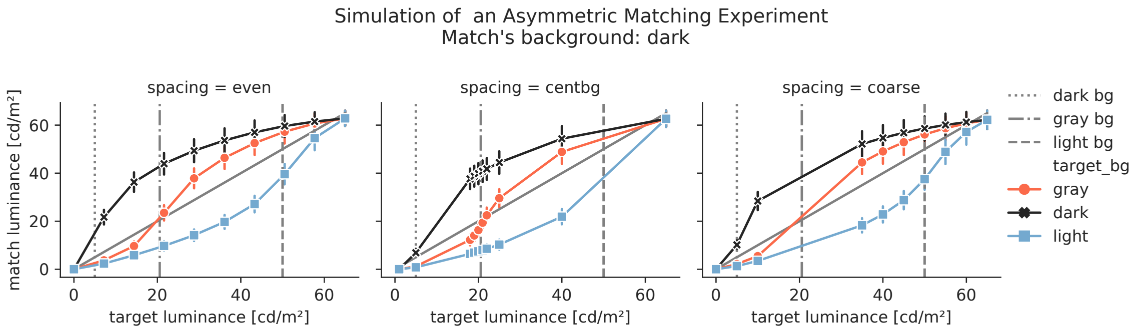

Figure 3.3: Asymmetric matching data of simulation with different spacings and a dark gray background of the match. The plot design and the legends are the same as in Figure 3.2. The matching data looks like the matching data in Figure 3.2. The data are not influenced by the different luminance of the match's background.

generated by changing the noise constant in `simu_constants.py` in the simulation programs[4].

3.1.3 Results of a "Symmetric" Matching Experiment

For a full comparison, I also simulated a matching experiment where target and match were both based on a ground truth function with CE. In a realistic experiment, this would be the case if both samples were presented in front of a homogeneous background. Technically, the experiment is also an asymmetric matching experiment if the luminance of the homogeneous backgrounds differ. However, to distinguish the two simulations I call this simulation "symmetric" matching experiment because the background pattern would be the same in a realistic experiment. Apart from using a ground truth function with CE for both target and match, the simulation procedure was the same as described above in Section 3.1.1.

I varied the luminance of the homogeneous background for both, target and match. In contrast to the asymmetric matching experiment above, the match's background luminance influences the matching data. Figure 3.4 shows the matching data for the even spacing of the luminance values. When target and match are both presented in front of the same homogeneous background, the non-linearity disappears. For instance, the left plot in Figure 3.4 shows the data obtained for a mid gray match background. The red line which is the matching data for a mid gray target background is a straight 45° line through origin. This is expected for an actual symmetric matching experiment. The black and blue line show the matching data with the same match background but a dark and light target background, respectively. These data show a non-linearity at the respective target background luminance.

The same pattern can be observed in the other two plots. The symmetric matching data, where the luminance of the match and target background is the same, is a straight line. The data for the other two asymmetric conditions show non-linearities. Accordingly, when the backgrounds of target and match differ in lightness, we can observe a type of non-linearity similar

[4]https://github.com/computational-psychology/brunn_ba2020/tree/main/programs_simulation

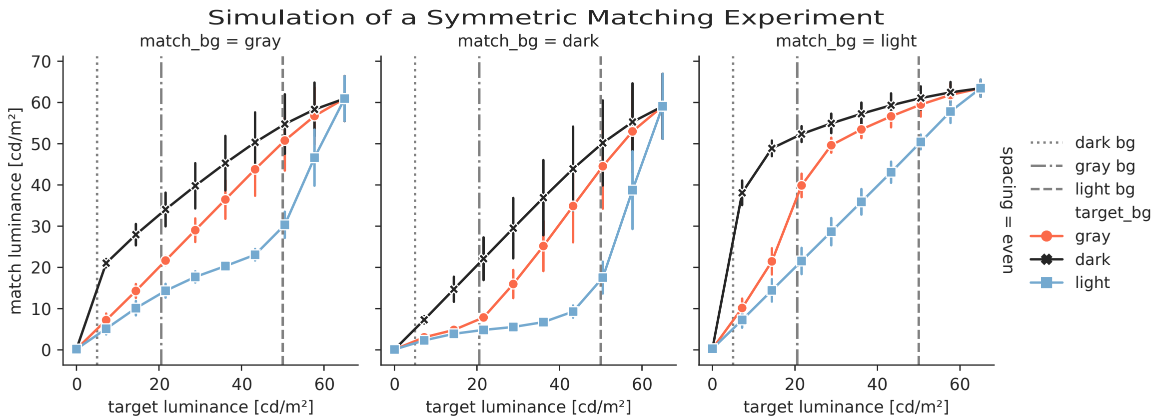

Figure 3.4: "Symmetric" matching data of simulation with even spacing. The plot design and the legends are the same as in Figure 3.2. The data are influenced by the different luminance of the match's background. The left, middle, and right panel show the matching data obtained with a mid gray, dark gray, and light gray match background, respectively. The matching data, where the luminance of the target's and match's background is the same, is a straight 45° line through origin. The other two lines show non-linearities at the respective background luminance.

to Takasaki's data in Figure 2.2. Figure 2 in the appendix shows the symmetric matching data for other spacings.

3.2 Simulation of an MLDS Experiment

In the next step, I simulate a scaling experiment since the previous simulation of an asymmetric matching experiment shows the expected results. I simulate an experiment with differently distributed luminance value to investigate the influence of the spacing on the perceptual scales. As a scaling method I use MLDS because this is one of the methods that was used to measure the perceptual scales which I reanalyze later. In the simulation, I calculate the perceived lightness for the luminance values with two different ground truth functions – either with CE (c.f. Equation 2.3) or without it (c.f. Equation 3.1). In MLDS, all targets are always presented in the same context. Therefore, I calculate the lightness perception of all the data points of one triad based on the same ground truth function. In this simulation, I only use the mid gray background luminance (20.56 cd/m^2) as the original data, for which Whittle's formula (c.f. Equation 2.3) was defined, was obtained with this background luminance. As in the previous simulation, I vary the spacing of the data points in luminance space.

In previous work, CE shows as a steepening of the slope around the background luminance (c.f. Figure 2.4). Due to this steepening, scales with and without CE should be discriminable as, for instance, the dashed line and the filled squares in Figure 2.4. In Section 3.2.4, I analyze whether the resulting perceptual scales, estimated with different ground truth functions, are always distinguishable, despite a particular spacing of the data points. Too few data points close to the background luminance might impede the recovery of the slope's steepening due to CE. It might be that data points close to the background luminance are needed to distinguish the scales base on this steepening. Apart from judging the ability to discriminate the scales, I also evaluate the similarity between the resulting perceptual scales and the ground truth functions in Section 3.2.3.

3.2.1 Simulation Procedure

The exact procedure of MLDS judgments is described below in Section 4.1.2. In short, the observers have to consider three targets – x_1, x_2, and x_3 – and need to compare the difference between $x_2 - x_1$ and $x_3 - x_2$. In the simulation, I compute the observer's decision variable d with

$$d = (x_3 - x_2) - (x_2 - x_1). \tag{3.2}$$

Then, I add the normally distributed noise ϵ with $\sigma = 0.1$, $\epsilon \sim N(0, \sigma^2)$, to the decision variable in order to simulate the observer's noise level. A value of $\sigma = 0.1$ corresponds to a realistic noise level for human observers (Aguilar & Maertens, 2020). If I add noise (ϵ) to the decision variable (d) I obtain Δ; $\Delta = d + \epsilon$. Based on Δ, I simulate the response variable R for the simulated observer. If $\Delta < 0$, then R = 0, indicating that the left difference ($x_2 - x_1$) was bigger. Otherwise, R = 1, indicating that the right difference ($x_3 - x_2$) was bigger. The MLDS estimation[5] is, then, performed on these sets with luminance values and the corresponding response variables. I repeat the same simulation 100 times for each spacing and then average over all runs of one spacing.

3.2.2 Results of an MLDS Experiment

Figure 3.5 shows the average perceptual scales for different luminance spacings. For a specific type of spacing, the scales of the simulation based on ground truth functions with or without CE are always plotted together in one plot. The orange crosses represent the data for the estimation based on the ground truth function without CE. The blue circles show the data for the estimation based on the ground truth function with CE. The orange and blue continuous lines show the ground truth functions. The dashed gray line indicates the background luminance.

The left panel of Figure 3.5 shows the data for the evenly distributed luminance values. Here the simulated scales lie on top of the continuous lines representing the ground truth functions. In the even spacing, the standard deviation is very small. The middle panel presents the simulated scales for the spacing with luminance values centered around the background luminance. The data based on the ground truth function with CE (blue) lies almost on top of the ground truth function. However, the scale without CE is negatively shifted along the y-axis. All the data points are below the ground truth function. The standard deviation is small for the data based on both ground truth functions. The right panel shows the data for the coarse spacing. Here, the scale with CE does not lie on top of the ground truth function. Additionally,

[5] https://www.rdocumentation.org/packages/MLDS

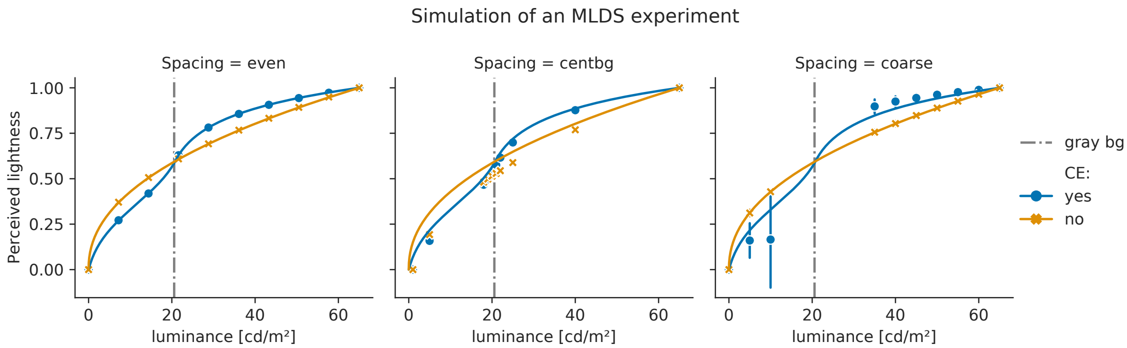

Figure 3.5: Perceptual scales obtained with MLDS in the simulation with different spacings. The perceived lightness (y-axis) is plotted against luminance (x-axis). The orange crosses and blue circles show the average perceptual scale values based on ground truth without and with CE, respectively. The continuous orange and blue lines show the corresponding ground truth functions.

the standard deviation is large in this condition. Conversely, the data without CE lie on top of the ground truth function and only have a small standard deviation.

Figure 3 in the appendix shows the data for the remaining spacings. The plots show the MLDS scales averaged over 100 runs. In the GitHub folder[6] the perceptual scale values of each individual run are saved together with the corresponding confidence intervals calculated with the MLDS bootstrap routine (`boot.mlds`[7]). The plots for the scales with the corresponding confidence intervals for the individual runs can be found in the GitHub folder or can be plotted with `simu_plot_mldsboot_runwise.py`[8].

3.2.3 Model Identification

As a metric for the influence of the different spacings on the perceptual scales, I evaluate the similarity between the resulting perceptual scales and the ground truth functions. In order to do so, I calculate the sum of the distances between the average perceptual scale values and the ground truth function. More concisely, I first calculate the 10 average scale values based on the perceptual scale values measured at the 10 luminance values of the respective spacing in the 100 simulation runs. This leads to the values $p_i(l_i)$ with $i \in \{1, \ldots, 10\}$, where l specifies the luminance values of the spacing. I calculate these values for the scales with and without CE[9]. The average perceptual scale values are identified with markers in Figure 3.5. In the next step, I select the y-values for the ground truth function (continuous lines in Figure 3.5) at the

[6] https://github.com/computational-psychology/brunn_ba2020/tree/main/data/simulation/MLDS/runs
[7] https://www.rdocumentation.org/packages/MLDS
[8] https://github.com/computational-psychology/brunn_ba2020/tree/main/programs_simulation
[9] Perceptual scale with (without) CE means a scale whose calculation was based on the ground truth function with (without) CE and does not refer to the visible effect.

luminance values of the specific spacing[10].

Thus, I have p_i for $i \in \{1, \ldots, 10\}$ for the perceptual scales and g_i for $i \in \{1, \ldots, 10\}$ for the ground truth functions. In the next step, I calculate the difference Δ_i between p_i and g_i for $i \in \{1, \ldots, 10\}$, once for the condition with CE and the other time without it. Finally, I sum up the distance values, once with and the other time without CE, so that:

$$sum_{\Delta} = \sum_{i=0}^{10} \Delta_i \text{ where } \Delta_i = |p_i - g_i| \tag{3.3}$$

I use these sums to assess the similarity of the perceptual scales to the ground truth. The fourth and fifth columns of Table 3.1 show the results of this calculation with and without CE, respectively.

In the left panel of Figure 3.5 we can see that the perceptual scale values for the evenly spaced luminance points are mostly lying on the ground truth functions and that the standard deviation is very small. Therefore, I take the summed distance of this spacing as a reference value and I compare the distances of the other spacings with this reference. The summed distance is approximately $6.07 \cdot 10^{-3}$ and approximately $9.41 \cdot 10^{-3}$ for the perceptual scale values with and without CE, respectively.

For the second spacing, where the luminance points were centered around the background luminance, the summed distance is significantly higher. For the perceptual scale based on the estimation with CE, the summed distance is approximately $304.87 \cdot 10^{-3}$; for the perceptual scale without CE, the summed distance is approximately $740.04 \cdot 10^{-3}$. The middle panel of Figure 3.5 shows the same tendency. Here the perceptual scale values without CE (orange) are all below the corresponding ground truth function. Apparently, this spacing is not suitable for recovering the exact ground truth function without CE. Possibly, more data points have to be distributed across the whole luminance range. However, despite the particular spacing, the course of the curve was recovered well only with a negative vertical shift compared to the actual ground truth function. The perceptual scale values with CE are closer to the ground truth function than in the condition without CE. Presumably, the data points around the background luminance recover the most specific characteristic, which is the slope's increase, and therefore, recover the curve well.

In the last spacing, namely *coarse*, the data points were far away from the background luminance in luminance space. The right panel of Figure 3.5 shows that the perceptual scale values with CE are different from the ground truth function. The summed distance with CE indicates the same, as it is higher than the distance for the evenly spaced data. For the perceptual scale with CE, the summed distance is approximately $395.93 \cdot 10^{-3}$; for the perceptual scale without CE, it is approximately $7.52 \cdot 10^{-3}$. The significant difference between the ground truth function with CE and the corresponding perceptual scale in this spacing implies that more data points around the background luminance are needed in order to reliably recover an underlying scale with CE. However, this spacing is suitable for recovering the ground truth function without CE. Here the summed distance is small and in Figure 3.5 the perceptual scale

[10] I calculated 10 000 perceived lightness values corresponding to a set of 10 000 luminance values, that were evenly distributed across the luminance range. Based on this set, I chose the perceptual scale value corresponding to the luminance values closest to the values of the current spacing.

Spacing	Avg_d	No.of runs below avg_d	$\sum$(distance) to ground truth	
			with CE	without CE
even (*reference*)	≈ 0.54	52	$\approx 6.07 \cdot 10^{-3}$	$\approx 9.41 \cdot 10^{-3}$
centered around bg	≈ 0.43	98	$\approx 304.87 \cdot 10^{-3}$	$\approx 740.04 \cdot 10^{-3}$
coarse	≈ 0.96	1	$\approx 395.93 \cdot 10^{-3}$	$\approx 7.52 \cdot 10^{-3}$

Table 3.1: Summary of simulation results. The first column indicates which spacing was used. The second column specifies the average distance between the scales (avg_d) across 100 runs. The third column indicates in how many runs the average distance was below the chosen reference (avg_d). The fourth and fifth columns provide the summed distance of the perceptual scale values to the corresponding ground truth function.

values lie mostly on the ground truth function with a small standard deviation.

Overall, the even spacing is the best at recovering both ground truth functions. The centered spacing also recovers the ground truth function with CE but performs badly at recovering the ground truth function without CE. Inversely, the coarse spacing recovers the ground truth function without CE very well but does not recover the function with CE. Hence, in an experiment that shall investigate CE an even spacing or a spacing with several luminance values around the background should be used.

Further developing the metric for the model identification would surely be beneficial for future work. As Table 3.1 shows, the summed distance of the coarse spacing with CE and of the centered spacing with CE do not differ significantly. However, by visual inspection of the middle and right panels in Figure 3.5, the difference between the perceptual scale values and ground truth is bigger for the coarse spacing than for the centered spacing. Most likely, the summed distances resemble each other because solely the distances at specific data points are summed and not among the whole luminance range. The metric could be refined to include the area between the scales among the whole range and not only the distance at the data points. Additionally, the variability of the values indicated by the standard deviation is neglected. The standard deviation in the coarse scales is higher than the deviation in the centered spacing. Presumably, this also influences the calculation of the summed distance.

Apart from these limitations, it was also not possible to calculate the function values based on the ground truth functions with CE (c.f. Equation 2.3) for particular spacings. As the y-values in Whittle's data (c.f. Figure 2.4) range between -15 and 15, a normalization routine was implemented. However, the calculation of the y-values directly with Equation 2.3 leads to distorted values, when using only a few values, that do not include the minimum (0) and maximum values (65), e.g., for the centered spacing. This should be further improved.

3.2.4 Evaluation of Discriminability

In the following, I analyze how well the scales based on different ground truth functions can be discriminated. As already explained in Section 3.2 the scales based on the different ground truth functions should be discriminable since the slope of the scale with CE increases at the background luminance. The two ground truth functions in Figure 2.4 are visually discriminable. Hence, as a preliminary, I state that the ground truth functions can be well discriminated. In the following, I analyze whether the simulated scales based on different ground truth functions are also distinguishable for all spacings. As a metric for the discriminability of the perceptual scales with and without CE, I calculate the average distance between the data points of the two scales. I first calculate the summed distance between the scales for each run. For this, I subtract the perceptual scale values:

$$d = \sum_{i=1}^{10} |p(i) - q(i)| \tag{3.4}$$

I do not interpolate between the perceptual scale values. I collect d for all 100 runs, and finally, average across all values in order to obtain the average distance, avg_d, to compare the distinguishability of the scales:

$$avg_d = \frac{\sum_{j=1}^{100} d_j}{100} \tag{3.5}$$

The two perceptual scale values with evenly spaced luminance values recover the ground truth functions very well (c.f. Section 3.2.3). Additionally, the standard deviation is very small in this spacing (c.f. Figure 3.5). For this reason, I use the avg_d of these data points as a reference.

For these evenly spaced luminance values, the average distance over 100 runs is approximately 0.54. Hence, 0.54 is the reference average distance for the following evaluation. This value does not indicate a degree of discriminability. Instead, an avg_d below the reference indicates that the discriminability is worse than that of the reference spacing. Conversely, a higher avg_d indicates a better discriminability. Additionally, I count in how many runs the average distance of the data points is below the reference average to determine the discriminability. For the reference spacing (evenly spaced values), the average distance is below the reference in 52 runs. This is an expected value as it corresponds to half of the runs. Therefore, less runs below the reference avg_d indicate a good discriminability, whereas more runs indicate a bad discriminability.

The second and third columns of Table 3.1 show the results of the evaluation of the discriminabilty. The second column indicates the average distance over 100 runs for each spacing, and the third column indicates in how many runs the average distance of the respective run was below the reference value 0.54.

The left panel of Figure 3.5 shows the perceptual scales for this evenly distributed spacing. In this plot, the difference of the two scales is clearly visible.

The middle panel of Figure 3.5 shows the perceptual scale for a luminance spacing where the data points are mostly centered around the background luminance. In this case, the scales are also visibly distinguishable. However, in 98 runs the average distance is below the reference value. This indicates that the scales for this spacing are mostly closer to each other than the scales for evenly spaced luminance points. Therefore, the scales can be confounded more easily. Nevertheless, the standard deviation plotted in Figure 3.5 is very small, and consequently, the two scales should be easily distinguishable in most runs.

The right panel of Figure 3.5 shows the coarse spacing where the data points are far away from the background's luminance in luminance space. For this spacing, the average distance is always higher than the reference value except for one run. This means that the perceptual scales based on the different ground truth functions are mostly "far" apart in most runs; therefore, the scales are clearly distinguishable. However, Figure 3.5 shows that the standard deviation is larger than in other spacings. This implies a large variance between runs. Consequently, it might happen that in one run the two perceptual scales do not differ.

Overall, the even spacing is the best for discriminating the scales based on different ground truth functions. According to the metric, the scales of the coarse spacing are also very well discriminable. However, owing to the visibly high standard deviation, the scales might be indistinguishable in some runs. The standard deviation is not considered in the calculation of the metric. According to the metric, the scales obtained with the centered spacing with centered luminance values are not well distinguishable. Nevertheless, in Figure 3.5 the perceptual scales can be very well visually distinguished. As the standard deviation of this spacing is very low, the scales should be discriminable in all scenarios.

In future work on this subject, it could be beneficial to include the standard deviation in the assessment of the discriminability. Currently, the distance is only calculated between the two scales of the same run. It could be useful to compare one scale based on one ground truth with the scales of all runs based on the other ground truth. Additionally, the metric could be refined to use an additional reference. For this, one could calculate the scale's avg_d with itself between the runs. Without adding any noise to the simulation, the resulting scales would be equal in each run and the avg_d between them would be zero. Owing to the added noise, the avg_d is expected to be greater than zero and would indicate a range where scales can be considered as equal and are therefore not discriminable. Distance values greater than avg_d would indicate that the two scales are different and therefore discriminable.

3.2.5 Analysis of the Slope

In the MLDS scales, CE is observed as a steepening of the slope around the background luminance. In the following, I analyze the slope of the MLDS scales to investigate whether it can be used as a metric for the occurrence of CE.

I calculate the slope values between each of the two luminance values of the particular spacing. Therefore, 10 MLDS scale values for the corresponding luminance values lead to nine slope values. For the calculation of the slope, I use the average scale values, marked in the plots, and do not interpolate between them. I also neglect the confidence intervals in calculating the slope. Hence, the slope s corresponds to:

$$s_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \text{ with } i \in \{1, \ldots, 9\} \tag{3.6}$$

where y corresponds to the MLDS scale values and x to the luminance values of the particular spacing.

Figure 3.6 shows the slope of the simulated MLDS scales. For one particular spacing, the slope of the scales based on ground truth functions with or without CE are always plotted

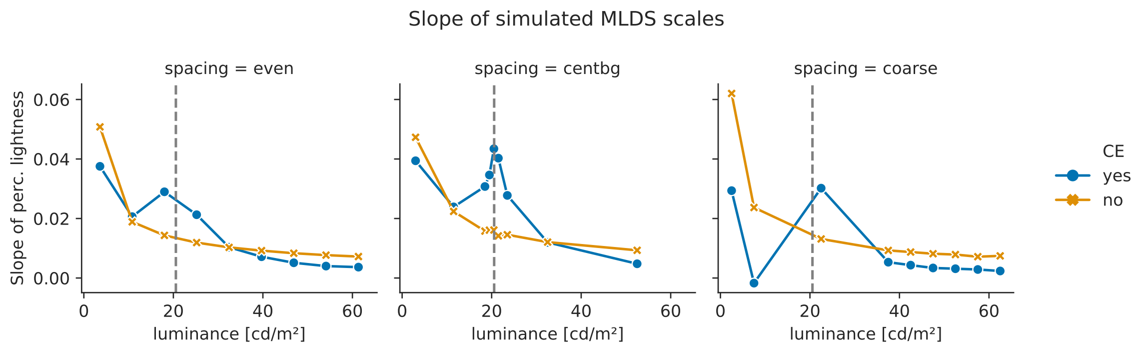

Figure 3.6: Slope of MLDS scales obtained in a simulation with different spacings. The slope of perceived lightness (y-axis) is plotted against luminance (x-axis). The orange crosses and blue circles show the slope of the scales based on ground truth without and with CE, respectively. In order to facilitate the comparison of the two ground truth functions the points are connected with straight lines.

together in one plot. The orange crosses represent the slope of the MLDS scales based on the ground truth function without CE. The blue circles show the slope of the MLDS scales based on the ground truth function with CE. In order to facilitate the comparison the actual slope values indicated with markers are connected with straight lines. The dashed gray line indicates the hypothetical background luminance.

The left panel of Figure 3.5 shows the slope for the evenly distributed luminance values. The middle and right panel present the slopes for the spacing with luminance values centered around the background luminance and the coarse spacing, respectively. Indeed, the slope based on the ground truth function without CE is increased around the background luminance. This is most obvious for the centered spacing in the middle panel of Figure 3.6. Nevertheless, the tendency is also visible in the even spacing. For the coarse spacing the slope is also increased. However, before increasing the slope drops significantly (2nd point). When looking at the data in Figure 3.5, I observe that the standard deviation is highest at this point. Possibly, the scale could not be estimated accurately in this area due to the specific spacing. Overall, the slope of the scales with CE is increased for all spacing and therefore the slope should be considered as an indicator for CE.

Figure 4 in the appendix shows the slope for the remaining spacings.

3.3 Summary of Simulations

In this chapter, I simulated an asymmetric matching experiment and an MLDS experiment based on two different ground truth functions with and without CE. In the asymmetric matching simulation, I verified that the simulation method is valid since it shows the expected results, and it can recover CE based on two different ground truth functions. In the next step, I simulated an MLDS experiment to investigate the effect of different spacings in luminance range on the resulting perceptual scales. I chose three different spacings for the simulation

– one with evenly distributed luminance values, one with luminance values centered around the background luminance, and one coarse spacing with luminance values “far” away from the background luminance in luminance space. I introduced two metrics in order to enable the evaluation of the similarity of the two scales to the ground truth functions and the evaluation of the discriminabilty of the two scales – with and without CE. Some spacings impede the exact recovery of the underlying ground truth functions; the highest deviations occurred for the centered spacing without CE and for the coarse spacing with CE. The ground truth function with CE was recovered well with the even and centered spacing. None of the spacings prevents the two perceptual scales – based on the two different ground truth functions – from being visually distinguished. Certainly, according to the metric the scales of the centered spacing are more similar to each other than the scales of other spacings. In future work, the metrics should be further developed to also consider the standard deviation. Overall, the results indicate that an even spacing or a spacing with several luminance values around the background should be used in an experiment investigating CE. Apparently, one value very close to the background luminance sufficed to recover CE correctly – as in the even spacing. Furthermore, the slope might be beneficial as an indicator of CE.

Chapter 4

Reanalysis of Previous Work

As all experiments observing CE are based on adjustment tasks, it is not evident whether the effect is a feature of the visual system or a consequence of the adjustment procedure. In order to find an answer to my research question, I reanalyze data obtained in an experiment based on a decision task. I use data collected in asymmetric matching, MLDS, and MLCM experiments by Aguilar and Maertens (2020). In Section 4.1, I describe how the authors obtained the data.

Afterward, I reanalyze the collected data. All programs used in this reanalysis and the corresponding data and plots are uploaded in GitHub[1]. Initially, I show the data from the asymmetric matching experiment where CE occurs in keeping with the expectation. Next, I analyze whether CE occurs in both MLDS and MLCM data as well. As MLDS and MLCM are based on decision tasks and difference judgments, the CE's occurrence in such data would imply that it is a built-in feature of the visual system and not a consequence of adjustment procedures.

As mentioned above, CE shows as a steepening of the slope of the perceptual scale around the background luminance. This steepening is only expected in the scales for center-surround (CS) stimuli with a homogeneous background (c.f. Ekroll et al., 2004). In the reanalysis, I observe a difference between the perceptual scales obtained with homogeneous and variegated backgrounds, as expected. The scales differ due to the steepening, similar to the difference recorded in the simulation in Section 3.2.

Additionally, I apply the distance metric developed in Section 3.2.4 (c.f. Equation 3.4), to the observers' perceptual scales obtained with MLDS and MLCM. Apart from that, I analyze the slope of the data to test it as a metric for the occurrence of CE. I examine whether the slope's magnitude is indeed increasing around the background luminance in order to have a better metric than only visual comparison. I also investigate how the pattern of the slope differs between the data for experiments with homogeneous as well as variegated backgrounds. Overall, I examine whether the distance metric and the slope could be used as indicators for the occurrence of CE.

[1] https://github.com/computational-psychology/brunn_ba2020

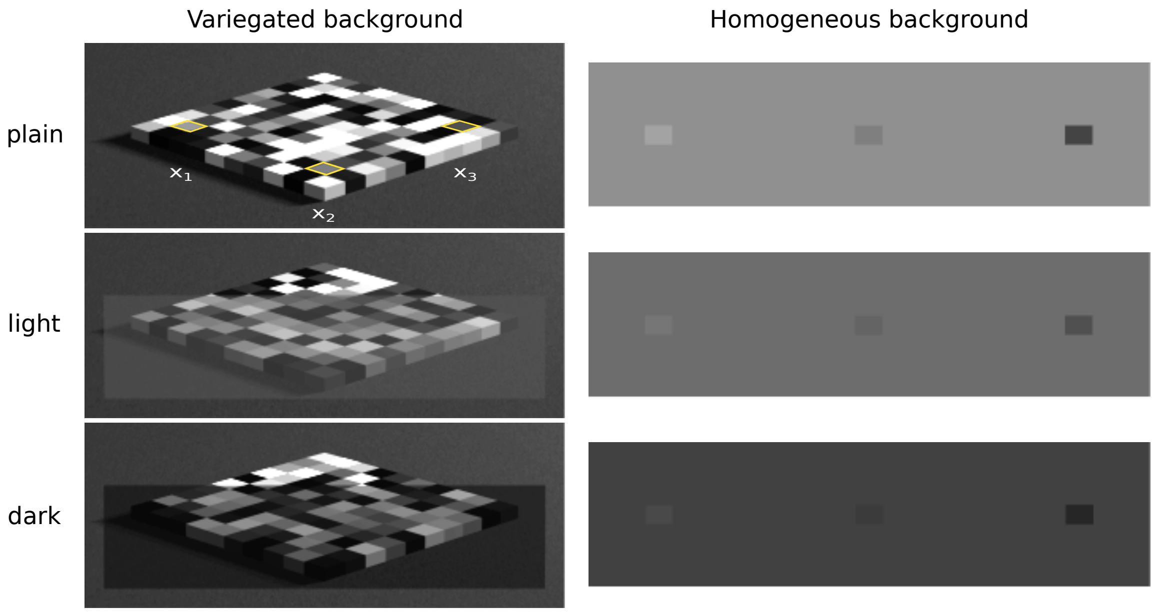

Figure 4.1: Stimuli used in the MLDS experiment. The left column shows different variations of the variegated checkerboard created by interposing transparencies. The right column shows variations of CS-stimuli. Here, instead of interposing transparencies, the background luminance is changed to create different viewing conditions. The background luminance equals the average luminance of the checks (behind transparencies) of the variegated checkerboard for the corresponding viewing condition. The stimuli are an approximation of how the stimuli of different background conditions would look in a calibrated monitor.

4.1 Data Collection

In the following, I briefly explain the experimental methods and procedure in Aguilar and Maertens (2020) before reanalyzing their data. Aguilar and Maertens (2020) examined whether the perceptual scales measured in different contexts can be meaningfully compared with each other. For measuring the perceptual scales, they used two statistical scaling methods – MLDS and MLCM. They used different viewing conditions – plain, dark, and light – for the experiment. These viewing conditions were created by either varying the background luminance or by interposing a transparency between the observer and the stimulus.

For creating the perceptual scales, the statistical scaling methods MLDS and MLCM were used. Both methods are explained briefly in Section 1.3.2. MLDS establishes a reasonable perceptual scale for one viewing condition (Aguilar & Maertens, 2020). All the elements of one stimulus (triad) are always presented in the same viewing condition. In addition, the minimum value of each scale is always anchored to zero by default in MLDS. MLCM is an enhancement of MLDS. In contrast to MLDS, MLCM sets the scale of one context as a reference and the other scales are determined relative to that reference. Only the minimum of the reference scale is anchored to zero. By doing so, an erroneous shift can be avoided.

4.1.1 Stimuli

In the following, I briefly describe the stimuli used in the experiment. Aguilar and Maertens (2020) used two kinds of stimuli: CS stimuli with a homogeneous background and variegated checkerboards. In the reanalysis, I expect CE to occur only for the data measured with a homogeneous background, as Ekroll and Faul (2009) found that adding variegation reduces or abolishes CE (c.f. Chapter 2).

Variegated Background

Aguilar and Maertens (2020) used variegated checkerboards as stimuli that induce lightness-constancy. The left columns of Figure 4.1 and Figure 4.2 show variations of the checkerboards. In MLDS and MLCM, three or two checks, respectively, are chosen as targets which the observer has to compare. The exact procedure is described in Section 4.1.2. In Figure 4.1, the targets are marked with a yellow outline. The three targets in MLDS are aligned in either descending or ascending order in such a way that $x_1 < x_2 < x_3$ or $x_1 > x_2 > x_3$, where x_i with $i \in \{1, 2, 3\}$ is the target's luminance. In MLCM, the two targets are at the position of x_1 and x_3 in MLDS. In order to avoid any differences in the direct surroundings of the targets, the same eight reflectance values are used for the adjacent checks, albeit in a shuffled order.

In order to create the different viewing conditions, a transparency is interposed between the observer and the checkerboard. These transparencies are created by using α blending. α specifies transmittance of the transparency. In α blending, the resulting luminance of a region, l' is created by combining the transparency's luminance, l_T, and the luminance of the region without transparency, l. Hence, the luminance of the region covered by the transparency is created with:

$$l' = \alpha \cdot l + (1 - \alpha) \cdot l_T \tag{4.1}$$

In MLDS, all the three elements of a triad are presented in the same viewing condition – i.e., with or without a transparency – in order to avoid any cross-context comparison. The second row in 4.1 shows an example of the light viewing condition, whereas the bottom one shows the dark viewing condition.

In MLCM, it is allowed to present one target in another environment – e.g., background, transparency – than the other target, as illustrated in Figure 4.2. This is called cross-context comparison.

Homogeneous Background

Additionally, Aguilar and Maertens (2020) used CS stimuli with a homogeneous background. The right column of Figure 4.1 displays exemplary MLDS stimuli with three targets that are presented in front of a homogeneous background. Instead of interposing any transparency, the different viewing conditions are created by using different types of background luminance.

The background's luminance equals the checks' mean luminance that are seen behind the transparency in the variegated checkerboard. The upper row corresponds to the viewing con-

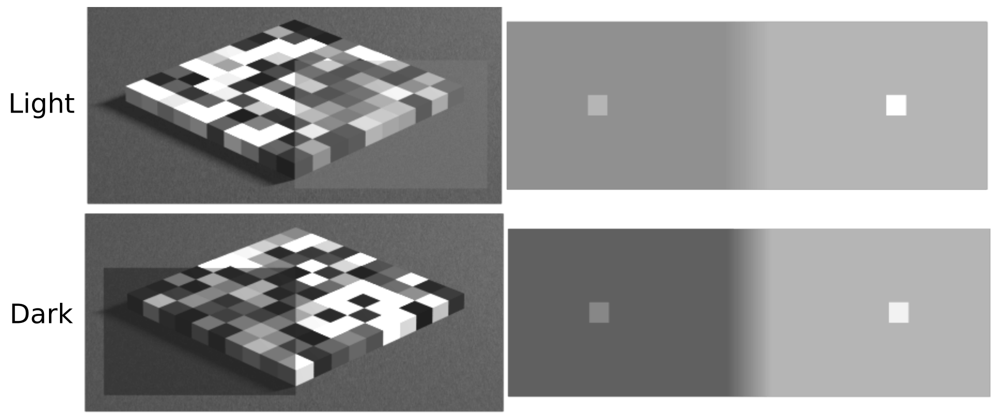

Figure 4.2: Stimuli used in the MLCM experiment. The left column shows different variations of the variegated checkerboard created by interposing transparencies. The upper row shows the example of a stimulus with a light transparency (background), while the lower row shows the example of a dark transparency (background). In MLCM, a cross-context comparison is allowed and therefore one target can be presented in another environment – background or transparency – than the other target. As the MLDS stimuli, these are approximations of stimuli presented on a calibrated monitor.

dition without any transparency, whereas the middle and lower rows show stimuli corresponding to the viewing condition with both light and dark transparency, respectively. In MLCM, the two targets can be presented in front of backgrounds with different levels of lightness. The right column of Figure 4.2 provides an example of these stimuli. The upper row shows stimuli corresponding to a light transparency, while the lower row shows stimuli corresponding to a dark transparency.

Matching Stimuli

For a full comparison, Aguilar and Maertens (2020) also included a matching task in their experiment. The stimulus looked similar to the one displayed in Figure 1.4. The external matching field (match) was presented above the target field (target) and embedded in a 5x5 variegated checkerboard. The mean luminance of these checks equaled the mean luminance of the lower variegated checkerboard seen without any transparency. In order to keep the same adjacent luminance between trials, the surrounding checkerboard of the match is solely rotated by 90° for creating variance. The target was either embedded in a variegated checkerboard at the position of x_2 (MLDS) (c.f. Figure 4.1) or presented in front of a homogeneous background.

4.1.2 Design and Procedure

Eight observers participated in the experiment. All the observers performed the experiments in the same order: MLDS, MLCM, asymmetric matching. Each of the unique trials was executed 10 times in order to mitigate the effect of perceptual noise and careless mistakes. The total number of trials varied across experiments.

MLDS Experiment

The method of triads was used in the MLDS experiment. Here, three squares are either embedded in the variegated checkerboard or presented on top of a homogeneous background. The yellow-framed squares in Figure 4.1 are the elements of the triad. The observers have to judge whether the difference between the left and the middle square is bigger than the difference between the right and the middle square. The targets could have 10 different luminance values, which resulted in a total of 120 different triads. Three different viewing conditions – plain, light, and dark – were used, and each set of triads was repeated 10 times. This resulted in a total number of 3600 trials: These 3600 trials were split into 10 blocks. The observers could do as many blocks as they wanted at once.

MLCM Experiment

In MLCM, the observers compared two targets during the experiment. They must indicate which target is perceived lighter. MLCM uses cross-context comparisons in order to create perceptual scales, which are all relative to one reference scale. In this experiment, the viewing condition without any transparency was chosen as a reference. Therefore, at least one of the targets was always presented in "plain" view (without transparency) in the cross-context comparisons and the number of necessary trials was reduced. As in MLDS, 10 different reflectance values were used for the targets. For the second dimension - the background - three variations were used, which results in 30 different stimuli values. This resulted in 435 possible pairs of stimuli but was then reduced to 335 possible pairs, as the cross-context comparison between the dark and light viewing conditions was excluded. The unique set of 335 trials was repeated 10 times, which resulted in a total number of 3350 trials. As in the MLDS experiment, these 3350 trials were split into 10 blocks. The observers could do as many blocks as they wanted at once.

Matching Experiment

In the matching experiment, the observer should adjust the luminance of the external matching field in such a way that the matching field and the embedded target in the checkerboard were perceived equally. As in MLDS and MLCM, 10 different reflectance values were used for the target in the matching experiment. The experiment was repeated for the three different viewing conditions, and each judgment was repeated 10 times as for the other experiments. This resulted in a total number of 300 trials.

4.2 Reanalysis of the Data

In the following, I reanalyze the data of Aguilar and Maertens (2020) to see if these data showed CE. In the original experiment, Aguilar and Maertens (2020) studied whether the perceptual scales measured in different contexts could be meaningfully compared with each other. They found that MLCM seems “better suited to measure perceptual scales across different viewing conditions” (Aguilar & Maertens, 2020). Additionally, CE was visible in their matching data,. The effect also seems to occur in the MLDS data obtained with a homogeneous background, but the data are not evident. Therefore, I reanalyze these data to see whether CE occurs in the perceptual scales obtained with both MLDS and MLCM.

Throughout the reanalysis, the observers are labeled with the same abbreviations - O1, $\cdots$, O8 - as in Aguilar and Maertens, 2020. The observers O3/GA and O2/MM are the authors of Aguilar and Maertens, 2020. First, I plot the data for the asymmetric matching data. Here CE occurs as it occurred in case of Takasaki (1966) or Ekroll et al. (2004), Ekroll and Faul (2009). Afterward, I plot the data for the MLDS and MLCM experiments. For MLDS, CE occurs for some observers. In all following plot grids, the observers are ordered by the visual inspection of the magnitude of CE in the MLDS scales obtained without any transparency (as in Figure 4.7). The magnitude of the effect decreases throughout observers.

I also analyze the distance between the scales obtained with MLDS and MLCM, calculated as in Section 3.2.4 with Equation 3.4.

4.2.1 Asymmetric Matching Data

In order to visualize the asymmetric matching data, the match’s luminance value on the y-axis is plotted against the target’s luminance on the x-axis. Figure 4.3 shows the matching data obtained in the experimental condition without any interposed transparency.

The orange crosses indicate the data for the experimental condition with a homogeneous background, while the blue circles indicate the data for the experimental condition with a variegated background. In order to facilitate the comparison of the two experimental conditions the data points are connected with straight lines. However, these lines do not show actual measurements. The dashed gray line indicates the target’s background luminance.

CE occurs in the data of four observers – O6, O4/MK, O2/MM, and O5. Here the data shows a strong non-linearity for the experimental condition with a homogeneous background. This was expected as CE also occurred in the asymmetric matching data of Takasaki (1966) and Ekroll and Faul (2009).

The data obtained in the experiment with an interposed dark and light transparency can be found in Section .2.1 in the appendix, but they are not analyzed here as the focus is on the perceptual scales to answer the research question.

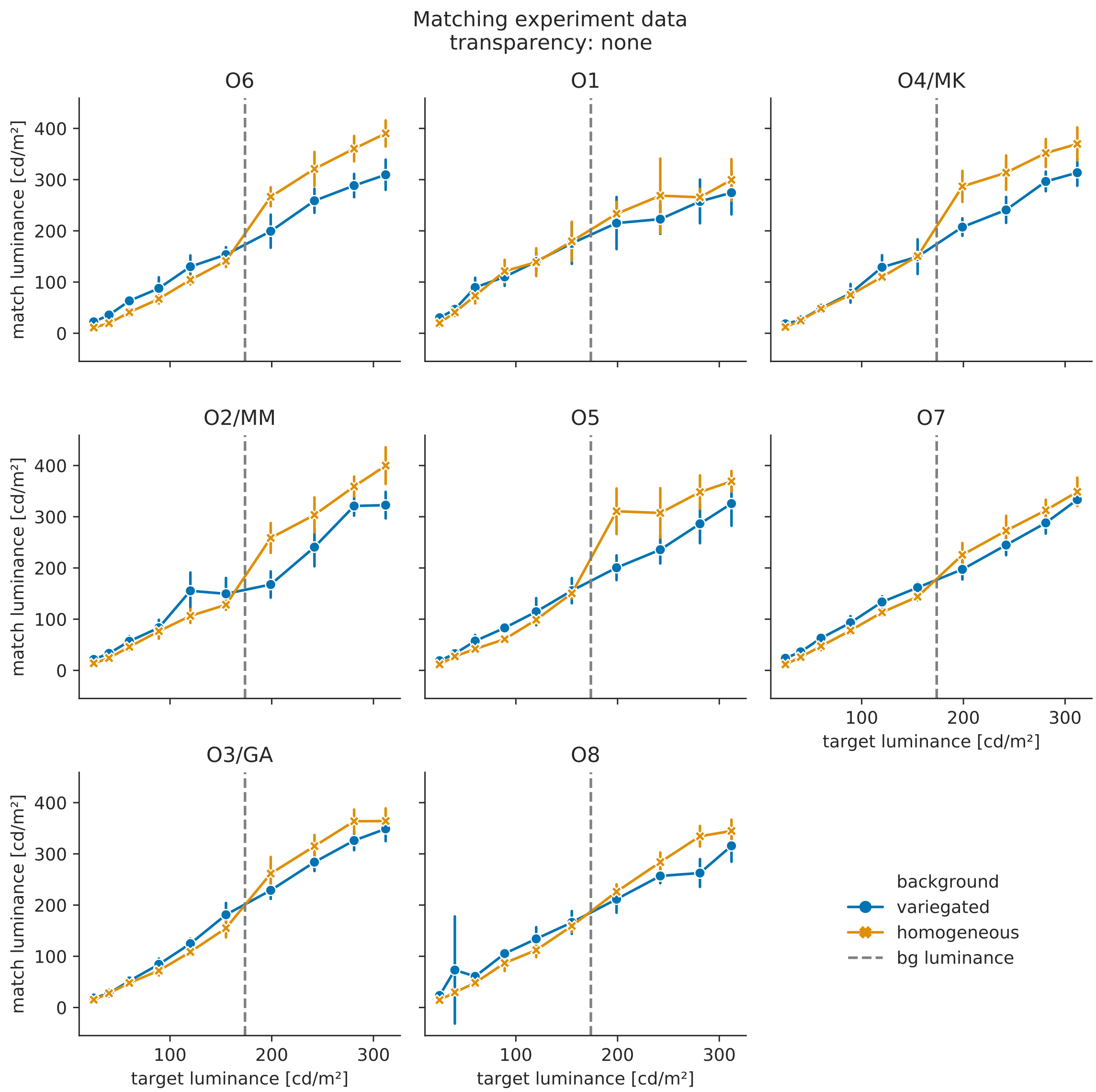

Figure 4.3: Matching data for all observers without any interposed transparency. The observer labels (O1, $\cdots$, O8) are the same as in Aguilar and Maertens (2020). The orange crosses show the data for the experiment with a homogeneous background, while the blue circles the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions, the data points are connected with straight lines. The dashed gray line indicates the target's background luminance. For O6, O4/MK, O2/MM, and O5, there is a strong non-linearity in the data for the homogeneous experimental condition.

background	O1	O2/MM	O3/GA	O4/MK	O5	O6	O7	O8
plain	0.8	0.76	0.31	0.6	0.85	1.17	0.65	0.25
dark	1.79	1.24	0.33	0.99	1.28	1.71	0.74	0.96
light	2.61	1.15	0.5	1.03	1.76	1.78	1.23	1.38

Table 4.1: Distance metric (c.f. Section 3.2.4) applied to MLDS data: It is clearly visible that the distance is larger for the observers whose data shows CE (e.g. O6). The background labels plain, dark, and light indicate the experimental condition without, and with a dark and light transparency, respectively. The values are approximated values.

4.2.2 Scales obtained with MLDS

In this section, I show the perceptual scales obtained with MLDS in order to clarify whether CE also occurs in the perceptual scales, measured with MLDS, for the homogeneous background. As expected, a difference between the perceptual scales for the variegated and the homogeneous backgrounds is visible. Presumably, this difference implies the occurrence of CE. In the experiment without any transparency, some perceptual scales for the two background conditions cross. This is a consequence of the slope's increase of the scale for the homogeneous background.

In order to compare the scales for both conditions, I normalize the maximum value of all perceptual scales to one. Additionally, I connect the perceptual scale values with straight lines in order to simplify the comparison of the scales for the different background patterns. However, these lines do not represent actual measurements. For the viewing conditions produced with interposed transparencies or a dark or light background, the luminance range, for which the perceptual scale was measured, is reduced by design. The vertical lines at the markers show the confidence intervals obtained with the bootstrapping routine implemented in the MLDS package in R[2]. The dashed gray line indicates the average background luminance.

In the following, I present the data of two exemplary observers, O6 & O8, to demonstrate the difference between the perceptual scales for the observers with and without CE. Next, I show the data for all the observers who participated in the experiment. As already mentioned above, a difference between the scales for the two different background patterns is visible. However, not all observers seem to experience CE. This is not surprising as Ekroll and Faul (2009) already stated that the data of some observers did not show the characteristics of CE.

Table 4.1 shows the results of the distance metric developed in Section 3.2.4 applied to the MLDS scales of each observer. The values are rounded to two decimals. The accurate values can be found in the GitHub[3].

Two Observers as an Example

In order to elucidate the difference between the perceptual scales caused by CE, Figure 4.4 shows the perceptual scales of O6 and O8. These perceptual scales were measured in the experiment without any transparency. The orange scale shows the data measured with a ho-

[2]https://www.rdocumentation.org/packages/MLDS, https://cran.r-project.org/web/packages/MLDS/MLDS.pdf
[3]https://github.com/computational-psychology/brunn_ba2020/tree/main/data/reanalysis/csv_data

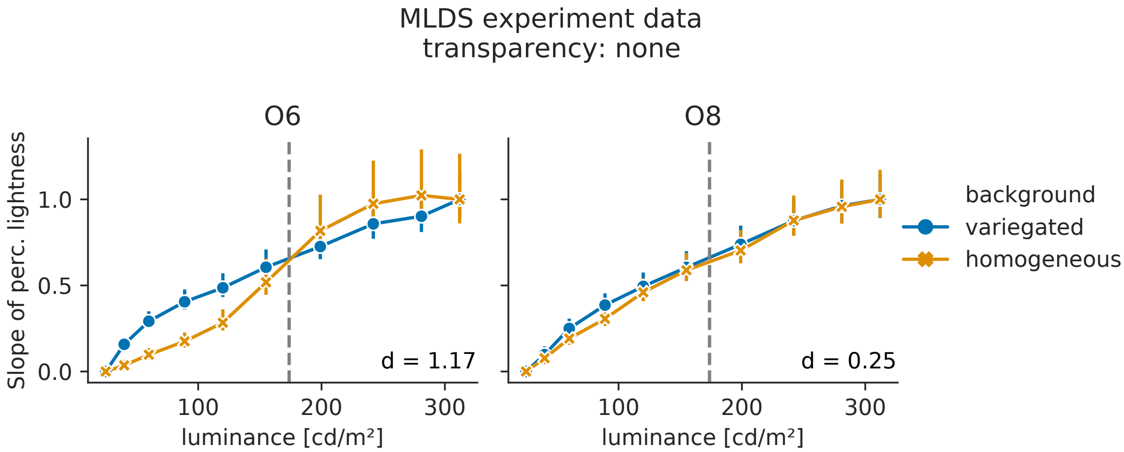

Figure 4.4: Perceptual scales measured with MLDS for two exemplary observers, O6 & O8. The perceptual scales of the two observers differ from each other due to an increase in the slope around the background luminance, which leads to a crossing of the scales for the variegated and homogeneous backgrounds for O6. In the perceptual scales of O8, no steepening is observable.

mogeneous background, while the blue scale represents the data obtained with a variegated background. The dashed gray line indicates the background luminance. The d in the bottom right corner indicates the distance between the two scales calculated as in Section 3.2.4. For O6, the perceptual scales for the different background patterns (homogeneous & variegated) intersect at the background luminance. The slope of the orange scale increases significantly around the background luminance. In contrast, the scales of O8 lie almost on top of each other and the slope does not increase close to the background luminance. Presumably, the increase in the slope that is visible for O6 is a consequence of CE. Hence, O6 is an observer with CE, while O8 is an observer without CE.

The distance between the perceptual scales also varies significantly between the two observers. The distance for O6 is approximately 1.17, whereas the distance for O8 is just 0.25. Consequently, the scales' distance of the observer who shows CE is significantly higher than the distance of the other observer.

Dark Transparency

Figure 4.5 shows the perceptual scales for the same two observers measured with an interposed dark transparency or a dark homogeneous background. For O6, there is still a visible difference between the two scales; however, the scales do not cross. For O8, the scales do not lie on top of each other anymore, but they are still parallel to each other. Hence, the slope of the two scales of O8 is approximately the same. As without any transparency, the distance between O8's scales (0.96) is lower than the distance for O6 (1.71).

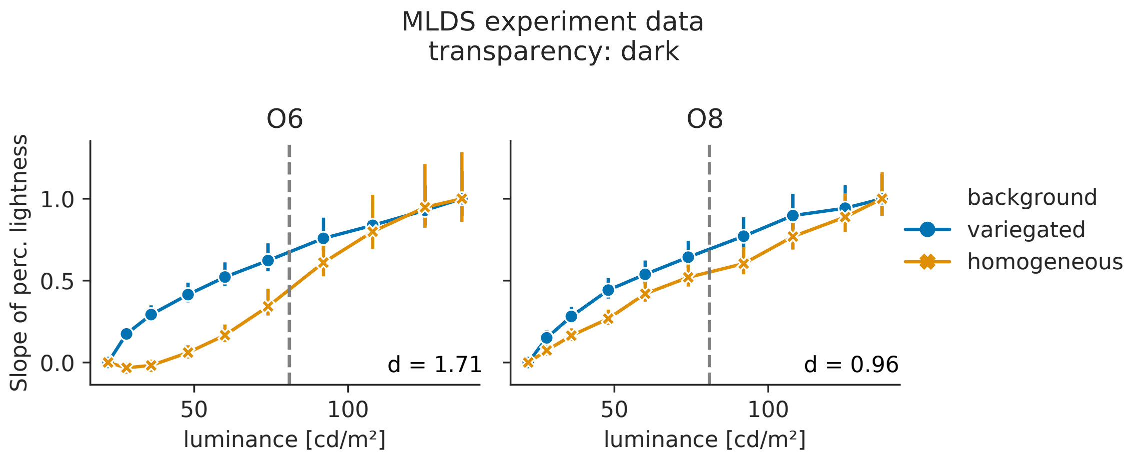

Figure 4.5: Perceptual scales for the two exemplary observers, O6 & O8, measured with MLDS in an experiment with interposed dark transparency or a dark homogeneous background. The perceptual scales of the two observers differ due to a differing slope. As opposed to the experiment without any transparency, the perceptual scales of O6 do not cross but still have different slopes. The scales for O8 are mostly parallel to each other; the slope is nearly the same.

Light Transparency

Figure 4.6 shows the perceptual scales measured for the same two observers with a light transparency or a light background. The scales look similar to the ones measured with a dark transparency. Only the luminance range of the two conditions is different. For O6, the distance metric (1.78) is similar to the dark transparency. For O8, the distance for the light condition (1.38) is higher than that for the dark condition. However, the distance of O6's scales is still higher than that of O8.

Undoubtedly, the slope of O6's perceptual scales for the different background patterns is different. However, a unique feature is missing to characterize CE precisely, such as the crossing of the perceptual scales in Figure 4.4.

The distance metric, developed in Section 3.2.4, might be a useful indicator. The scales' distance of O6 - the observer with CE - is always higher than that of O8, the observer without CE. In the experimental condition without any transparency, the two scales differ the most for O6 and are the most similar for O8. In this condition, the difference between the two observers is the biggest.

Additionally, the slope analyzed in Section 4.3 might be beneficial for finding a unique characteristic to detect CE.

In the following, I show the data of all observers to demonstrate how many scales show patterns similar to O6.

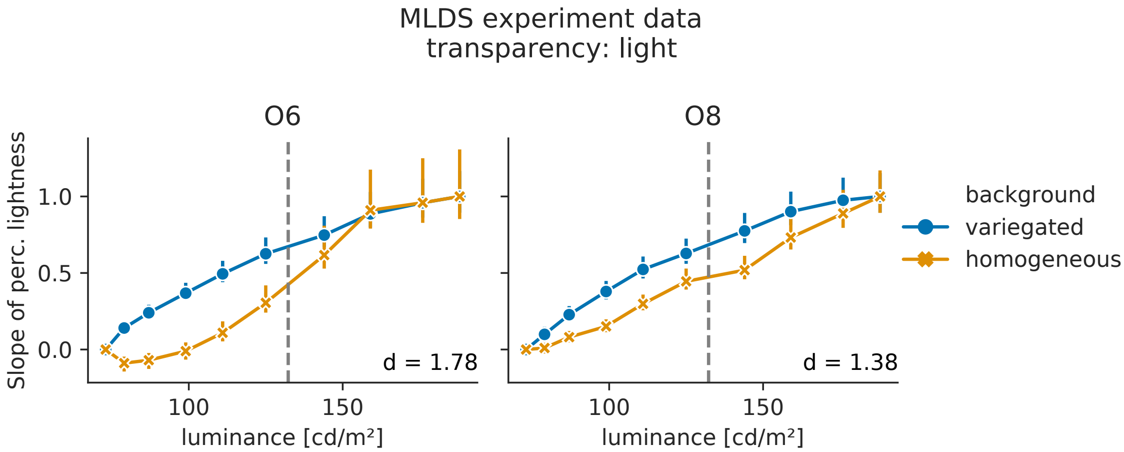

Figure 4.6: Perceptual scales for the two exemplary observers, O6 & O8, measured with MLDS in an experiment with interposed light transparency or a light homogeneous background. The perceptual scales of the two observers differ. The slope of the O6's scales differs between them. In the perceptual scales of O8, no steepening is observable.

Comparison of All Observers

Figure 4.7 shows the perceptual scales of all the observers measured with MLDS in the experiment without any interposed transparency. The orange crosses indicate the data for the context with a homogeneous background, while the blue circles display the data for the context with a variegated background. Again, I connected the perceptual scale values with straight lines in order to compare the two experimental conditions. Indeed, the slope of the perceptual scales differs for most observers. Only the perceptual scales of O3/GA and O8, presented in the lower row, lie almost on top of each other, despite the different background patterns. For O6, O1, O4/MK, and O2/MM, the perceptual scales cross as in Figure 4.4 for O6. For O5 and O7, the perceptual scales do not cross at the background luminance. In fact, for O7, the scales do cross, even though at a darker luminance value than the background luminance. Hence, the standard deviation might lead to this crossing and not CE.

Figure 4.8 and Figure 4.9 show the perceptual scales for all the observers obtained in the experiment with a dark or light transparency (background), respectively. In Figure 4.8, the perceptual scales differ, but only the scales of O4/MK cross at the background luminance. For the other observers, the characteristic steepening around the background luminance is not visible. The scales of O3/GA and O8 lie almost on top of each other.

The data obtained with the light transparency look similar to those obtained with a dark transparency. For O4/MK and O2/MM, the steepening of the slope at the background luminance is visible. The scales of O4/MK cross at the background luminance. The scales' slopes of O6, O1, O5, and O7 differ for the different background patterns. Nevertheless, the scales do not cross. For O3/GA, the scales still lie almost on top of each other. However, for O8, the slope decreases around the background luminance.

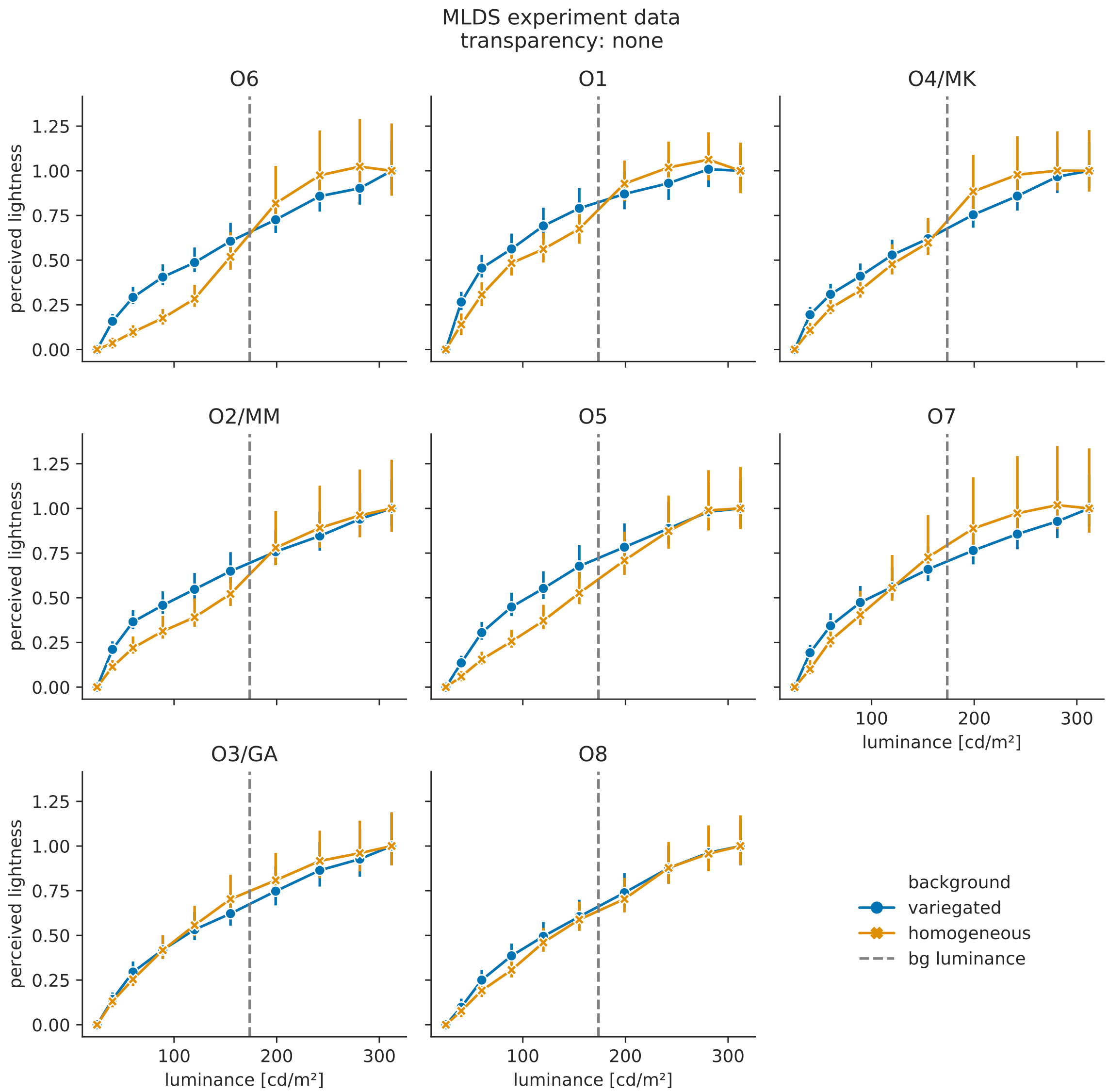

Figure 4.7: Perceptual scales measured with MLDS for all observers without any interposed transparency. The perceptual scales are plotted against luminance. The orange crosses show the perceptual scale values for the experiment with a homogeneous background, while the blue circles display the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the perceptual scale values are connected with straight lines. For O6, O1, O4/MK, and O2/MM, a non-linearity is observable in the data for the homogeneous background (orange).

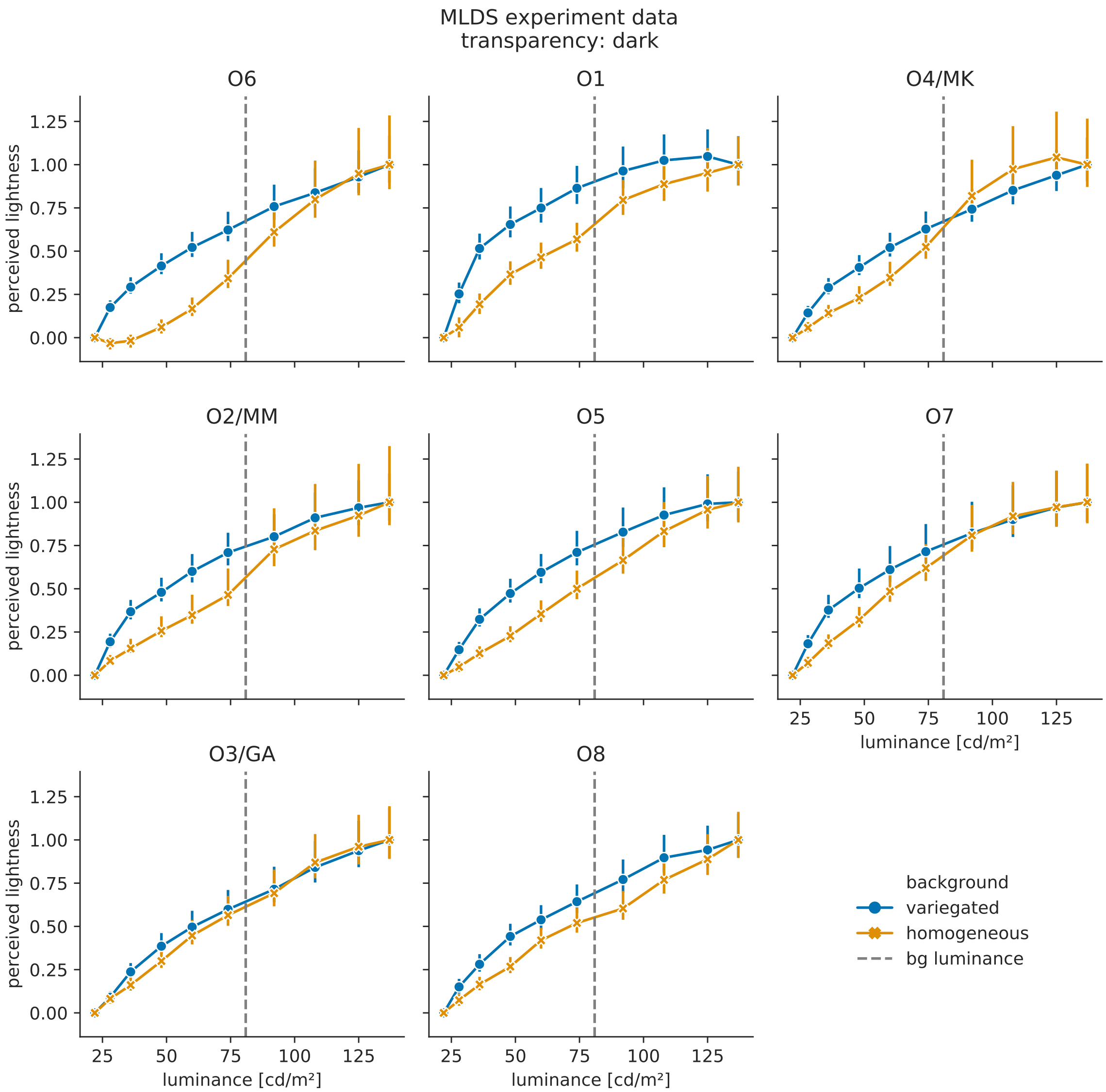

Figure 4.8: Perceptual scales measured with MLDS for all observers with a dark transparency. The plot design and the legends are the same as in Figure 4.7. For most observers, the perceptual scales obtained with different background patterns differ from each other. Only for O3/GA, the scales almost overlap. For O8, the perceptual scales are almost parallel; hence, the slope is mostly the same for both scales.

Evaluation of the Distance Metric

As already stated, Table 4.1 shows the values calculated with the distance metric for all the observers. For the viewing condition without any transparency ("plain"), the distance is the

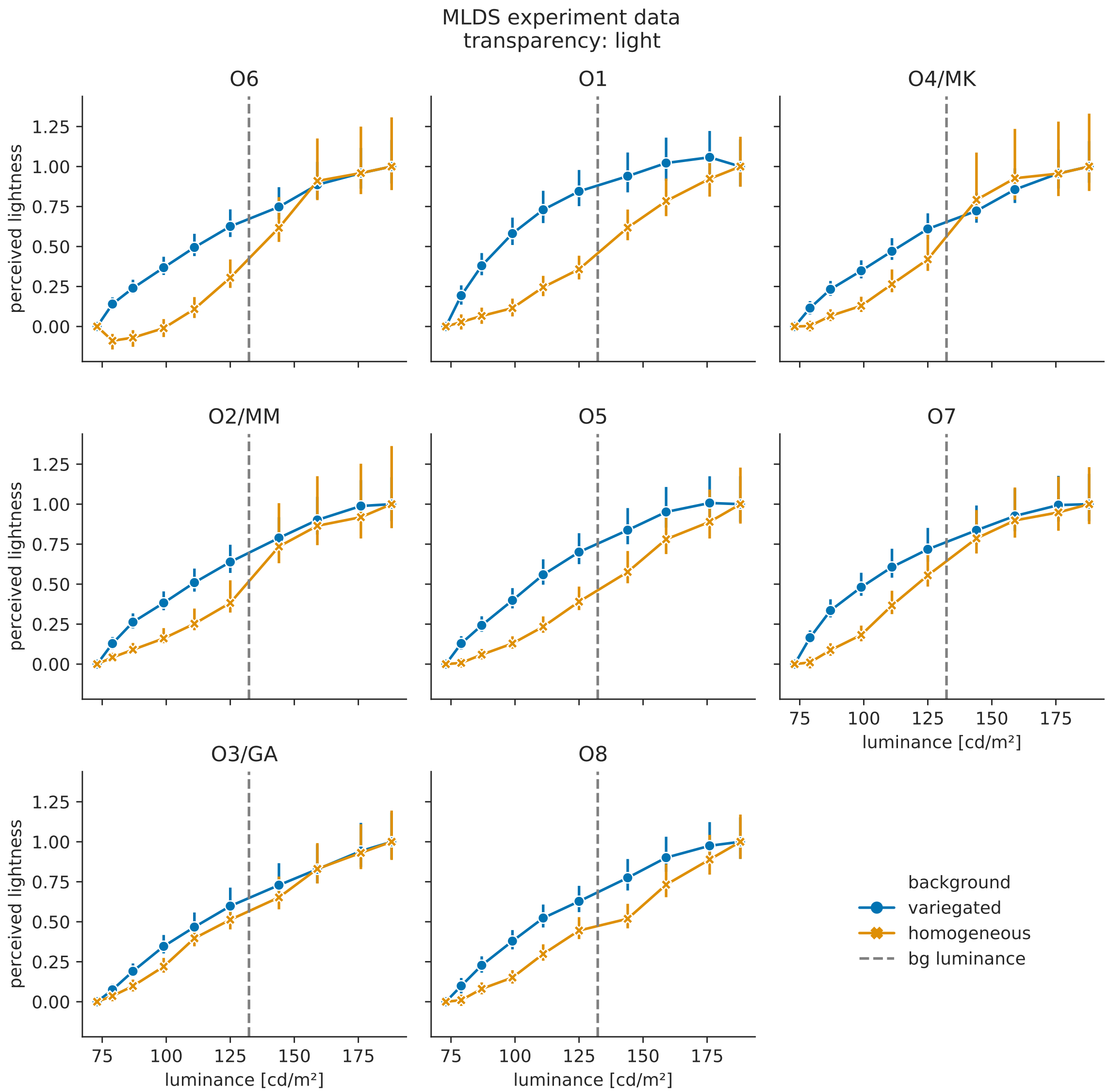

Figure 4.9: Perceptual scales measured with MLDS for all observers with a light transparency. The plot design and the legends are the same as in Figure 4.7. The scales look similar to the ones obtained with a dark transparency in Figure 4.8. Overall, the scales differ more than in Figure 4.8.

highest for O6, followed by O5, O1, O2/MM, O7, O4/MK, O3/GA, and O8 in descending order. It is noticeable that this order does not correspond to the ordering of the observers in Figure 4.7. By visual inspection, the magnitude of CE seems larger for O1 and O4/MK than for O5 and O7. However, according to the calculated distance, the value of O5 is higher than that of

background	O1	O2/MM	O3/GA	O4/MK	O5	O6	O7	O8
plain	0.9	0.33	0.63	0.52	0.91	0.62	0.39	0.71
dark	1.22	0.78	0.73	0.75	0.57	0.96	0.58	1.08
light	1.14	1.55	2.05	1.49	1.35	0.71	1.09	0.89

Table 4.2: Distance metric (c.f. Section 3.2.4) applied to the MLCM data: The values do not show a clear tendency as for MLDS. It is not possible to tell whose data shows CE based on the distance. The background labels are the same as in Table 4.1. The values are approximated values.

O1. Moreover, the value of O7 higher than that of O4/MK. However, the distance values of O5 and O1, and O7 and O4/MK differ only slightly.

For the viewing condition with a dark transparency, the observers are ordered according to the calculated distance: O1 > O6 > O5 > O2 > O4 > O7 > O8 > O3. For the viewing condition with a light transparency, the observers are ordered as in the dark viewing condition, except that O7 and O8 are swapped. Overall, the calculated distance fits well with the visible magnitude of the effects in Figure 4.8 and Figure 4.9.

Four out of the eight observers (O6, O1, O4/MK, O2/MM) clearly show CE in the MLDS scales obtained without any transparency. The distance metric for these four observers is also large. For two observers, O5 and O7, CE is not so obvious by visual inspection, but the distance metric indicates the occurrence of CE. In the MLDS scales with a dark and light transparency, the same four observers as without any transparency show CE. For the observers O5 and O7, CE is not so clear by visual inspection, since the "characteristic" steepening of the scales is missing. According to the distance metric, O5's distance is higher than that of O6, O2, and O4. Hence, when assuming that a high distance indicates CE, O5 would have CE. It is not possible to make a clear statement about O7 and O8 based on the visual inspection and the distance metric. One observer, O3/GA, does not seem to have CE at all.

In order to develop a better metric than visual inspection for the occurrence of CE I analyze the slope of the perceptual scales in Section 4.3.2.

4.2.3 Scales obtained with MLCM

In this section, I present the reanalyzed data that was obtained with the second statistical scaling method – MLCM. The arrangement and design of these plots are the same as for MLDS (c.f. Section 4.2.2). Figure 4.10 shows the data measured without any interposed transparency.

Table 4.2 shows the values calculated with the distance metric for the perceptual scales obtained with MLCM.

Interestingly, in Figure 4.10 the steepening of the curve, which is clearly visible in the MLDS scales, is not visible here. From my point of view, MLCM and asymmetric matching are more similar than MLDS and matching when considering the task that the observers have to perform. The observers have to compare the targets pairwise and must decide which target appears lighter. Hence, the appearances of the two targets have to be compared. One may argue that this process could be internally based on adjustment procedures. Consequently, I would have expected to see the effect in these data.

In Table 4.2 there is no clear tendency as it was the case in MLDS. For the condition without any transparency (plain), the calculated distance for O8, the exemplary observer without CE (c.f. Section 4.2.2), is even larger than the distance for O6, the exemplary observer with CE. In fact, the values would be expected to be ordered reversely.

One possible reason that CE is not observed could be the high noise level of the data. The confidence intervals in Figure 4.10 are larger than those in the MLDS scales. Another reason might be the nature of the scaling process, which is similar to Thurstonian scaling (L. Maloney & Yang, 2003). The method only records which of the two targets appears lighter. Hence, only the relation between the two targets is judged, but the actual magnitude of difference in perception is neglected. This could affect the occurrence of CE. In order to recover the steepening of the slope, one has to detect that the difference between two values is larger than the difference between two other values. Hence, disregarding the differences' magnitudes could affect the occurrence of CE, as any steepening of the slope might be missed. MLDS might be better suited to recover CE because the observers judge the magnitude of the targets' differences instead of only judging their appearances.

The perceptual scales for MLCM, measured with a dark or light transparency, can be found in Section .2.2 in the appendix.

4.3 Analysis of the Slope

In the following I analyze the scales' slope numerically to have a better metric than only visual inspection. Based on the results of the slope of the simulated scales in Section 3.2 I expect to see an increase of the slope at the background luminance (c.f. Figure 3.6). I investigate whether the slope can be used as an additional metric for the occurrence of CE. Since the slope's increase is visible in the MLDS scales of some observers (c.f. Figure 4.7, 4.8, 4.9), the slope might serve as an indicator for CE. Additionally, I investigate whether the slope has a particular pattern for the data of experiments with a homogeneous background and a variegated background.

I calculate the slope in the same way as described in Section 3.2.5. Thus, 10 perceptual scale values for the corresponding luminance values lead to nine slope values. I use Equation 3.6 for the calculation and x corresponds to the luminance values used by Aguilar and Maertens (2020). The perceptual scale values are inserted in y. I use the perceptual scale values, marked in the plots, and do not interpolate between them. I also neglect the confidence intervals in calculating the slope.

In the plots, the observers are aligned in the same order as in the reanalysis of the data, where the observable magnitude of CE in the MLDS scales without any transparency was decreasing. Hence, I also expect that the increase in the slope around the background luminance decreases gradually across the observers. The orange crosses show the slope of the data for the experiment with a homogeneous background, while the blue circles display the slope of the data for the experiment with a variegated background. Again, I connect the data points with straight lines in order to simplify the visual comparison.

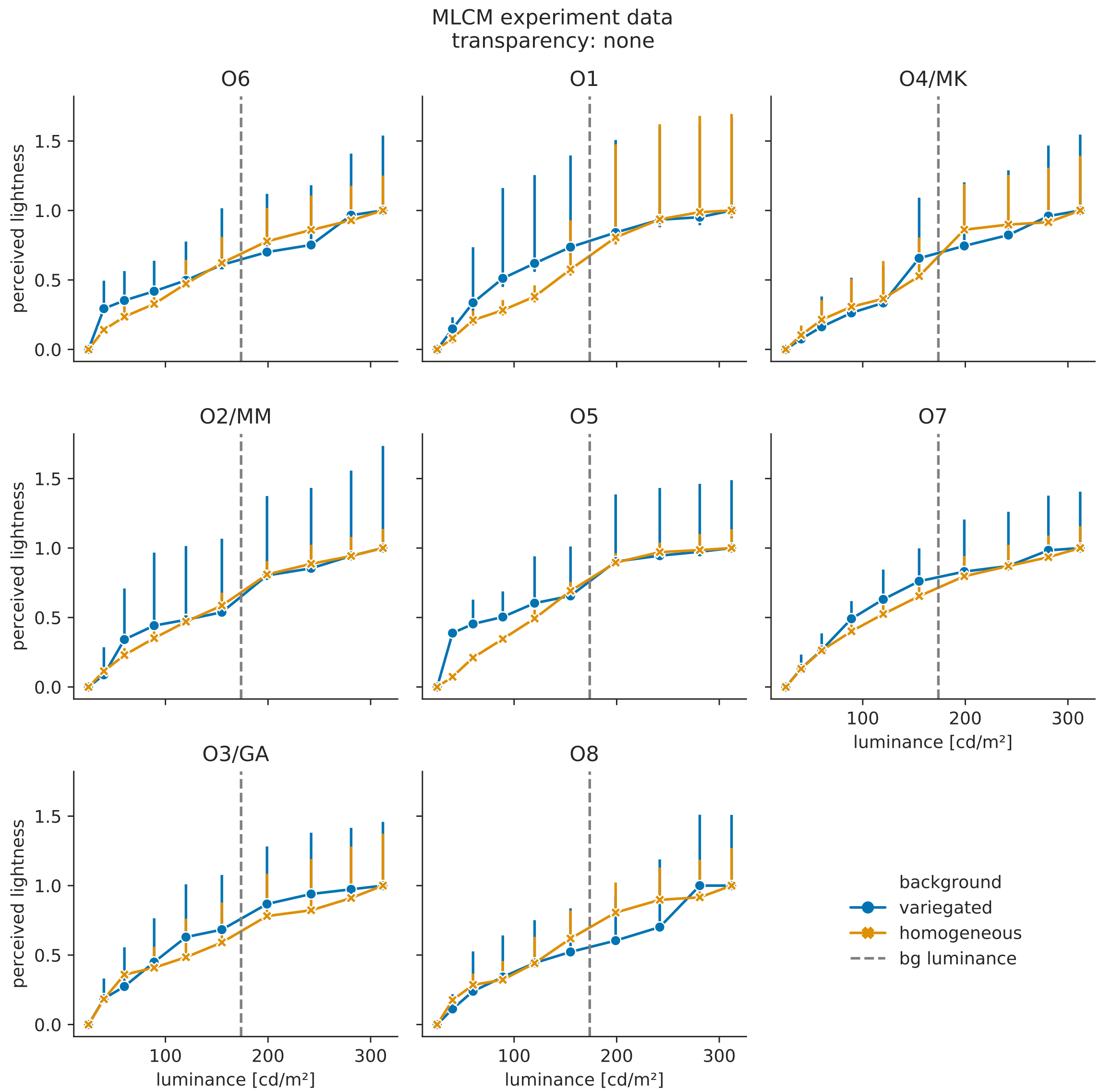

Figure 4.10: Perceptual scales measured with MLCM for all observers without any interposed transparency. The plot design and the legends are the same as in Figure 4.7. It is not possible to make a clear statement about the occurrence of CE in these scales.

4.3.1 Slope of Asymmetric Matching Data

Figure 4.11 shows the slope of the matching data without any interposed transparency. The slope significantly increases around the background luminance for the four observers whose data showed CE in Section 4.2.1 (O6, O4/MK, O2/MM, O5). Interestingly, the slope also in-

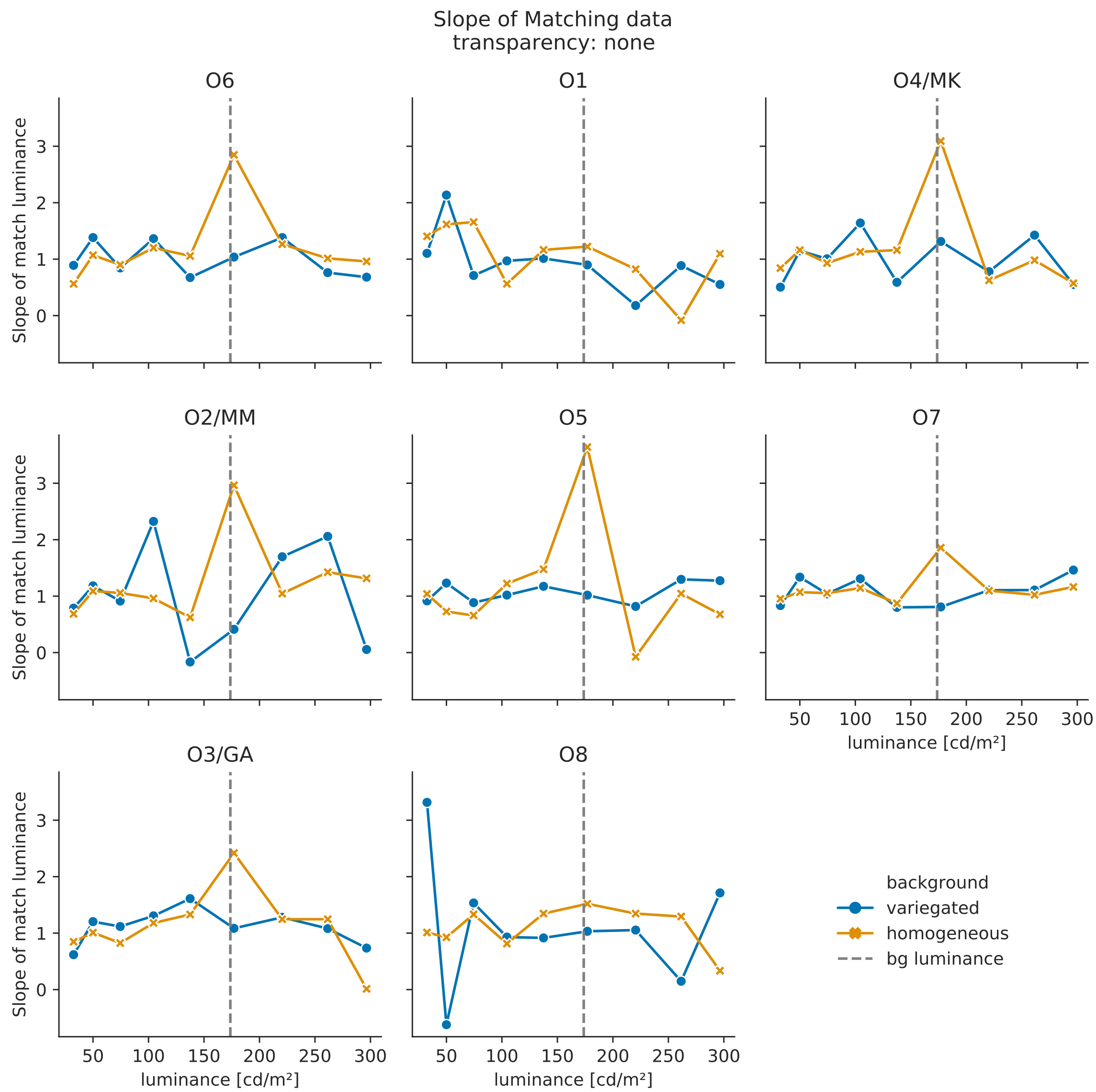

Figure 4.11: Slope of the matching data for all observers without any interposed transparency. The orange crosses show the slope of the data for the experiment with a homogeneous background, while the blue circles display the slope of the data for the experiment with a variegated background. In order to compare the two experimental conditions, the points are connected with straight lines. For O6, O4/MK, O2/MM, and O5, a significant increase in the slope occurs around the background luminance.

creases for O3/GA while the difference in the matching data is not so clearly distinguishable. This indicates that the slope should be analyzed as an additional metric for CE.

The slope of the matching data with a dark and light transparency is shown in Section .2.1 in the appendix.

4.3.2 Slope of MLDS Scales

In the following, I analyze the slope of the perceptual scales presented in Section 4.2.2. Figure 4.12, Figure 4.13, and Figure 4.14 show the slope for the experimental condition without any interposed transparency and with a dark and a light transparency, respectively.

Indeed, the slope increases around the background luminance for the four observers whose perceptual scales show CE (O6, O1, O4/MK, and O2/MM). For O5 and O7, whose perceptual scales also differ from each other in Figure 4.7, the slope does not increase at the background luminance. However, a difference in the slope's pattern is observable. The slopes of O3/GA and O8 are mostly the same among the background conditions. This was expected because the data almost overlapped in Figure 4.7.

Additionally, a pattern of the slope is recognizable. For the homogeneous background (blue), the slope values are mostly about the same value, except of the slope's increase around the background luminance. For the variegated background, the slope's pattern is different since the slope decreases monotonically. Humans are more sensitive in dark luminance ranges than in light luminance ranges – this leads to a pattern of perception, which is similar to the best-fitting power law, depicted by the dashed line in Figure 2.4. The significant decrease in the slope in the dark luminance range is a consequence of the non-linear human perception. Matching data does not recover this pattern.

In MLDS, the decreasing pattern is the weakest for the observers who do not show CE – namely O3/GA and O8. For these two observers, the slope's pattern is mainly identical between the two different background conditions.

Dark Transparency

Figure 4.13 shows the slopes of the perceptual scales measured in the experiment with a dark transparency (background). Just like without any transparency, the slope increases at the background luminance for O6, O1, O4/MK, and O2/MM. For these observers, the change in the slopes is clearly visible, although their perceptual scales did not cross in Figure 4.8, except for O4/MK. This increase should be considered as a unique feature to identify the occurrence of CE.

For O7, the slope of the homogeneous condition is also higher than that of the variegated condition. This correlates with the results of the distance metric (c.f. Table 4.1) which also indicated CE. However, for O5, the distance metric indicates the CE's occurrence, but the slope does not visibly increase in Figure 4.13. Presumably, the distance is large because the two scales are almost parallel and shifted along the y-axis. Hence, the distance is large between all data points but not as a consequence of the slope's increase and a resulting steepening as, for instance, for O4/MK.

The observable pattern is similar as in the slopes of the perceptual scales without any transparency. For the dark transparency with a homogeneous background, all the slope values are about the same or they increase until reaching the peak at the background luminance and then

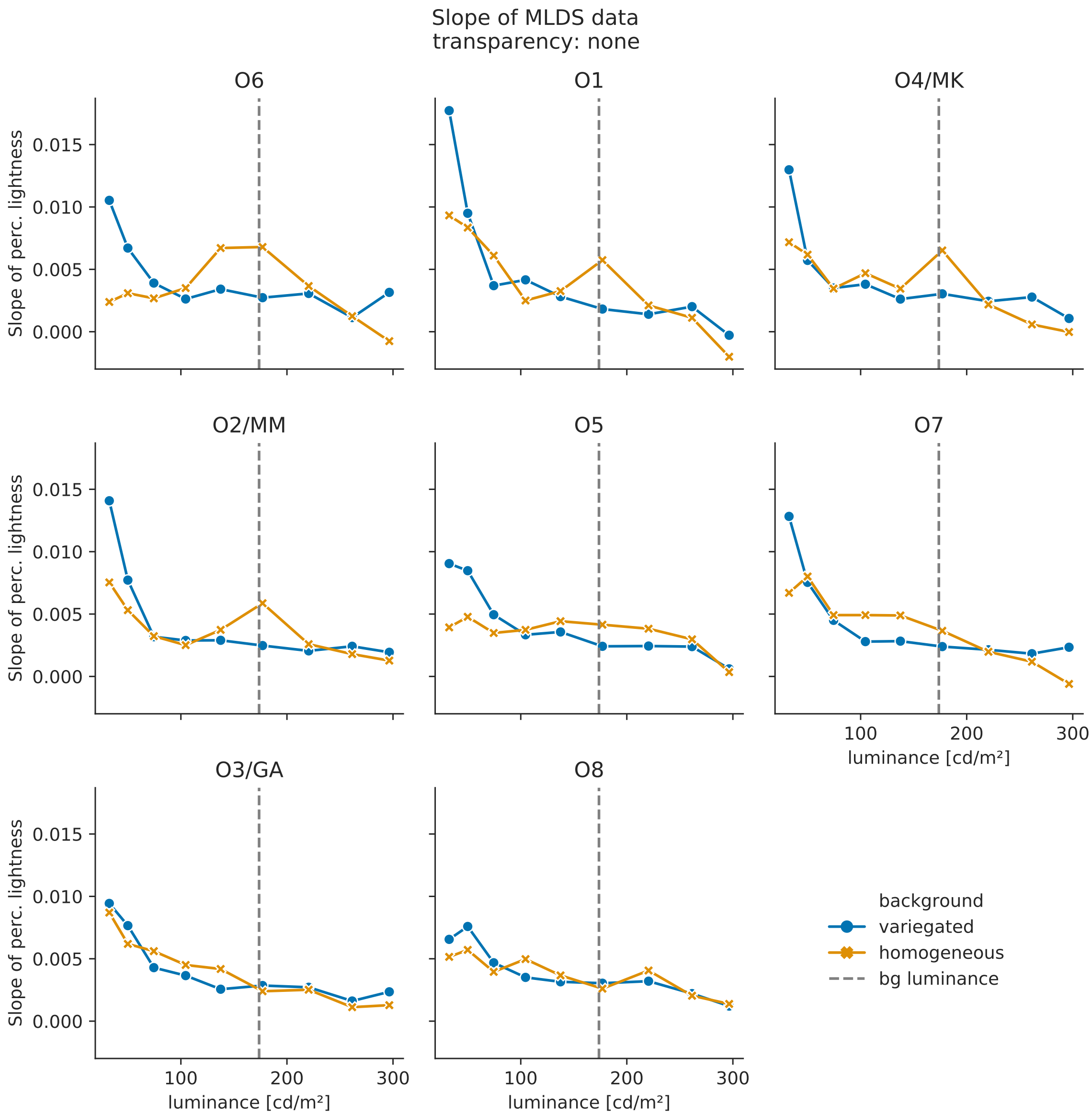

Figure 4.12: Slope of the perceptual scales measured with MLDS for all observers without any interposed transparency. The orange crosses show the slope of the data for the experiment with a homogeneous background, while the blue circles display the slope of the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions, the points are connected with straight lines. For O6, O1, O4/MK, and O2/MM, the slope for the homogeneous condition increases around the background luminance.

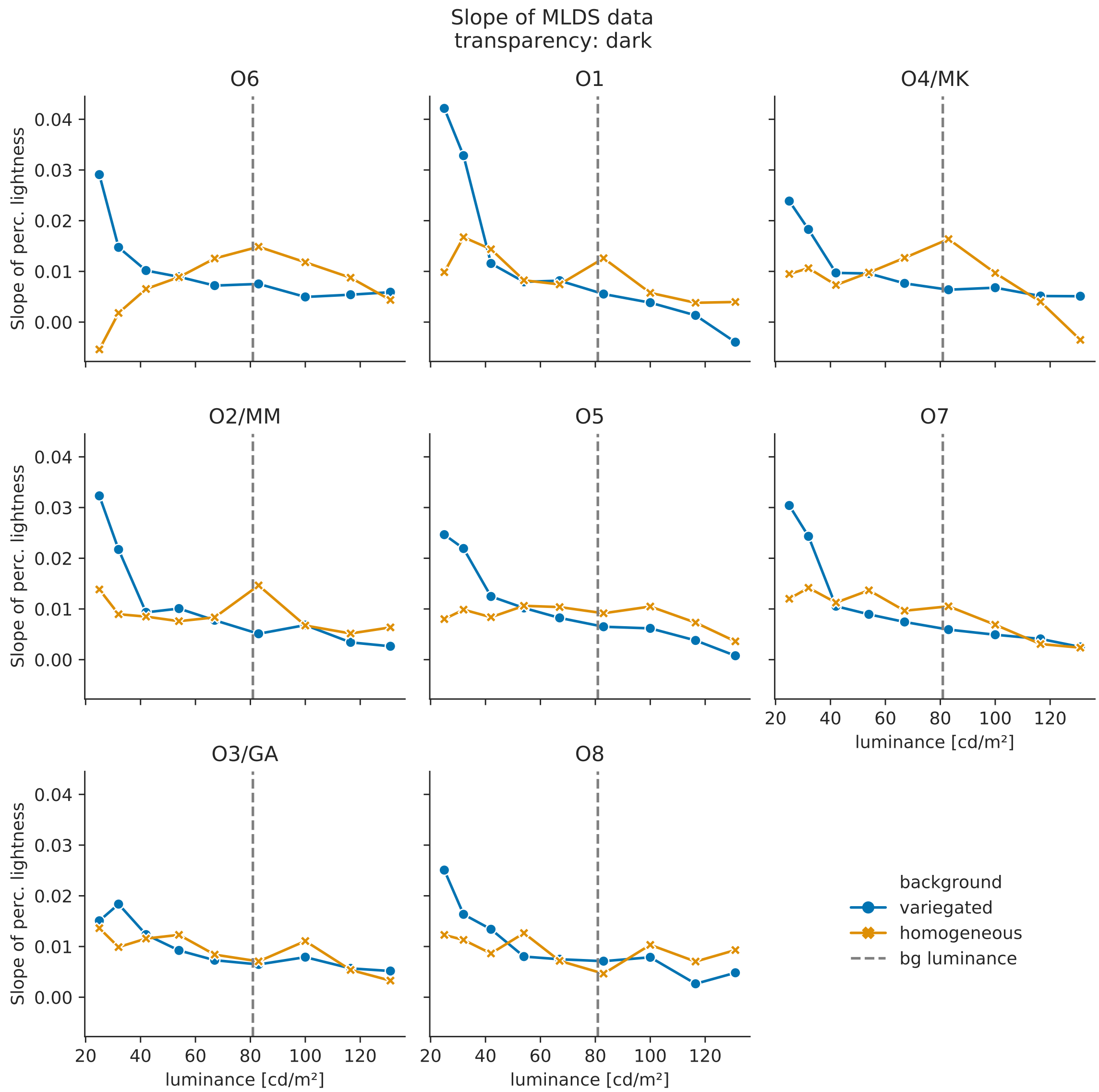

Figure 4.13: Slope of the perceptual scales measured with MLDS for all observers with a dark transparency. The plot design and the legends are the same as in Figure 4.12. For O6, O1, O4/MK, and O2/MM a significant increase in the slope occurs around the background luminance.

they decrease again. The slope in case of the experiment with a variegated background (orange) resembles the respective slope without any interposed transparency. It is monotonically decreasing corresponding to the non-linear human perception.

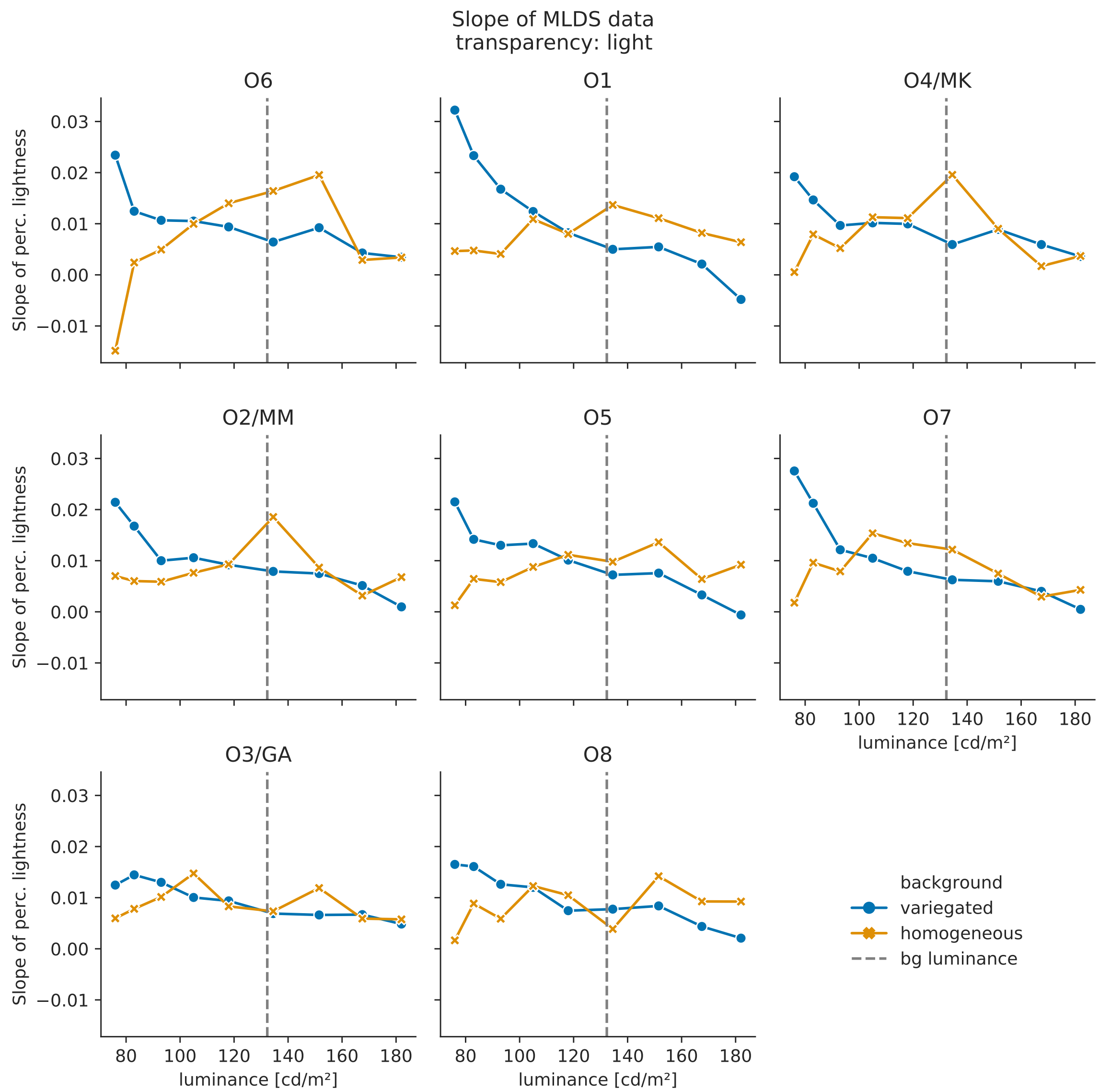

Figure 4.14: Slope of the perceptual scales measured with MLDS for all observers with a light transparency. The plot design and the legends are the same as in Figure 4.12. For O6, O1, O4/MK, and O2/MM, the slope increases around the background luminance. For O7, the slope of the homogeneous condition is also higher than that of the variegated condition at the background luminance; however, it does not increase compared to the previous values.

Light Transparency

Figure 4.14 shows the slopes of the perceptual scales measured in the experiment with a light transparency (background). It is difficult to detect a pattern in the slopes for the homogeneous background. Indeed, for the four observers whose slopes indicated CE beforehand (O6, O1, O4/MK, and O2/MM), the slope at the background luminance is higher than for the lower luminance values. However, for O6, for instance, the slope increases further until the background's consecutive luminance value. For O4/MK and O2/MM, the increase in the slope at the background luminance is the clearest. This steepening in the slope was also visible in the perceptual scales (c.f. Figure 4.9). For other observers, it is difficult to make a clear statement based on the slope of the perceptual scales.

4.3.3 Slope of MLCM Scales

Figure 4.15 shows the corresponding slope of the perceptual scales obtained with MLCM without any interposed transparency. As expected based on the perceptual scales in Figure 4.10, a steepening in the slope, as for MLDS, is not visible for MLCM. The slope of the MLCM data with interposed transparencies is presented in Section .2.2 in the appendix.

4.4 Summary of the Reanalysis

In the data of the asymmetric matching experiment, CE occurs for most observers. In the reanalysis of the perceptual scales obtained with MLDS, CE occurs for about half of the observers. The resulting difference between the perceptual scales for the respective background conditions (homogeneous, variegated) is the clearest in the experimental condition without any transparency. It should be noted that the observers that show CE do not coincide completely in the MLDS scales and the asymmetric matching data. For instance, O1's matching data does not have a non-linearity but in the MLDS scale the slope is increased at the background luminance.

For four observers – O6, O1, O4/MK, and O2/MM – the perceptual scales cross at the background luminance. This crossing correlates with a slope's increase of the scale measured with the homogeneous background. Plotting the slope confirms the visible slope's increase. The slope of O6, O1, O4/MK, and O2/MM increases considerably at the background luminance. The slope for the variegated background mostly decreases monotonically due to the higher sensitivity of humans in dark luminance range. The slope might be a useful indicator for the occurrence of CE.

However, in the MLDS scales for the experiment with a dark and light transparency, the slope's pattern is not as obvious as in case of the scales without any transparency. For a light transparency, the perceptual scales for the different background patterns are different between each other for O6, O1, O4/MK, and O2/MM. However, the "characteristic" steepening of the slope at the background luminance is absent, except for O4/MK and O2/MM.

Hence, the slope might be used as an indicator for the occurrence of CE. The increase in the slope indicates CE, but conversely, a constant slope does not prove that the observer does not have CE.

An additional indicator might be the distance between the perceptual scales for the different background patterns. Apparently, the distance is larger if CE occurs than when it does not. However, the indications based on the slope and the distance metric are contradictory in some

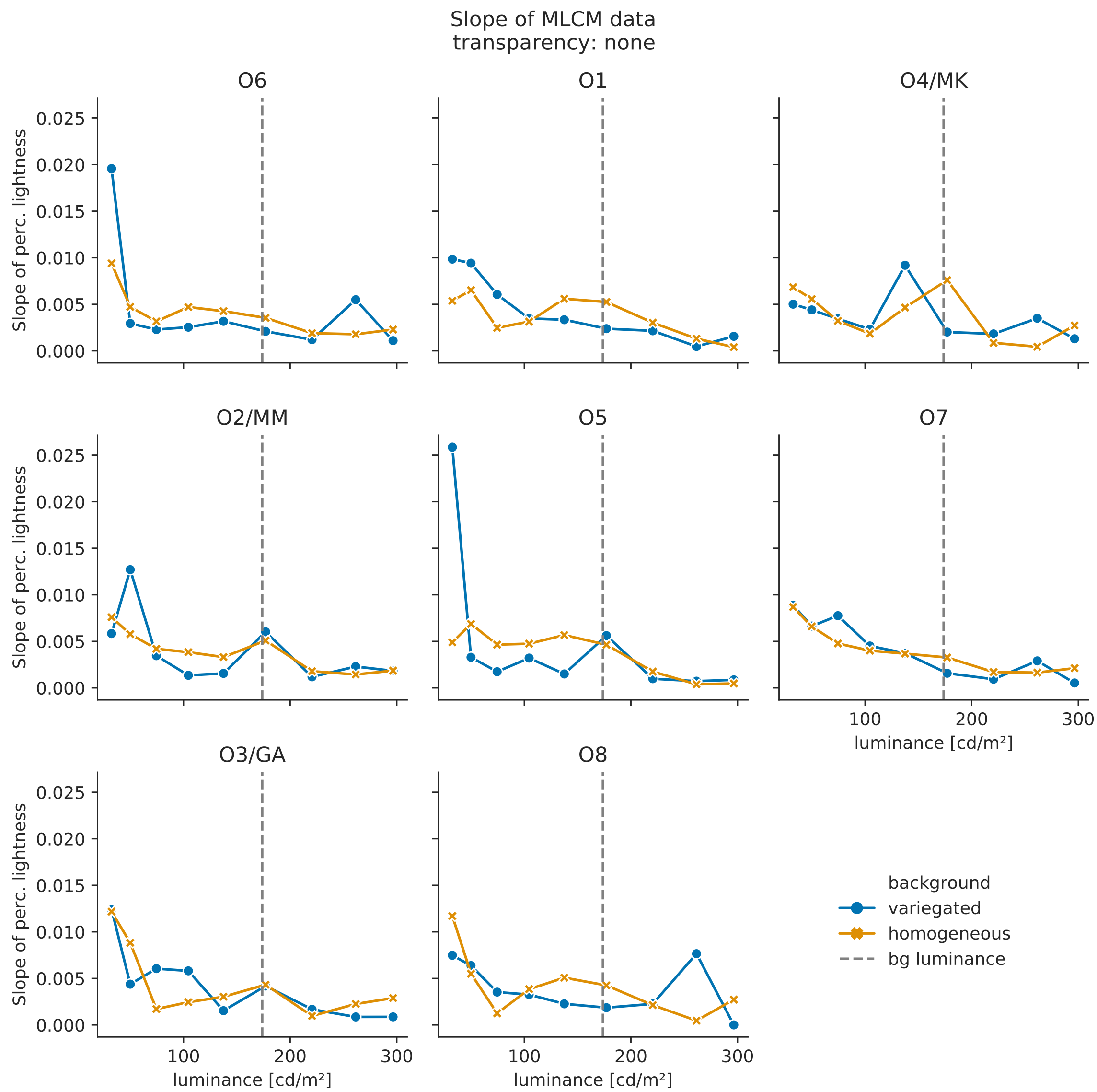

Figure 4.15: Slope of the perceptual scales measured with MLCM for all observers without any interposed transparency. The plot design and the legends are the same as in Figure 4.12. A steepening of the scale is not clearly visible in these plots.

cases. For instance, the distance of O5 is very large and therefore indicating CE but the slope of O5's scales is not increased.

In the MLCM data, CE was barely visible. In Section 4.2.3, I discussed two possible reasons for the effect's absence. One reason for that might be the high noise level of the data. Another reason might be that the observers only judge the relationship between the targets' appearances instead of the appearances' differences in this method. As a consequence, the steepening around the background luminance could be missed.

However, the CE's absence in this method does not influence the answer of my research question. Considering the task that the observers have to perform, MLDS is very different from asymmetric matching. The difference between two pairs of targets has to be compared and one element (x_2) is shared by both pairs. Consequently, the occurrence of CE in the data obtained with MLDS indicates that CE is a feature of the visual processing.

Chapter 5

Discussion

In this thesis, I investigated whether CE is an artifact of methods based on adjustment tasks or a feature of the visual system. For this, I simulated two experiments to investigate the effect of differently distributed luminance values on perceptual scales; I also reanalyzed the previous work by Aguilar and Maertens (2020).

In the simulations, described in Chapter 3, I simulated an asymmetric matching experiment and an MLDS experiment based on two different ground truth functions – with and without CE. I used three different spacings of the luminance values that are evenly distributed, coarse, and centered around the background luminance. Apart from that, I developed two metrics to evaluate the similarity with ground truth functions and the distinguishability of scales based on different ground truth functions. Additionally, I analyzed the slope of the simulated scales to determine whether the slope could possibly serve as an indicator for CE.

The spacings differed in their ability to correctly recover the ground truth functions. The evenly distributed spacing recovered both ground truth functions correctly. In contrast, perceptual scales obtained with the centered spacing only recovered the ground truth scale with CE correctly, but they were shifted vertically compared to the function without CE. Inversely, the coarse spacing correctly recovered the ground truth function without CE, but it did not recover the function with CE.

The second metric indicated the discriminability of the two perceptual scales based on ground truth functions with or without CE. In that metric the ground truth functions are considered discriminable. The scales of the evenly distributed spacing were taken as a reference to compare the other scales, as they correctly recovered the ground truth functions with a low standard deviation. According to the metric, the coarse spacing is always discriminable. However, the data show a high standard deviation which may cause situations where the scales are not discriminable. In contrast, the centered spacing is not well discriminable according to the metric. However, the data show almost no standard deviation, and the two scales are well discriminable by visual inspection.

Analyzing the slope of the simulated scales indicates that it might be a useful indicator for CE. The slope of the scale with CE is significantly increased at the background luminance. This tendency is the clearest for the centered spacing which also recovers the ground truth function with CE accurately.

In Chapter 4, I reanalyzed data from previous work by Aguilar and Maertens (2020). I reanalyzed asymmetric matching data and perceptual scales measured with MLDS and MLCM. CE occurred for most of the observers in the asymmetric matching data and for about half of the observers in the MLDS scales. CE was not observable in the MLCM scales. According to the metric, the distance was higher for the observers with CE than for those without. Hence, the distance metric might serve as a possible indicator for CE but still needs further improvement. Apart from that, I analyzed the slope of the data and the scales. For most of the observers whose data (scales) showed CE, the slope increased at the background luminance. Additionally, the slope's pattern differed between the background conditions. The slope for the variegated background was mostly monotonically decreasing, whereas the slope for the homogeneous background was mostly about the same value.

5.1 Open Questions

When simulating an MLDS experiment, as described in Chapter 3, I found that it is not always possible to accurately recover a ground truth scale with certain spacings. The spacing with luminance values centered around the background luminance could not correctly recover the ground truth scale without CE. However, it did recover the ground truth with CE. Presumably, the scales with CE are well recovered because the data points close to the background record the most particular characteristic of CE – the steepening of the slope. Inversely, the coarse spacing without any luminance values close to the background correctly recovered the ground truth with CE. However, it did not recover the ground truth without CE. This indicates that the data points around the background luminance are needed to capture the steepening of the slope caused by CE. Simultaneously, several luminance points distributed across the whole domain are needed in order to recover the exact shape of the function without CE. The even spacing recovered both ground truth functions. Apparently, one data point close to the background was sufficient to recover CE when the other data points are evenly distributed across the remaining range. In general, the simulation could serve as a tool to determine the appropriate spacing of the luminance values before conducting the actual experiment. Depending on the exact objective, different spacings might be useful.

As already mentioned, the metrics developed in Chapter 3 need further improvements in order to make better statements about the actual distinguishability of the scales and the scales' similarity with ground truth functions. One major constraint is the neglect of the variability in the simulation, captured by the standard deviation. Additionally, the metrics are only applied to the data points and do not interpolate between the data. It might be useful to consider the area below the scales to solve this issue. Apart from that, an additional reference would be beneficial in order to determine the distinguishability of the scales. The current metric only indicates whether two scales are discriminable compared to the even spacing but does not quantify how discriminable the scales are. As explained in Section 3.2.4 the scale's average distance to itself could serve as such an additional reference.

In the reanalysis in Chapter 4, CE occurred, as expected, in the matching data for the experimental condition with a homogeneous background. The effect occurred for more than half of the observers. Additionally, CE occurred in the reanalyzed scales obtained with MLDS by Aguilar and Maertens (2020). The slope of these scales increases around the background luminance. In MLDS, the observers judge the magnitude of the targets' difference in appearance instead of adjusting the targets' appearance like in asymmetric matching. Hence, the occurrence of CE in MLDS scales indicates that the effect is a processing characteristic of the visual system and not only a consequence of adjustment tasks. Ekroll et al. already found that adding variegation to the background would reduce CE – this has been confirmed in the reanalysis.

As expected, the steepening was only visible in the data for the homogeneous background. CE is observable in about half of the observers' MLDS scales. As mentioned in Chapter 2, Ekroll and Faul (2009) stated that the individual difference among observers seems to be significant. The magnitude of CE differs among the observers, but it also differs for a single observer among experimental conditions – e.g., when interposing transparencies. In the perceptual scales, CE was the clearest in the scales for the experimental condition without any transparency. Here the perceptual scales for different background conditions even crossed at the background luminance. In the scales for the experiments with dark or light transparencies, the scales did not cross but could still be visually distinguished.

As an additional indication for the occurrence of CE, I analyzed the slope of the data. For the observers whose data showed CE, the slope increased significantly at the background luminance. For the experimental conditions with an interposed dark or light transparency, the "characteristic" steepening was not as obvious as for the condition without any transparency. Still, the slope can serve as an indication for CE. The increase in the slope indicates CE; however, conversely, a constant slope does not prove that the observer does not have CE.

However, the indications of the slope and of the distance metric do not coincide for all observers regarding CE, as e.g. O5. For this observer the distance metric indicates CE whereas the slope does not increase significantly and hence, does not indicate CE. Consequently, the appropriateness of the indicators should be further investigated.

CE did not occur in the perceptual scales obtained with MLCM. As I already argued in Section 4.4, this could be attributed to the fact that the method measures only the relationship between two targets and does not quantify the amount of difference between them. MLCM uses paired comparison, which is based on the internal evaluation of $|x_2 - x_1|$. The observers compare the appearance of these two targets and judge which one is perceived lighter.

In contrast, MLDS uses the method of triads, where two intervals are compared. The observers compare the intervals $|x_3 - x_2|$ and $|x_2 - x_1|$ and indicate which interval is perceived larger. This interval comparison allows the establishment of a "metric", i.e. when $|x_3 - x_2|$ and $|x_2 - x_1|$ are confused equally often, this means that both intervals are about the same size. This indicates that x_3 is as far as x_1 from x_2.

Due to the paired comparison in MLCM, this type of estimation is not possible with this method. MLCM, as Thurstonian Scaling, must rely on stimuli that are close together and are often confused, to estimate a "metric" for the perceptual scale (L. Maloney & Yang, 2003). If the stimuli x_1 and x_2 are confused equally often this indicates that the stimuli are perceived equally light. If one stimulus is always judged as lighter it is not possible to deduce how much lighter this stimulus is. In this case, it is only possible to establish an ordered, but not a quantitative relationship between the two targets. Figure 5.1 illustrates this problem. By pairwise comparison within one context, the relationship $x_1 < x_2 < x_3$ and $o_1 < o_2 < o_3$ is

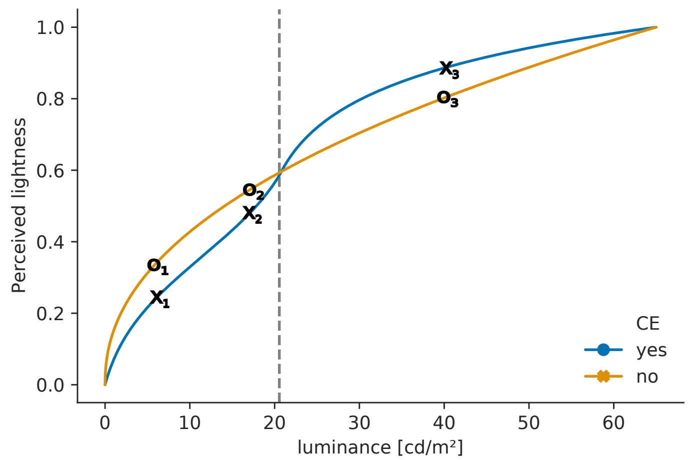

Figure 5.1: Example to illustrate the scaling in MLCM. By pairwise comparison within one context (blue or orange underlying function) the order $x_1 < x_2 < x_3$ and $o_1 < o_2 < o_3$ can be made. Cross-context comparisons lead to the order $x_1 < o_1 < x_2 < o_2 < o_3 < x_3$. However, the amount of difference between the targets is neglected.

obtained. By pairwise comparison across context, the relationship $x_1 < o_1 < x_2 < o_2 < o_3 < x_3$ can be obtained. However, depending on this relationship, the perceptual scales might be created in such a way that x_3 would just be marginally bigger than o_3 whereas in the actual scale x_3 could have been significantly bigger.

The amount of necessary trials in the statistical scaling methods increases rapidly with an increase of the data points. Consequently, in MLCM it is not feasible to use enough data points to establish a "metric" that qualifies the magnitude of difference between the perceptual scale values. For this reason, MLCM might not be able to recover CE as the slope's increase has to be recovered and therefore the magnitude of the difference is important. In general, MLDS seems to be better for further investigation of the effect.

5.2 Future Work

It might be true that not all observers have CE, but just as well it could be true that CE is very small for some observers and therefore is not always recovered. As the simulation shows it is not possible to accurately recover a ground truth function with CE with coarse spacing. Hence, if an observer's magnitude of CE is smaller, a tighter spacing might be necessary to recover CE in their scale. In future work, it would be interesting to test whether CE occurs for more observers in MLDS scales with a tighter spacing. Indeed, the experiment of Aguilar and Maertens (2020) was not designed to investigate CE and the luminance values were far away from the background luminance in luminance space. Hence, this might have caused the absence of CE in the scales of some observers. However, it is difficult to increase the amount of luminance values, as the amount of necessary trials increases rapidly and hence becomes unfeasible.

The occurrence of CE in the perceptual scales obtained with MLDS indicates that the effect is a feature of the visual system. In future work, it is necessary to understand the characteristics of the effect as CE influences the visual perception under certain conditions. Hence, for modeling the whole visual system, it should be investigated which conditions exactly cause CE and how to model the effect.

For instance, as stated in Ekroll et al. (2011), it would be interesting to scrutinize whether individual differences in the perception of CE correlate with differences in the perception of transparencies. This could indicate that CE and transparency perception share a common mechanism. Overall, it would be good to obtain experimental data for a higher number of observers.

Additionally, the exact influence of the amount of the background's variegation could be investigated. For that, the targets can be presented in front of a homogeneous background to assess the observer's magnitude of CE. Next, the background can be gradually variegated and the development of CE can be recorded. This could provide information about how much variegation is needed to influence CE.

Another open question is whether CE develops over time. For the simultaneous contrast, Ekroll and Faul (2009) implemented an instantaneous spatial mechanism with Takasaki's formula. They stated that "phenomena characteristic of crispening such as gamut expansion effect [$\cdots$] are typically observable immediately upon inspection of a demonstration" (Ekroll & Faul, 2009). It would be interesting to explore in future work whether this assumption holds. In order to test the CE's time dependency, the stimuli described in Section 4.1.1 could be used. The center-surround (CS) stimulus (homogeneous background) provokes a perception with CE, while the variegated checkerboards do not provoke such perception. Both targets should be close to the background luminance and should be similar to each other. When presenting these targets in front of the two backgrounds, the targets in front of the homogeneous background look more different due to CE. The observer has to judge which target is lighter. This task should be easier for the targets in front of the homogeneous background. The stimuli are, then, presented for different time intervals between 50 and 1000 ms. One can analyze whether the observers' answers are more often correct for the CS stimulus than for the variegated checkerboard. This would indicate that the discriminability of the targets in front of the homogeneous background is enhanced compared to the variegated checkerboard. If the targets could be discriminated in a better way for the CS stimulus, CE is probably a built-in feature of an early stage of humans' visual processing. Conversely, an enhanced discriminability, which occurs only after a certain time, would indicate that CE builds up over time and is part of a later stage in the visual processing.

Since my results indicate that CE is a feature of the visual system, it should be included in future models of the system. For this, it is necessary to understand the effect's characteristics; some exemplary ones are described above. In general, future models should aim to regard CE and similar phenomena in their predictions. The same applies for scaling methods. Scaling methods have the advantage over matching methods that they create a perceptual scale. However, it is important that methods that measure the observers' perceptual scales are able to recover CE and similar phenomena. MLCM, for instance, seems unsuitable to measure CE and therefore, should be handled with care. Aguilar and Maertens (2020) state that MLCM seems useful for measuring perceptual scales across different viewing conditions. Nevertheless, the absence of CE in the reanalysis of the MLCM scales is a drawback. The possible reasons, which lead to the fact that CE cannot be measured with MLCM, could also prevent other phenomena from being seen. In contrast, MLDS is able to recover CE and the method seems more beneficial than MLCM. However, MLDS has its own disadvantages. For instance, the comparison of scales between viewing conditions is complicated since all scales are anchored to zero by design (Aguilar & Maertens, 2020). Undoubtedly, both methods need further investigation to clarify their appropriateness for measuring perceptual scales in certain conditions. Possibly, each method is only useful when studying certain problems; e.g., MLDS is useful when a metric scale with a quantitative order of the data points is needed whereas MLCM is useful when comparing conditions. In future work on CE the method should be chosen carefully.

Additionally, the spacing of the data points should be chosen with care. The simulations could serve as a tool to determine a beneficial spacing of the data points before conducting the actual experiment. This could be also done in experiments that investigate other effects than CE.

Overall, it is necessary to take CE into consideration in future experiments since the effect might lead to non-linearities in matching data and scales. Therefore, CE should be always taken into account, also in experiments that do not intend to investigate the effect. Experimenters should give consideration to CE when interpreting their data.

Chapter 6

Conclusion

In the course of this thesis, I analyzed whether CE is a consequence of adjustment tasks or if it is a built-in feature of the visual processing. For this purpose, I reanalyzed the perceptual scales obtained by Aguilar and Maertens (2020) with MLDS and MLCM. Additionally, I simulated an asymmetric matching and an MLDS experiment with differently distributed luminance values.

In the simulations, I investigated the influence of the spacing on the occurrence of CE. More precisely, I investigated whether a perceptual scale based on a ground truth scale with CE is always distinguishable from a perceptual scale based on a ground truth scale without CE. Additionally, I analyzed the similarity of these perceptual scales to their corresponding ground truth functions. Nevertheless, the metrics for the evaluation need further improvement. However, the results clearly indicate that multiple data points around the background luminance were needed in order to recover a ground truth function with CE.

In the reanalysis, CE occurs in the scales of about half of the observers. As mentioned above, the individual differences among the observers seem significant. Additionally, the experiment of Aguilar and Maertens (2020) was not designed to measure CE, and the effect might be absent due to the experiment's design. This assumption is supported by the results of the simulations, as data points close to the background luminance were needed to correctly recover the ground truth function with CE. This implies that CE might be absent for some observers because the spacing was too coarse to recover the observer's individual magnitude of CE. Overall, the results of this reanalysis indicate that CE is a characteristic of the visual processing and that the effect does not depend on adjustment tasks.

Additionally, I analyzed the data's slope in order to investigate whether this could be an indicator for the occurrence of CE. For most observers whose data showed CE, the slope was increased at the background luminance. The slope seems to be an appropriate indicator for CE, but it is not sufficient as a single piece of evidence. Another possible indicator for CE is the distance metric that indicates the distinguishability of two scales.

In order to define the characteristic patterns and indicators of CE, future research should analyze the nature and cause of the effect. In Chapter 5, I presented several suggestions for future work. For defining the characteristics of CE, the influence of variegating the background and the time dependency of the effect could be investigated.

List of Abbreviations

CE Crispening Effect

bg background

centbg luminance values centered around the background

MLDS Maximum Likelihood Difference Scaling

MLCM Maximum Likelihood Conjoint Measurement

CS center-surround

Bibliography

Aguilar, G. & Maertens, M. (2020, April). Toward reliable measurements of perceptual scales in multiple contexts. *Journal of Vision*. doi:10.1167/jov.20.4.19

Bressan, P. & Actis-Grosso, R. (2006). Simultaneous lightness contrast on plain and articulated surrounds. *Perception*, 445–452. doi:10.1068/p5247

Brown, R. O. & MacLeod, D. I. A. (1997, November). Color appearance depends on the variance of surround colors. *Current Biology*, *7*(11), 844–849. doi:10.1016/s0960-9822(06)00372-1

Ekroll, V. & Faul, F. (2009, September). A simple model describes large individual differences in simultaneous colour contrast. *Vision Research*, *49*(18), 2261–2272. doi:10.1016/j.visres.2009.06.015

Ekroll, V. & Faul, F. (2013, March). Transparency perception: the key to understanding simultaneous color contrast. *J. Opt. Soc. Am. 30*(3), 342–352. doi:10.1364/JOSAA.30.000342

Ekroll, V., Faul, F., & Niederee, R. (2004, July). The peculiar nature of simultaneous colour contrast in uniform surrounds. *Vision Research*, *44*, 1765–1786. doi:10.1016/j.visres.2004.02.009

Ekroll, V., Faul, F., & Wendt, G. (2011, February). The strengths of simultaneous colour contrast and the gamut expansion effect correlate across observers: evidence for a common mechanism. *Vision Research*, *51*(3), 311–322. doi:10.1016/j.visres.2010.11.009

Evans, R. M. (1974, November). *The perception of colour*. New York: John Wiley & Sons Inc.

Faul, F., Ekroll, V., & Wendt, G. (2008, March). Color appearance: the limited role of chromatic surround variance in the “gamut expansion effect”. *Journal of Vision*, *8*(3), 30–30. doi:10.1167/8.3.30

Kane, D. (2017). Can ’crispening’ be explained by contrast gain? In *Human vision and electronic imaging 2017* (pp. 182–187). Burlingame, CA, USA: Ingenta. doi:10.2352/ISSN.2470-1173.2017.14.HVEI-141

Kane, D. & Bertalmio, M. (2019). A reevaluation of whittle (1986, 1992) reveals the link between detection thresholds, discrimination thresholds, and brightness perception. *Journal of Vision*, *19*(1). doi:10.1167/19.1.16

Kingdom, F. A. A. & Prins, N. (2016). *Psychophysics - a practical introduction* (2nd ed.). USA: Academic Press.

Knoblauch, K. & Maloney, L. T. (2012). *Modeling psychophysical data in r*. Use R. New York: Springer. doi:10.1007/978-1-4614-4475-6

Maloney, L. & Yang, J. (2003, February). Maximum likelihood difference scaling. *Journal of Vision*, *3*(8), 573–85. doi:10:1167/3.8.5

Masin, S. C. & Idone, A. M. (1981). Studio sperimentale sulla percezione della trasparenza con figura e sfondo achromatici e omogenei. *Giornale Italiano di Psicologia*, *8*, 265–277.

Rogers, B. (2017, October). *Perception - a very short introduction*. New York, USA: Oxford University Press.

Smith, V. C. & Pokorny, J. (1996, October). Color contrast under controlled chromatic adaptation reveals opponent rectification. *Vision Research*, *36*(19), 3087–3105. doi:10.1016/0042-6989(96)00035-1

Soranzo, A., Lugrin, J.-L., & Wilson, C. J. (2013, June). The effects of belongingness on the simultaneous lightness contrast: a virtual reality study. *Vision Research*, *86*, 97–106. doi:10.1016/j.visres.2013.04.012

Takasaki, H. (1966, April). Lightness change of grays induced by change in reflectance of gray background. *Journal of the Optical Society of America*. *56*(4), 504–509. doi:10.1364/JOSA.56.000504

Thomson, G. & Macpherson, F. (2017, July). Adelson's checker-shadow illusion. *The Illusions Index*. Retrieved from https://www.illusionsindex.org/ir/checkershadow

von Kries, J. A. (1905). Handbuch der physiologie des menschen. (Chap. Die Gesichtempfindungen, Vol. 3, pp. 109–282). Braunschweig: Vieweg.

Whittle, P. (1992, August). Brightness, discriminability and the crispening effect. *Vision Research*, *32*(8), 1493–1507. doi:https://doi.org/10.1016/0042-6989(92)90205-W

Wiebel, C. B., Aguilar, G., & Maertens, M. (2017, April). Maximum likelihood difference scales represent perceptual magnitudes and predict appearance matches. *Journal of Vision*, *17*, 1–14. doi:10.1167/17.4.1

Appendix

.1 Simulations

In Chapter 3, I simulated an asymmetric matching and an MLDS experiment based on two different ground truth functions. Additionally, I changed the spacing of the luminance values. Table 1 provides the luminance values used in the particular spacings. The spacing whose luminance values were centered around the background luminance in Chapter 3 is labeled with “centbg”. The coarse spacing whose data is shown in Chapter 3 is labeled with “coarse”. The plots in the following sections use the same labels that are defined in this table.

	even	centbg	coarse	coarse1	coarse2	coarse3
1	0	1	0	1	1	0
2	7.2	5	5	5	5	5
3	14.4	18	10	10	10	10
4	21.6	19	35	15	30	15
5	28.8	20	40	25	40	20
6	36.1	21	45	34	45	35
7	43.3	22	50	40	50	45
8	50.5	25	55	46	55	50
9	57.7	40	60	55	60	55
10	65	65	65	65	65	65

Table 1: Luminance values, $[cd/m^2]$, used in the different spacings. The label “centbg” indicates the spacing where the luminance values were centered around the background luminance, presented in Chapter 3. The label “coarse” means the coarse spacing presented in Chapter 3.

.1.1 Asymmetric Matching

Figure 1 shows all matching data of the asymmetric matching simulation . Here, all six spacings are shown. The match luminance (y-axis) is plotted against the target luminance (x-axis). The black, red, and blue line show the data for the target presented in front of a light, mid gray, and dark background, respectively. The vertical lines at the markers show the standard deviation of the matching data over 100 runs. The left column shows the data obtained with a mid gray match background. The middle and right columns show the data for a dark and light gray match background, respectively.

It is noticeable that the matching data do not change depending on the match background.

.1.2 Symmetric Matching

Figure 2 shows the matching data of the “symmetric” matching simulation for all spacings. The plot design and the arrangement are the same as in Figure 1. It is noticeable that the matching data changes depending on the match background in this simulation. The matching data, where the luminance of the target’s and match’s background is the same, is a straight 45° line through origin. The other two lines show non-linearities at the respective target background luminance.

.1.3 MLDS

Figure 3 and Figure 4 show the simulated MLDS scales and the slope of these scales for all spacings. For one particular spacing, the scales of the simulation based on ground truth functions with or without CE are always plotted together in one plot. The orange crosses represent the scale (scale’s slope) for the estimation based on the ground truth function without CE. The blue circles show the scale (scale’s slope) for the estimation based on the ground truth function with CE. In Figure 3 the continuous lines represent the ground truth functions. In Figure 4 the slope values are connected with straight lines in order to facilitate the comparison of the two conditions. The dashed gray line indicates the hypothetical background luminance.

The slope’s increase at the background is the clearest for the centered spacing. However, it is visible for all other spacings as well.

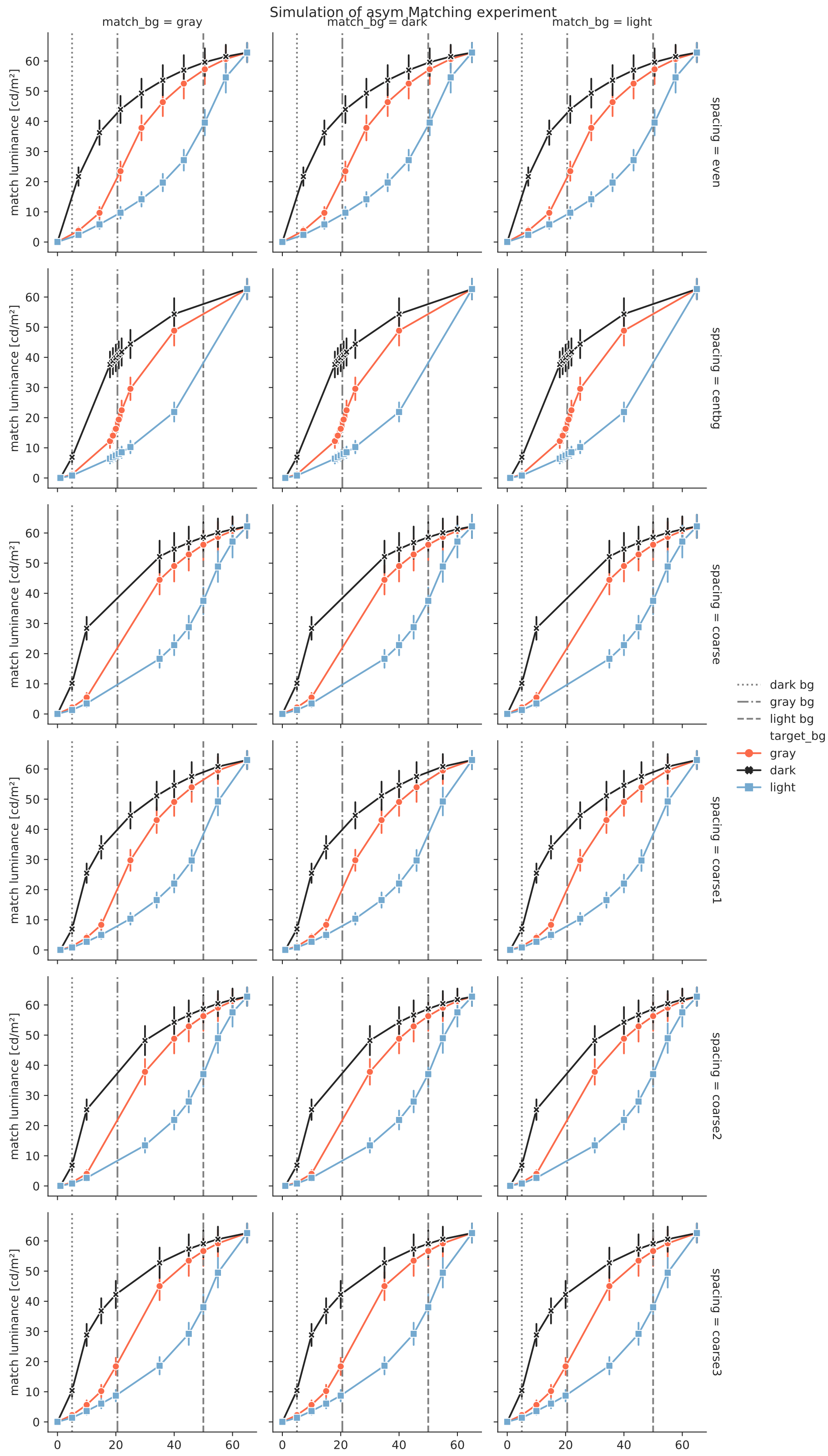

Figure 1: Asymmetric matching data of simulation for different spacings. The plot design and the legends are the same as in Figure 3.2. The data are not influenced by the different luminance of the match's background. The left, middle, and right panel show the matching data obtained with a mid gray, dark gray, and light gray match background, respectively.

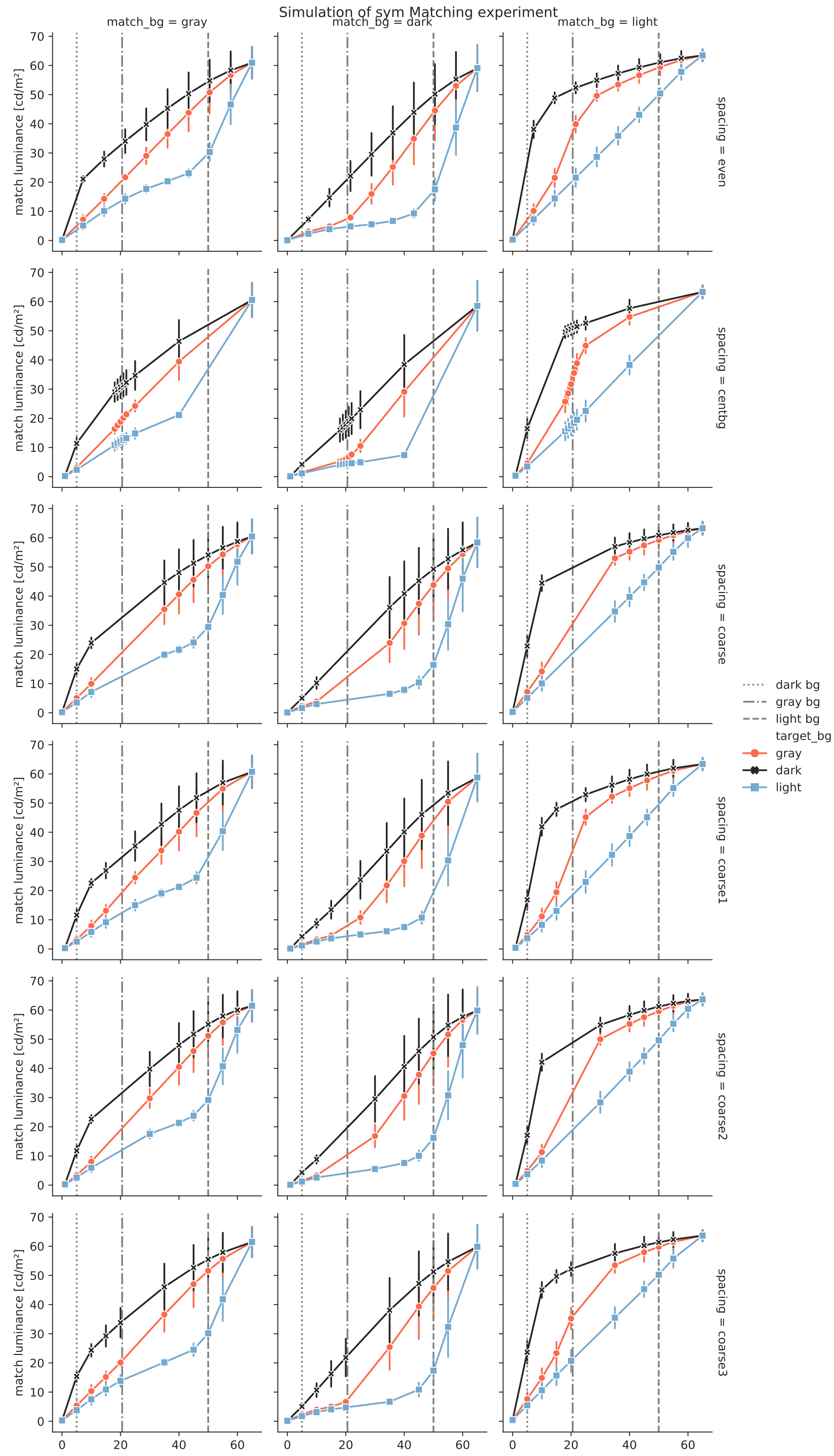

Figure 2: "Symmetric" matching data of simulation for different spacings. The plot design and the legends are the same as in Figure 3.2. The data are influenced by the different luminance values of the match's background. The left, middle, and right panel show the matching data obtained with a mid gray, dark gray, and light gray match background, respectively. The position of the non-linearity varies depending on the match's background.

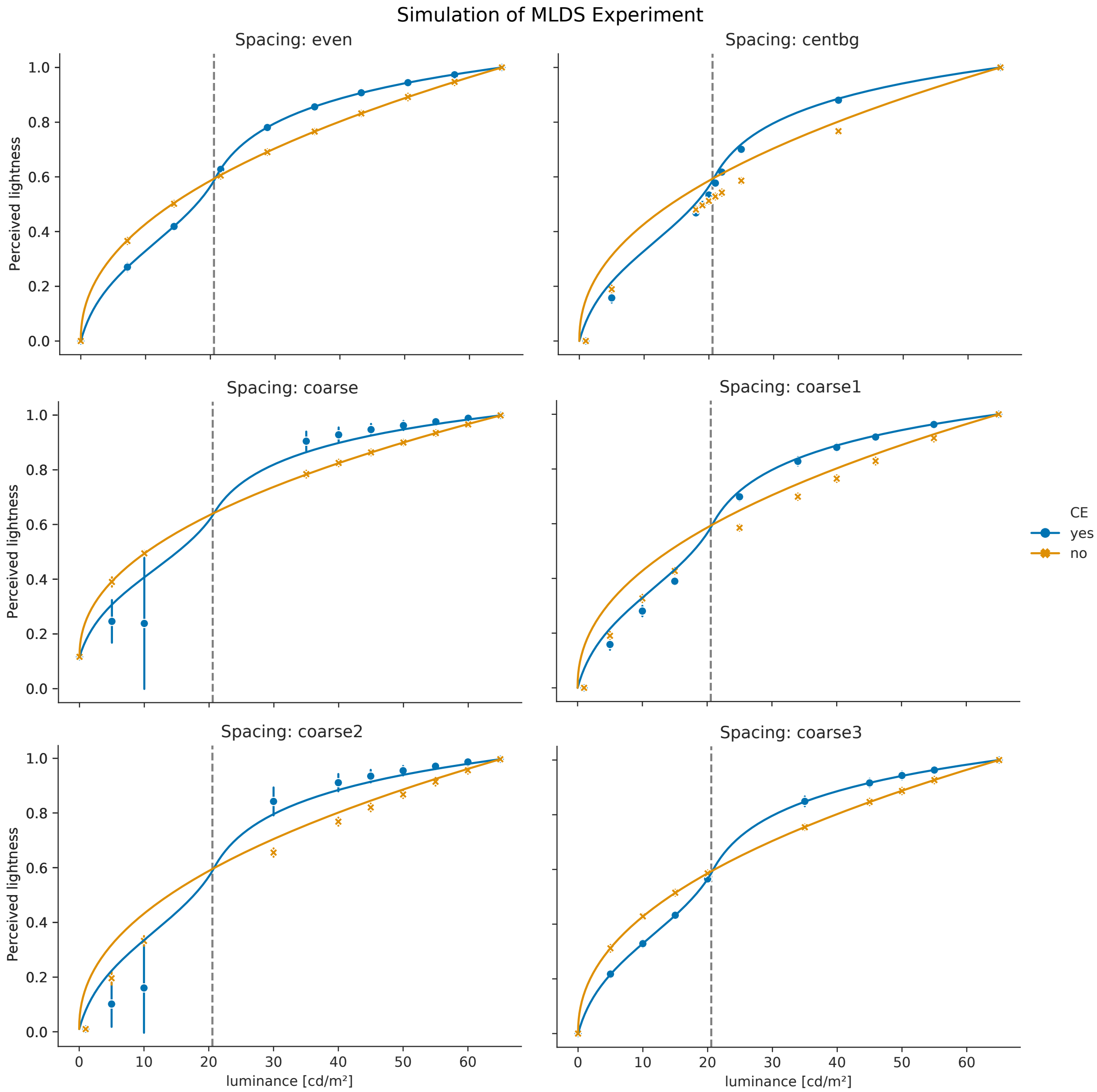

Figure 3: Perceptual scales obtained in the simulated MLDS experiment. Each plot shows the scales measured with a different spacing of luminance values. The perceived lightness (y-axis) is plotted against luminance (x-axis). The orange crosses and blue circles show the perceptual scale values based on ground truth without CE and with CE, respectively. The continuous orange and blue lines show the corresponding ground truth functions.

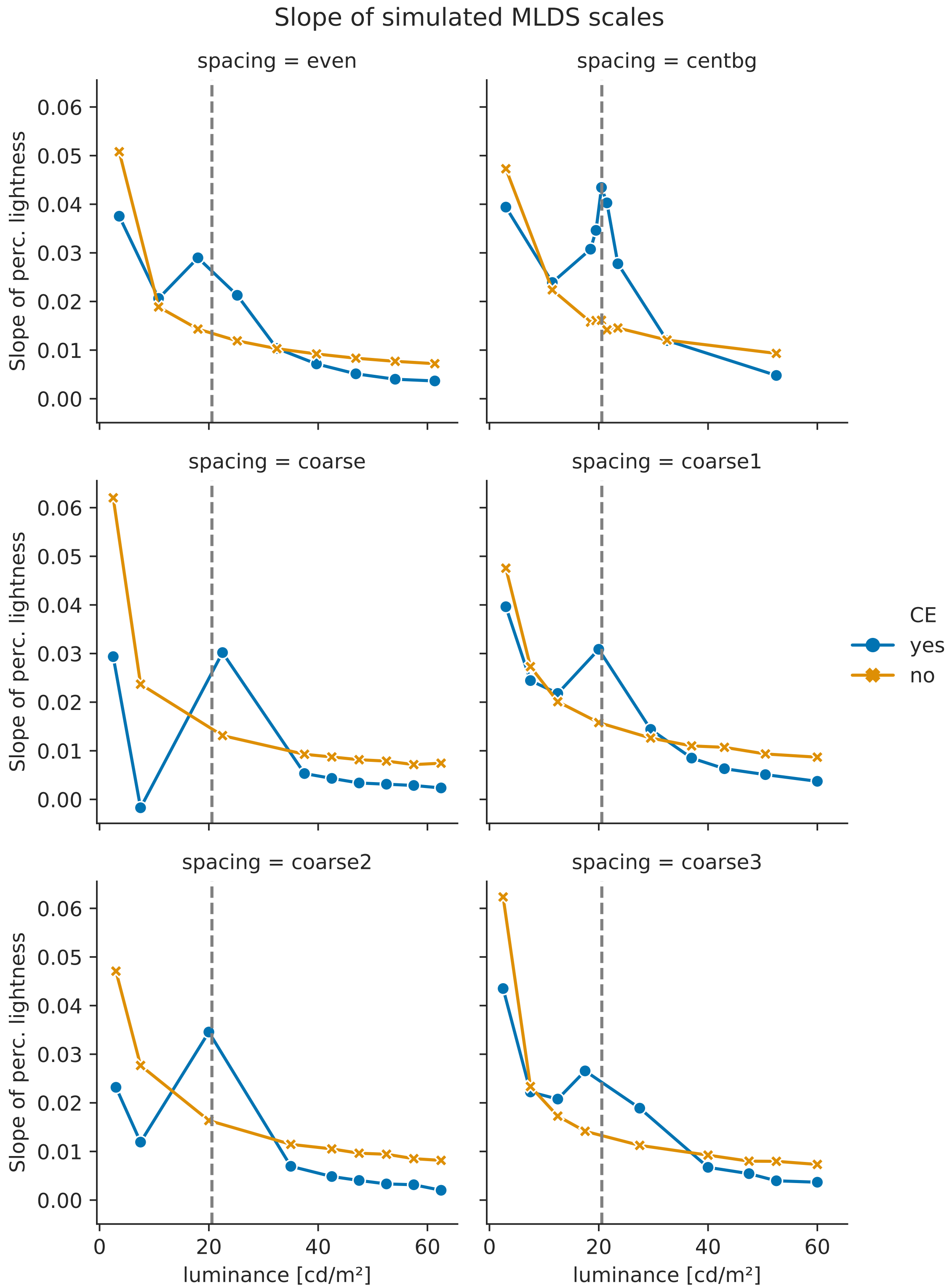

Figure 4: Slope of MLDS scales obtained in a simulation with different spacings. The slope of perceived lightness (y-axis) is plotted against luminance (x-axis). The orange crosses and blue circles show the slope of the scales based on ground truth without and with CE, respectively. In order to facilitate the comparison of the two ground truth functions the points are connected with straight lines.

.2 Reanalysis of Previous Work

As already described above, the observer labels (O1, $\cdots$, O8) in the following plots are the same as in Aguilar and Maertens (2020). In the following, I present the asymmetric matching data and the MLCM scales measured with a dark and light transparency. Additionally, I show the corresponding slope.

.2.1 Asymmetric Matching Experiment

In the following, I present asymmetric matching data measured with an interposed dark and light transparency and the corresponding slopes.

Data

Figure 5 and Figure 6 show the asymmetric matching data measured with a dark and light transparency, respectively. The matching luminance (y-axis) is plotted against the target luminance (x-axis). The orange crosses show the data for the experiment with a homogeneous background, while the blue circles display the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the data points are connected with straight lines. The dashed gray line indicates the target's background luminance.

The non-linearity of the data is not as obvious as in the data obtained without any transparency. In Figure 5 (dark transparency), a non-linearity at the background luminance is visible for O6, O4/MK, O2/MM, O5, O7, and O8. This is interesting as O8 does not show CE in the MLDS scales. As in the data measured without any transparency, O1 and O3/GA do not show the effect. However, in Figure 6 (light transparency) the non-linearity is not clearly visible for any of the observers.

Slope

Figure 7 and Figure 8 show the slope of the matching data presented previously. The orange crosses show the slope of the data for the experiment with a homogeneous background, while the blue circles display the slope of the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the points are connected with straight lines.

In Figure 7 (dark transparency), the slope is increases visibly at the background luminance for all the observer whose data showed CE. Just as for the matching data in Figure 6, Figure 8 shows no clear tendency for the occurrence of CE in the slope of the matching data for the light transparency.

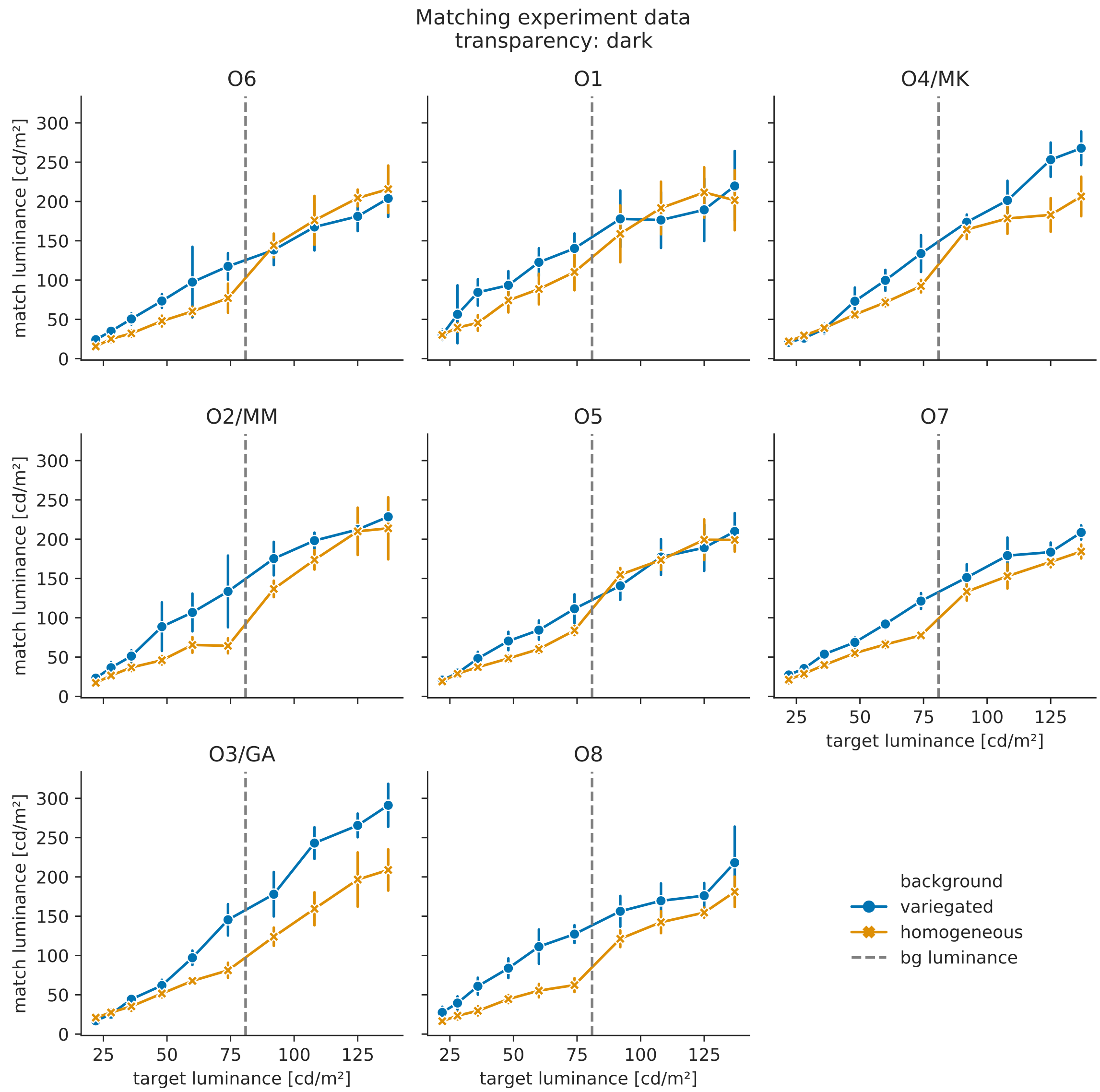

Figure 5: Matching data for all observers with a dark transparency. The orange crosses show the data for the experiment with a homogeneous background, while the blue circles display the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the data points are connected with straight lines.

.2.2 MLCM

In the following, I present scales measured with MLCM with an interposed dark and light transparency and the corresponding slopes.

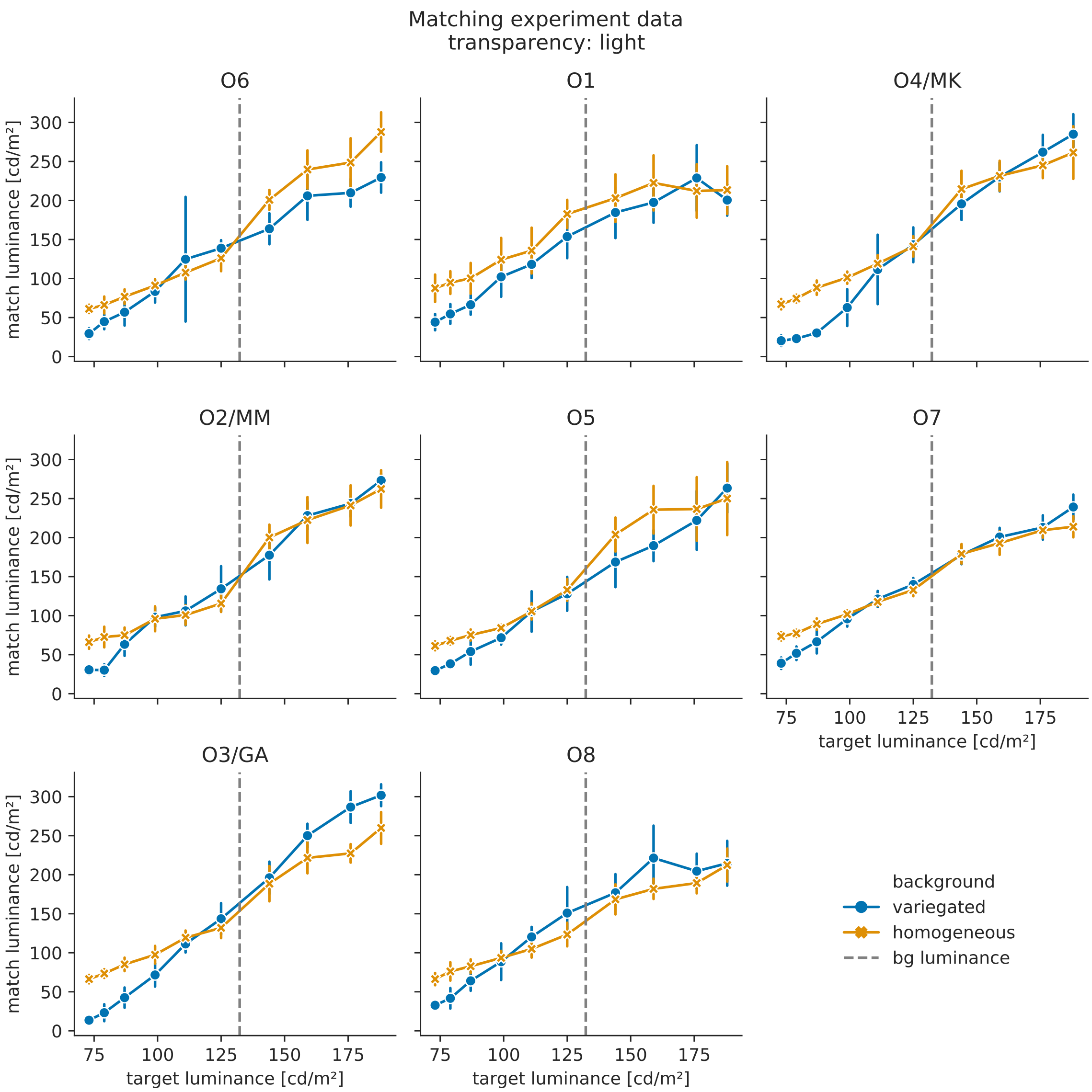

Figure 6: Matching data for all observers with a light transparency. The plot design and the legend are the same as in Figure 5. For O6, O4/MK, O2/MM, and O5, a strong non-linearity occurs in the data for the homogeneous experimental condition.

Data

Figure 9 and Figure 10 show the scales for a dark and light transparency, respectively. The perceptual scales are plotted against luminance. The orange crosses show the perceptual scale values for the experiment with a homogeneous background, while the blue circles display the

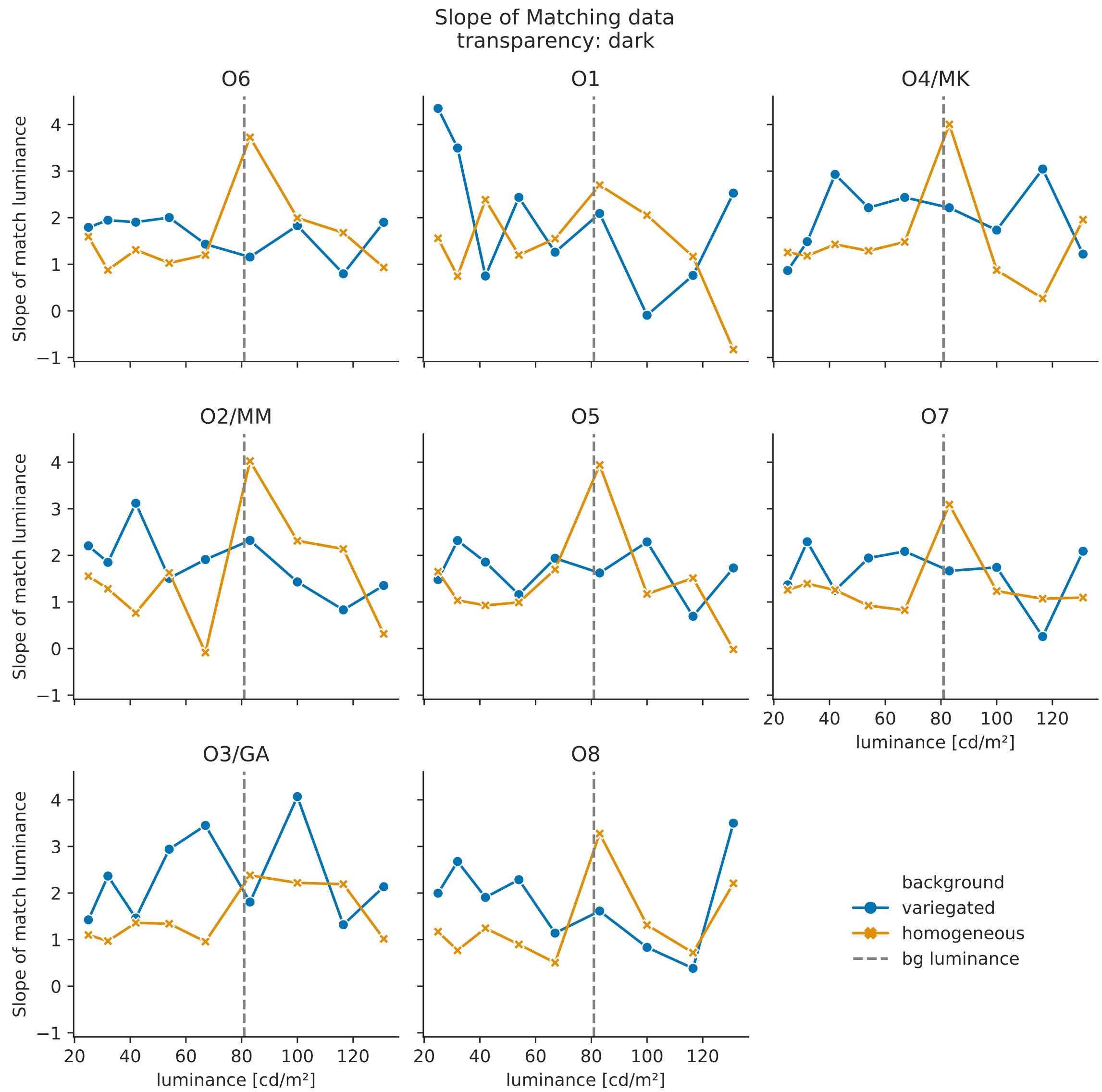

Figure 7: Slope of the matching data for all observers with a dark transparency. The orange crosses show the slope of the data for the experiment with a homogeneous background and the blue circles the slope of the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the points are connected with straight lines.

data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the perceptual scale values are connected with straight lines. As well as in the MLCM scales measured without any transparency, CE does not occur.

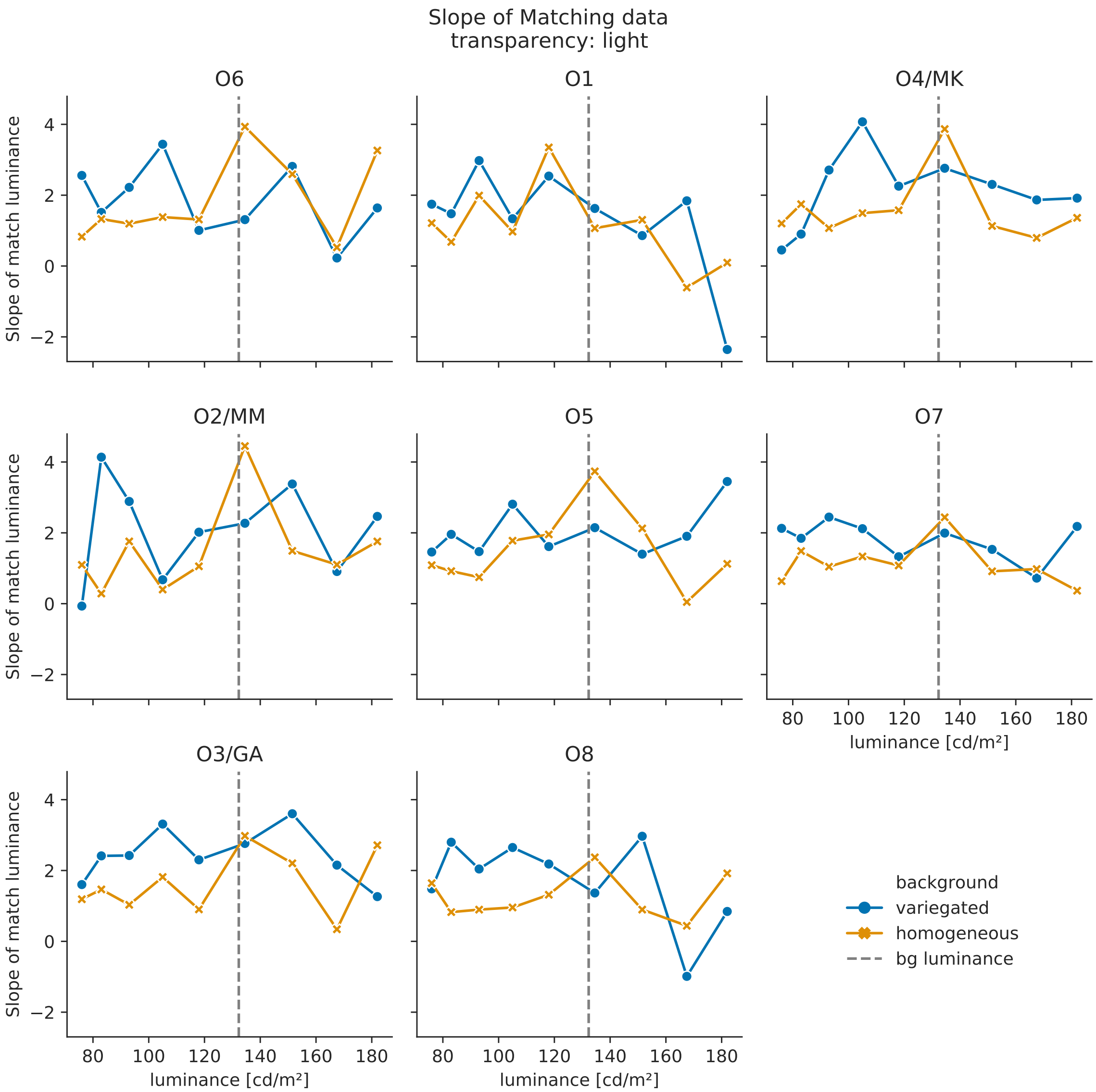

Figure 8: Slope of the matching data for all observers with a light transparency. The plot design and the legend are the same as in Figure 7.

Slope

Figure 11 and Figure 12 show the slopes of the scales for a dark and light transparency, respectively. The orange crosses show the slope of the data for the experiment with a homogeneous background, while the blue circles display the slope of the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the points are connected with straight lines. Like in the MLCM scales, the slopes do not show a clear tendency indicating CE.

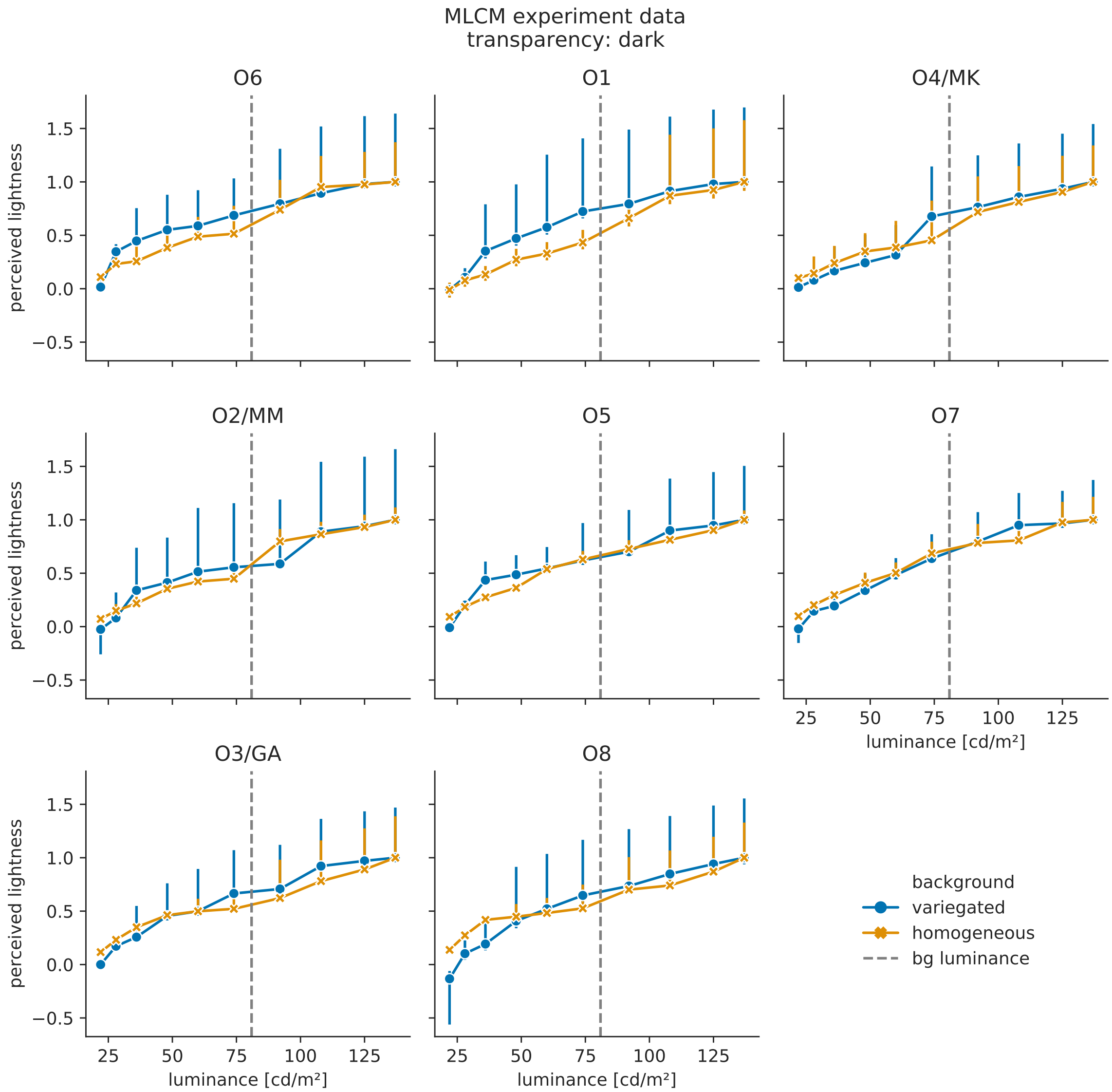

Figure 9: Perceptual scales measured with MLCM for all observers with a dark transparency. The perceptual scales are plotted against luminance. The orange crosses show the perceptual scale values for the experiment with a homogeneous background, while the blue circles display the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the perceptual scale values are connected with straight lines.

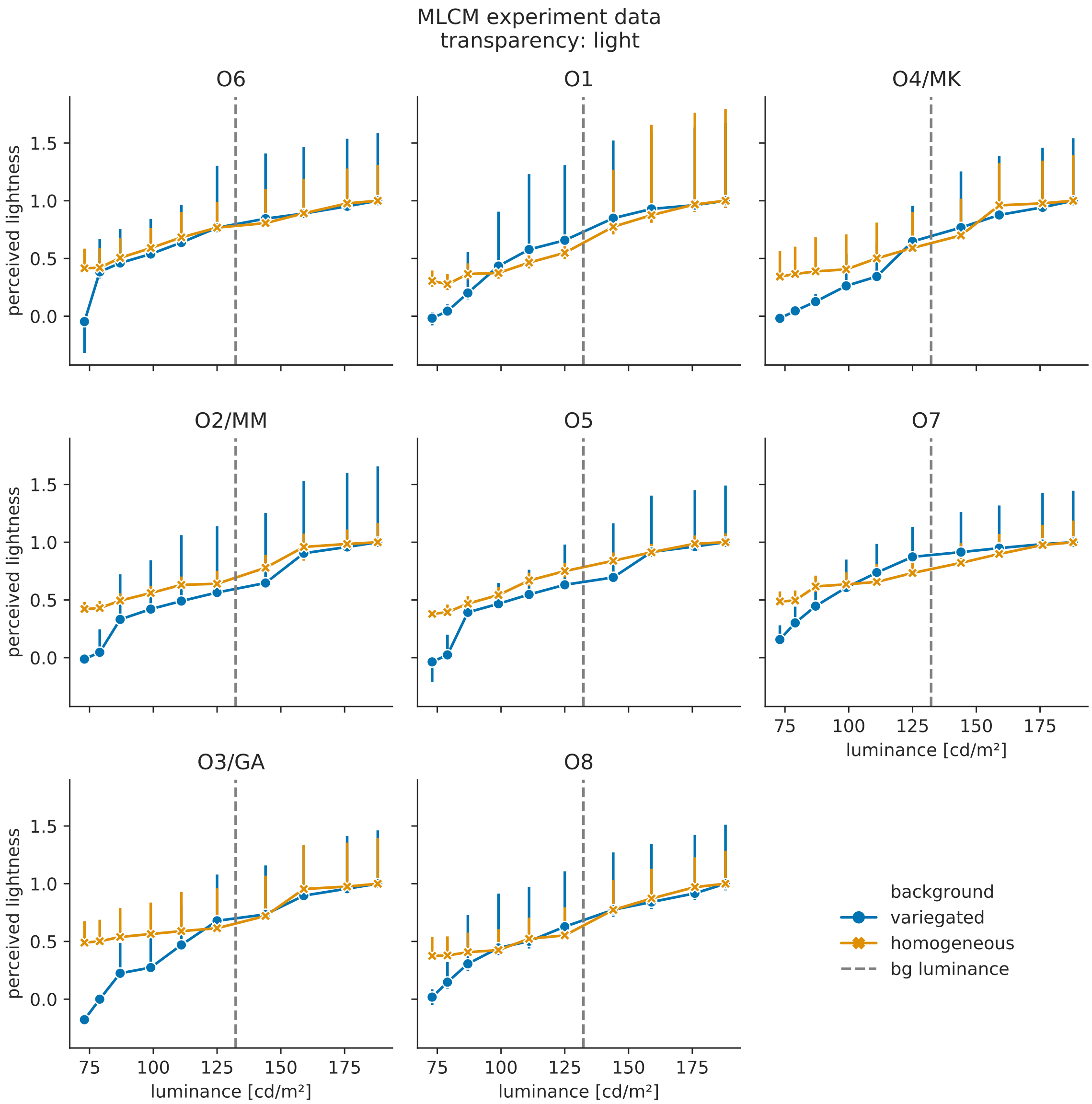

Figure 10: Perceptual scales measured with MLCM for all observers with a light transparency. The plot design and the legend are the same as in Figure 9.

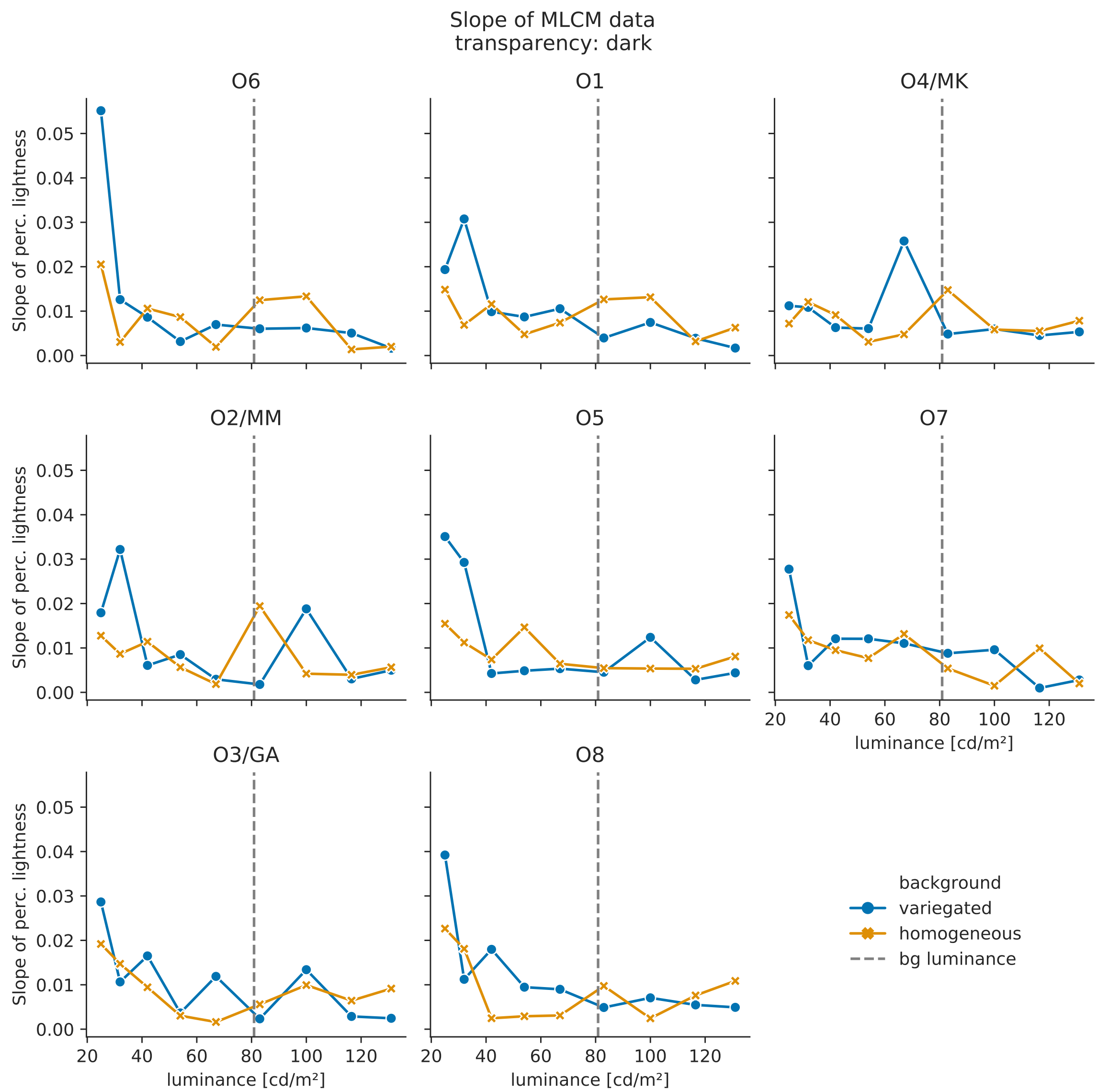

Figure 11: Slope of the perceptual scales measured with MLDS for all observers with a dark transparency. The orange crosses show the slope of the data for the experiment with a homogeneous background and the blue circles the slope of the data for the experiment with a variegated background. In order to facilitate the comparison of the two experimental conditions the points are connected with straight lines.

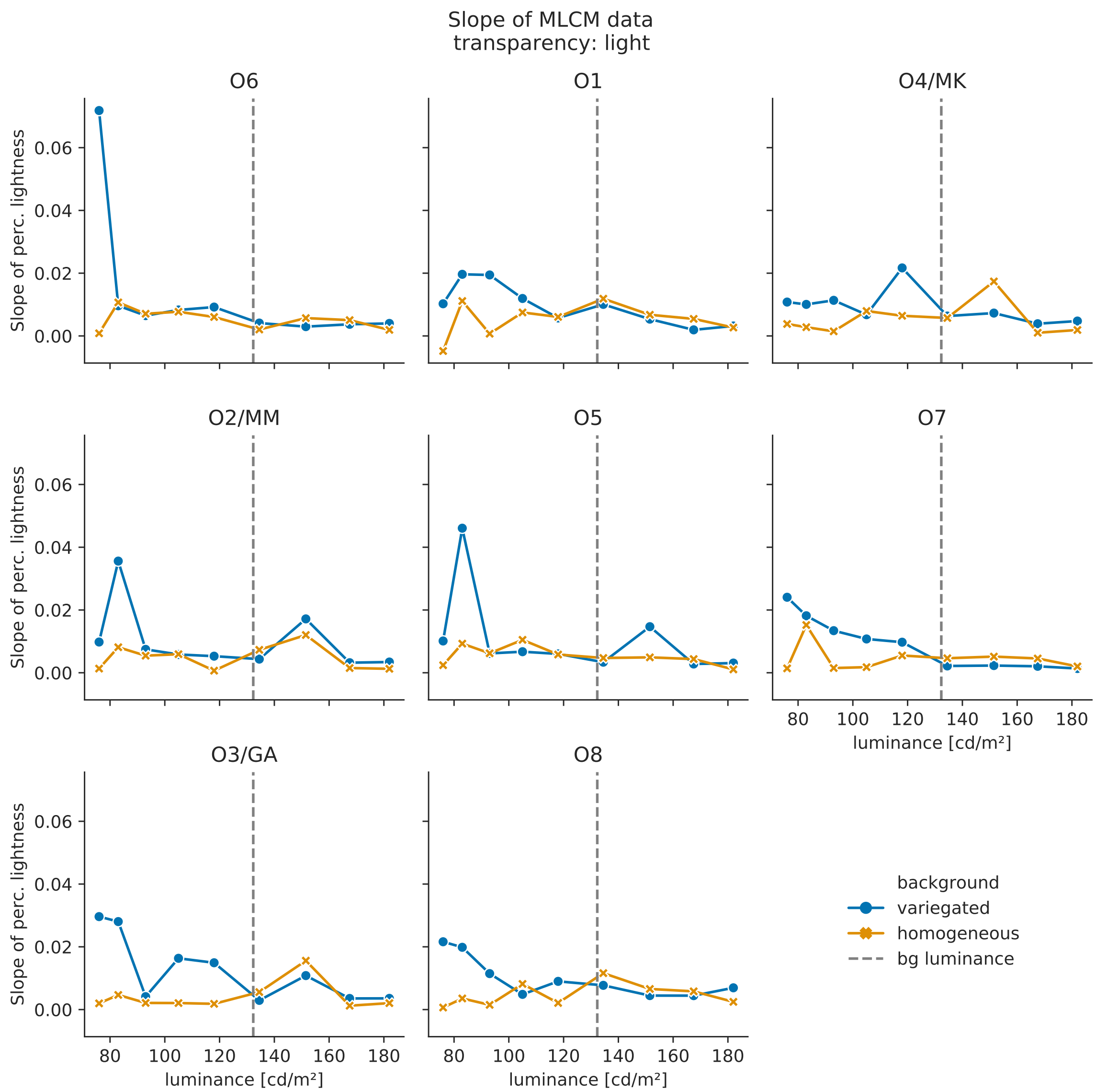

Figure 12: Perceptual scales measured with MLCM for all observers with a light transparency. The plot design and the legend are the same as in Figure 9.