Technische Universität Berlin

Bachelor thesis:

Relating discrimination with perceived magnitude on simultaneous brightness contrast displays

Jannis Ole Röhl

Matriculation Number: 469035

Computer Science (B.Sc.)

12. December 2025

First examiner: **Dr. Guillermo Aguilar**

Second examiner: **Prof. Dr. Guillermo Gallego**

Eigenständigkeitserklärung - Affidavit

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe. Alle Stellen die den benutzten Quellen und Hilfsmitteln wörtlich oder inhaltlich entnommen sind, habe ich als solche kenntlich gemacht.

Sofern KI-Tools verwendet wurden, habe ich Produktnamen, Hersteller, die jeweils verwendete Softwareversion und die Art der Nutzung (z.B. sprachliche Überprüfung und Verbesserung der Texte, systematische Recherche) benannt. Ich verantworte die Auswahl, die Übernahme und sämtliche Ergebnisse des von mir verwendeten KI-generierten Outputs vollumfänglich selbst.

Die Satzung zur Sicherung guter wissenschaftlicher Praxis an der TU Berlin vom 15. Februar 2023. (Link zur Satzung vom 15. Febuar 2023) habe ich zur Kenntnis genommen.

Ich erkläre weiterhin, dass ich die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt habe.

Berlin, den 12.12.2025

J. Röhl

....

Unterschrift

Zusammenfassung

In dieser Studie wurde untersucht, wie der simultane Helligkeitskontrast die Beziehung zwischen wahrgenommener Helligkeit und Empfindlichkeit im Bezug auf Helligket beeinflusst. In Anlehnung an das Grundkonzept von Shi and Eskew (2024) habe ich die Methodik modifiziert und um den simultanen Helligkeitskontrast erweitert. Es wurden zwei Experimente mit mehreren Beobachtern durchgeführt. Im ersten Experiment wurden Maximum-Likelihood-Conjoint-Measurements (MLCM) verwendet, um Wahrnehmungsskalen für Kontrastinkremente und -dekremente für jeden Teilnehmer individuell zu ermitteln. Im zweiten Experiment führten die Beobachter eine Zwei-Alternativen-Wahl (2AFC) zur Unterscheidung von Sockeln und Tests bei verschiedenen Kontraststufen durch, wobei alle Kombinationen von Kontrastinkrement- und -dekrement, Sockeln und -Tests verwendet wurden. Anschließend wurde eine modifizierte Naka-Rushton-Gleichung an die Kontrastdekremente und eine kubische Funktion an die Kontrastinkremente angepasst. Schließlich wurde das Vorhersagemodell von Shi and Eskew (2024) verwendet, um die Wahrnehmungs-skalen auf die Vorhersage der Schwellwerte zu testen. Die Ergebnisse zeigen, dass die Wahrnehmungsskalen die Schwellen gut vorhersagen trotz der Anwesenheit des simultanen Helligkeitskontrastes. Die Vorhersagegenauigkeit variierte jedoch zwischen den Beobachtern und war weniger konsistent als in Shi and Eskew (2024), was darauf hindeutet, dass der simultane Helligkeitskontrast einen unterschiedlichen einfluss auf unterschiedliche Beobachter hat.

Abstract

The current study investigated how the simultaneous brightness contrast influences the relationship between perceived intensity and sensitivity. Following the general idea of Shi and Eskew (2024), I modified the methodology to incorporate the simultaneous brightness contrast. Two experiments were conducted with multiple observers. In the first experiment, maximum likelihood conjoint measurement (MLCM) was used to estimate perceptual scales for contrast increments and decrements across observers. In the second experiment, observers performed two-alternative forced-choice (2AFC) pedestal discrimination at different contrast levels, with all combinations of contrast increment and decrement pedestals and tests. Then, a modified Naka-Rushton equation was fitted to contrast decrements and a cubic function to contrast increments. Finally, Shi and Eskew (2024) prediction model was used to test the perceptual scales on predicting the pedestal discrimination thresholds. The findings demonstrate that the estimated perceptual scales offer a good prediction of discrimination thresholds, even in the presence of simultaneous brightness contrast. However, prediction accuracy in my study varied across observers and was less consistent than in Shi and Eskew (2024). Suggesting that simultaneous brightness contrast affects the relation between perceived intensity and sensitivity differently for observers.

Generative AI disclosure

Generative AI was used in the writing process of this thesis. I used Anthropic (2025) to grammatically correct my writing and gain inspiration for improving my style. The information and content were always written by me, and the tool was only used for gaining inspiration for rephrasing my sentences. The resulting answers from the model were always evaluated in terms of grammatical and content correctness. The generative AI model was used for chapter 4, 5, 7 and the abstract. If further proof is needed, I can submit the whole prompt journal.

Contents

1 Introduction

In our daily lives, we encounter different stimuli that we can not only perceive but also distinguish. For example, we can perceive something as visually bright, but we can also decide between two visually bright stimuli which one appears as more intense. Two underlying concepts to describe this perception of stimuli are sensitivity and perceptual intensity.

Sensitivity is referred to as the minimal difference between stimuli that is just detectable (Lu and Dosher, 2013) and is also often described as the just noticeable difference (JND). To find the JND between two stimuli, threshold measurements are used. In these measurements, a reference stimulus is compared to a test stimulus. A test stimulus is the sum of the reference stimulus and an offset for the measured physical variable. It is the observer's task to choose the test stimulus and not the reference stimulus (An experiment, where neither the reference nor the test stimulus is the null stimulus, is generally referred to as a discrimination experiment (Kingdom and Prins, 2016)). These experiments are always performance based, therefore, there is a clear right and wrong answer.

Perceptual intensity is the subjective strength of a stimulus in relation to its physical manipulation and can be quantified through scaling (Kingdom and Prins, 2016). Scaling, or perceptual scaling, refers to a collection of psychophysical methods for collecting data that describes the relationship between a stimulus and human perception. One example of this is the maximum likelihood difference scaling (MLDS) model, developed by Maloney and Yang (2003), which utilizes a statistical model to estimate perceptual scale values.

Unlike the estimation of sensitivity, which is based on thresholds, the estimation of perceptual intensity uses suprathreshold comparisons. The latter can be defined as an experiment involving stimuli that are each above there own detection threshold (Kingdom and Prins, 2016). With stimuli that are definitely distinguishable, the experiments are appearance based. In other words, they focus on the apparent magnitude, not on discrimination, as is the case with performance based experiments.

The psychophysical concepts of sensitivity and perceptual intensity are the key to investigate other related phenomena. One example of their involvement is the qualitative difference between increments and decrements in contrasts of stimuli by human achromatic vision. The observer's perception of increases in the stimulus intensity are called increments, while the perception of decreases in intensity are called decrements. This qualitative difference between increments and decrements is shown in many studies related to human visual perception (Chichilnisky and Wandell, 1996; Rudd and Zemach, 2004) .

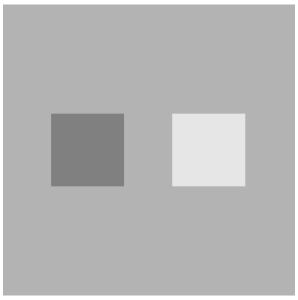


Figure 1: A gray background with two squares. The left square is an decrement and the right square is an increment in relation to the background.

1.1 Previous work by Shi & Eskew

Shi and Eskew (2024) investigated the qualitative difference between increments and decrements. Firstly, they separated the two variations of the stimuli and estimated a perceptual scale for each of them. Secondly, they tested predictions of pedestal discrimination on the two perceptual scales.

In this context, pedestal discrimination refers to the distinction between a pedestal stimulus and a test stimulus, where the pedestal stimulus is not null and the test stimulus is the combination of the pedestal and an added or subtracted offset (Figure 2).

To estimate the perceptual scales, Shi and Eskew (2024) used the MLDS model, with collected data from an experimental approach. The stimuli for the experiments contain a mid-gray background, a little fixation cross and one square in each quadrant of the fixation

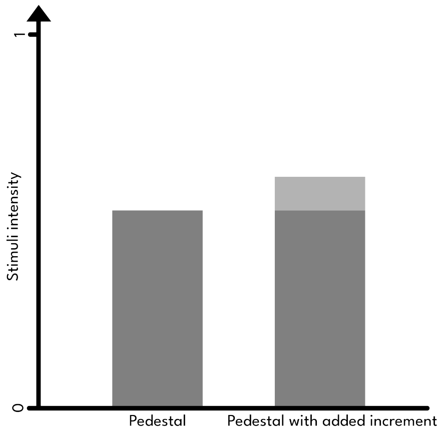


Figure 2: Conceptualization of pedestal discrimination stimuli. The left bar represents the pedestal and the right bar the test contrast which is in this case the pedestal combined with an increment.

cross. The four squares are either increments or decrements and definitely distinguishable from one another (Figure 3 (a) shows decrements and (b) increments). Participants were asked to compare the top two squares with the bottom two squares with each other and had to choose the pair that appeared more similar in contrast. They estimated two perceptual scales from the collected answers. One for the increments and one for the decrements.

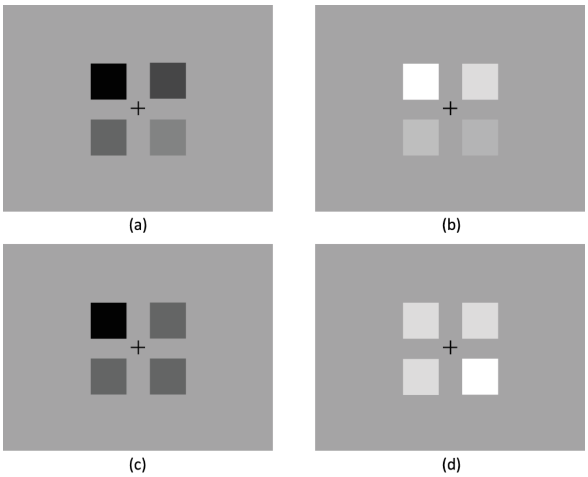


Figure 3: Stimuli examples from Shi and Eskew (2024). (a) is an incremental and (b) an decremental perceptual scale example. (c) is an increment pedestal with a decrement test and (d) is an increment pedestal with increment test example of pedestal discrimination stimulus.

Shi and Eskew (2024) measured the sensitivity with pedestal threshold measurements using a similar stimulus as in the perceptual scales experiment (the stimulus contains a mid-gray background with a fixation cross and four squares, each in one of the quadrants of the fixation cross). All four squares are either increments or decrements and three of the four squares have the same contrast (pedestal). The last square is the pedestal combined with a test contrast that is either incremental or decremental (Figure 3 (c) shows a decrement pedestal with decrement test and (d) a increment pedestal with increment test). The observers were asked to to choose the side with the test contrast.

To test the perceptual scales, Shi and Eskew (2024) assumed that the thresholds at a pedestal level are inversely proportional to the rate of change of the perceptual scale at the pedestal contrast level (Shooner and Mullen, 2022). To test this assumption, Shi and Eskew (2024) created a prediction model on this assumption. Therefore, they evaluated the goodness of the fit by calculating the square mean of the correlation between the predicted and obtained thresholds across pedestal contrasts, as well as the root-mean-square error (RMSE).

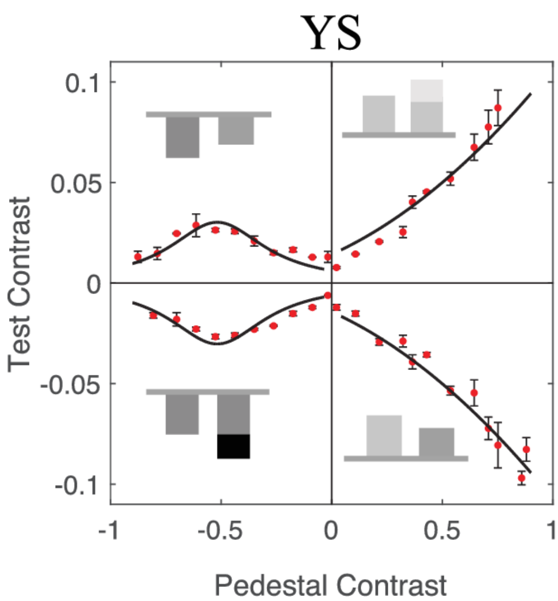


Figure 4: One result from Shi and Eskew (2024). The red dots represent the mean pedestal discrimination thresholds. The black graphs are the predicted thresholds from the estimated perceptual scales. The correlation between the estimated and measured values is clearly visible as the overlay between them.

Shi and Eskew (2024) discovered two major findings. First, both experiments revealed the replicate of the asymmetry between increments and decrements. Second, and more importantly, they discovered the strong predictive accuracy of contrast discrimination from the perceptual scales. This suggests that the thresholds are inversely related to the derivative of the perceptual scale curve. (Figure 4).

1.2 Simultaneous brightness contrast

The human perception of the qualitative difference between increments and decrements is not constant and varies depending on the circumstances. There are certain stimuli that alter the perceptual intensity, for instance optical illusions. The simultaneous brightness contrast is one example of optical illusions, which affects the magnitude of the human perception.

It contains two gray squares with the same stimulus intensity, one placed on a white and one placed on a black background (Figure 5). Through the underlying effect of lateral inhibition, the gray square with the white background appears less bright than the one with the black.

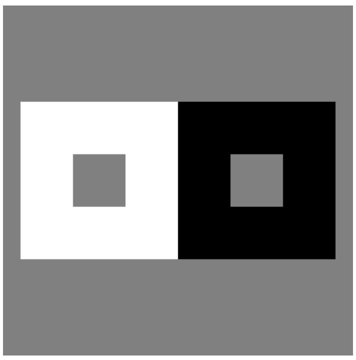


Figure 5: The Simultaneous brightness contrast with physical equal bright gray patches on both sides.

1.3 Research question

The simultaneous brightness contrast has an influence on the perceptual intensity of the gray squares contained in its stimuli. Combined with the results by Shi and Eskew (2024), this leads to the following research question: How is the relation between sensitivity and perceived intensity influenced by the simultaneous brightness contrast?

In order to successfully investigate the relation between perceptual intensity and sensitivity of brightness in the context of the simultaneous brightness contrast, I created a two-part experiment with four participants combining the main approach from Shi and Eskew (2024) with the simultaneous brightness contrast. First, I estimated incremental and decremental perceptual scales for each observer in the simultaneous brightness contrast. Following this, I tested the prediction of pedestal discrimination against measured thresholds.

2 Methods

The experiment was conducted in two parts. The first part involved the estimation of perceptual scales using MLCM (Ho et al., 2008, Knoblauch and Maloney, 2012). The second part involved the measurement of thresholds using the same stimuli intensities as in the first part of the experiment.

2.1 Observers

Four observers participated in the experiment. Two of them were naive (DD and LW) and the other two were informed (JR the author and JV an affiliate). All participants had normal or corrected to normal vision. The naive participants received a compensation of 12,50 € per hour and completed the experiment within four hours.

2.2 Apparatus

The stimuli for this experiment were presented on a monitor in a darkened room. I used a JVC monitor (376 mm x 301 mm, 1024 x 768 px) operating at 60 Hz with a 16 bit grayscale range (65536 levels). The monitor had a luminance range from 0.62 cd/m^2 to 507 cd/m^2. The participants sat 76 cm away from the monitor, which resulted in approximately 32 pixels per viewing degree. The presentation of the stimuli was handled via a python library HRL (https://github.com/computational-psychology/hrl).The participants were given a chinrest, to assure a stable viewing position, and a keypad, to submit their answers.

2.3 Stimuli

For both experiments, the stimuli were created with stimupy (Schmittwilken et al. (2023)) using the same spatial layout. The spatial layout contained two squares with a diagonal of 1.5°. Each square was placed into the center of a 5 x 7° surrounding area, with the two areas placed next to each other. A gray background of 152 cd/m^2 filled the rest of the monitor. The surrounding areas were either black (0.62 cd/m^2) or white (507 cd/m^2). Each stimulus was presented until a decision was made by the observer and between two stimuli a small

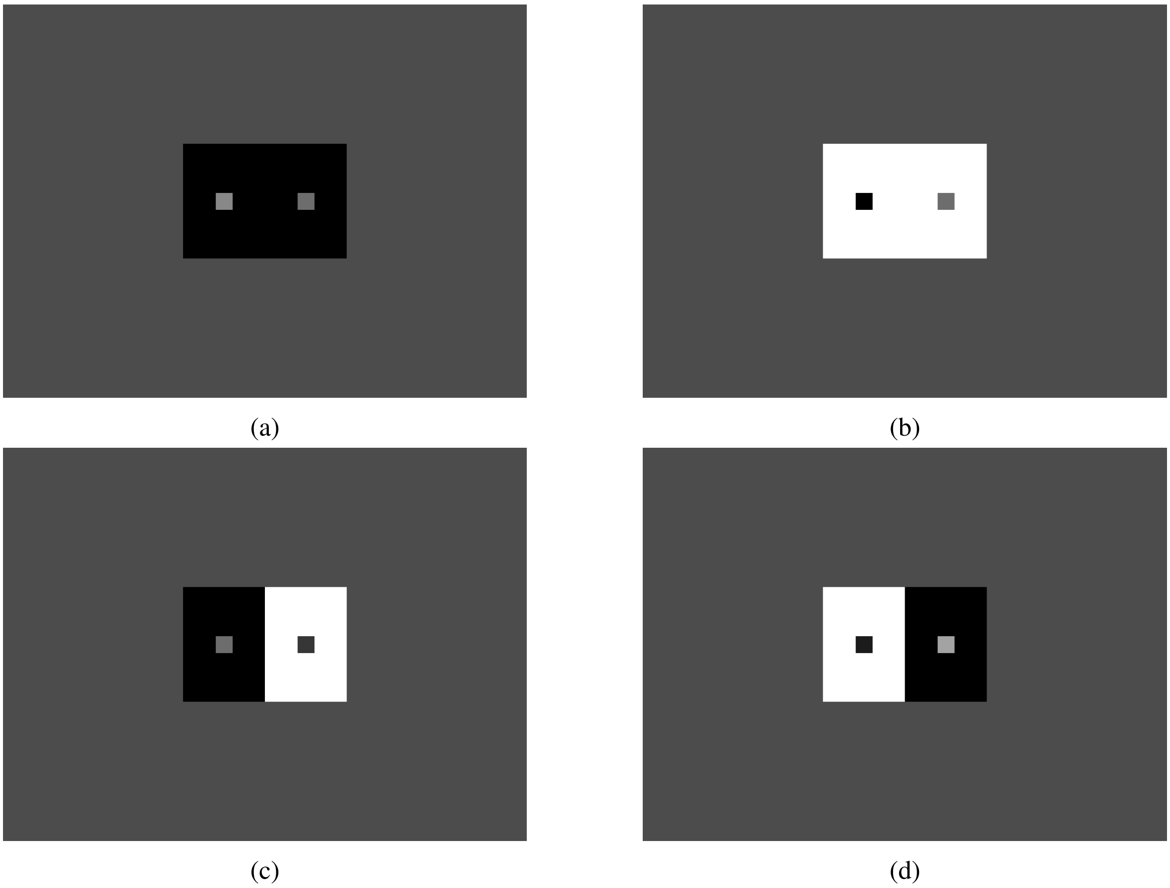


Figure 6: Example stimuli from the experiments. Figure 6a shows two gray squares on a black surrounding and is a increment test. Figure 6b shows a decrement test. Figure 6c and Figure 6c are example stimuli used for the MLCM experiment and contain a mixed surrounding (one gray square paired with a black surrounding and the other square paired with a white surrounding)

fixation cross was displayed for (250 ms). The luminance of the gray squares varied for the experiments and is described further in the corresponding sections.

The experiment made use of the weber contrast (Westheimer, 1985). The weber contrast is used to calculate the difference between the gray square and the corresponding surrounding areas of the stimulus. The contrast can be defined as the difference between the luminance of the gray square and the surrounding area, divided by the luminance of the surrounding area ($\Delta l/l$).

2.4 Perceptual scale estimation with MLCM

To measure perceptual scales, the MLCM (Ho et al., 2008, Knoblauch and Maloney, 2012) model was preferred over the MLDS (Maloney and Yang, 2003) model, as MLCM can yield perceptual scales from different judgments at the same time. As a result, we are able to estimate a perceptual scale with consideration to the context (black or white surrounding).

The approach used to estimate the perceptual scales can be compared to Vincent et al. (2024). At first, observers were given two minutes to adopt to the background luminance of 152 cd/m^2. They were then presented with different stimuli (layout in 2.3, examples in Figure 3) and had to report whether the left or right gray square appeared brighter. The participants reported their decision via a keypad.

Eight luminance intensities were sampled linearly between 0.79 cd/m^2 and 380 cd/m^2 and used for the stimuli. Each of the eight intensities was paired with a black (0.62 cd/m^2) or white (507 cd/m^2) surrounding. This resulted in 16 variations (8 luminances x 2 surroundings). All variations were paired with one another and judged by the observers. This resulted in 120 judgments, also named trials. The 120 trials were combined in a random order in a block. Observers were tasked to repeat five blocks for this experiment.

2.5 Threshold measurements with pedestal discrimination

The method of constant stimuli, or method of constants (Kingdom and Prins (2016) was used to estimate thresholds. Observers were presented with different stimuli and had to judge whether they perceived the left or right gray square as brighter. They used a keypad to report their decision.

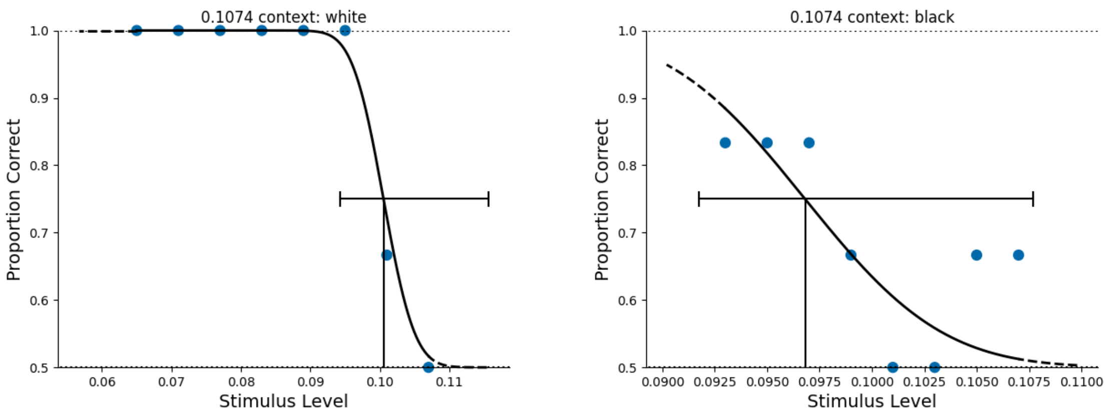
Thresholds were estimated for the 16 variations from 2.4. The luminance of the gray squares served as pedestals in this experiment. Incremental and decremental thresholds were estimated for each variation, with exception for the lowest stimulus intensity of the gray square (only incremental) and the highest stimulus intensity of the gray square (only decremental).

For a threshold measurement, eight test contrast were linear selected between a minimum and maximum offset from a certain pedestal. Each of the eight test contrasts was presented with the pedestal on the same surrounding. The observes judged if they perceived the left or right one as brighter and used a keypad to report there answer. Each comparison was repeated six times which results in 48 trials (eight test contrasts x six repeats) per block. This process was repeated for every threshold estimated on a black surrounding (incremental and decremental, blocks in a random order) and then for every threshold on a white surrounding (incremental and decremental, blocks in a random order). This resulted in 28 blocks (16 variations, 12 with two threshold estimations, four with only one).

After a first run of all 28 blocks (14 increments and 14 decrements) the ranges for the test contrasts were adjusted to a more suitable range for the observer. Then all 28 blocks were repeated again.

All four observers completed this process separately.

The strategy behind the adjustment was to maximize the data points that are between 50 % to 100 % correct distinction. To achieve this goal I raised the maximum difference between the pedestal and test contrasts, given the highest values were below 100 % correct



(a) (b)

Figure 7: Two examples for test contrast range adjustments. Figure 7a shows a result from the first run of LW. The tested range resulted in six out of eight test contrasts where distinguished. Therefor were the maximum difference for the test contrast lowered. Figure 7b shows also a result from the first block of LW. A correct distinction wasn't reached. Therefore, the maximum difference for the test contrasts was increased between the two runs.

distinction. I lowered the maximum difference, if the three most different test contrasts were all completely distinguishable by the observer. When the lowest two to three differences were all higher than 50 % correct distinction I lowered the minimum difference for the pedestal.

All the settings for the pedestal discrimination can be found in the appendix (Appendix 7.3).

2.6 Perceptual scale analysis

The analysis was conducted for each observer individually. The five data-sets from each observer were combined. Afterwards, the perceptual intensities were calculated from the data-set. In order to calculate the perceptual intensities, the MLCM implementation in R (R Core Team, 2024, Knoblauch and Maloney, 2012, Ho et al., 2008) was used. The implementation has different models for the estimation. The additive model and the full model were both estimated. Then, the better fitting model was chosen and used for the estimation.

The perceptual intensities were split for both surroundings. I min-max-normalized the two data-sets individualy and used the values to fit the same functions used in Shi and Eskew (2024). Therefore, I fitted a modified Naka–Rushton (Michealis–Menton) equation (1a) for the increments and a (1b) for the decrements.

$$P_+ = \left[1 + \frac{m}{C_{max} - C_{min}}\right] \times \frac{(C_+ - C_{min})}{(C_+ - C_{min}) + m} \tag{1a}$$

$$P_- = b \times C_-^3 + d \times C_-^2 + e \times C_- + f \tag{1b}$$

The equation (1a) was used to fit the incremental perceptual scale (P_+). The function has only one free variable m. C_+ is the contrast from the data-set, C_{min} is the minimum contrast of all tested increment stimuli and C_{max} is the maximum contrast from all increment stimuli. The decremental perceptual scale (P_{-}) was fitted with (1b). C_{-} refers to the decrement contrasts used in MLCM and all other variables are free.

I fitted the decremental perceptual scale with boundaries for the four free variables. The ranges for the bounds were estimated from the fit results from Shi and Eskew (2024) and were as follows:

	b	d	e	f
Lower bound	1	$-\inf$	1	$-\inf$
Upper bound	inf	-1	inf	0

Table 1: The bounds for the MLCM fit, estimated from the results from Shi and Eskew (2024).

2.7 Pedestal discrimination analyses

The pedestal discrimination analysis was conducted for every observer individually. For every pedestal two data-sets were collected. These two data-sets were combined and pooled using the pooling utility from psignifit (Zito et al., 2025, Schütt et al., 2016). Then, the same module was used to estimate a psychometric function for the pooled data from a pedestal. The threshold was taken from 75 % right distinction from the estimated psychometric function.

This process was repeated for each observer, for each pedestal and each threshold direction (increment or decrement) individually.

2.8 Combined analyses with prediction model

The combined analyses were done separately for every participant and followed the concept by Shi and Eskew (2024) (Figure 8 shows the concept visualized).

The first half of the analysis consisted of the plotting of the test contrasts. The test contrasts are the differences between the pedestal contrasts and the contrasts for the estimated

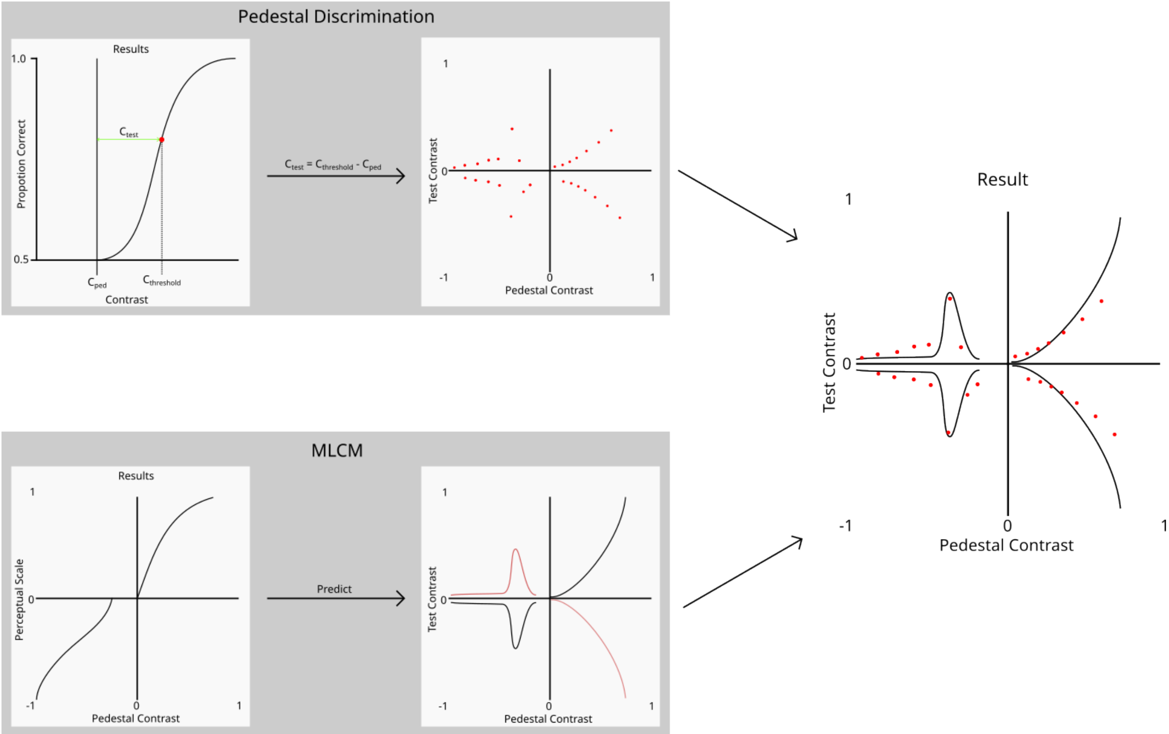


Figure 8: The concept of the combined evaluation. The top row shows the process for the pedestal discrimination results. The bottom row shows the process for the MLCM results. The right side displays the combined results.

threshold. This was done for every estimated threshold for each observer (Figure 8 upper left part).

Afterwards, I used the prediction model from Shi and Eskew (2024) to plot the prediction of the test contrasts in relation to the pedestal contrast into the same coordinate system. The prediction model ((2)) has one free variable and was fitted through the measured threshold contrasts. The incremental prediction was fitted with the increment test contrasts and the decremental prediction with the decrement test contrasts. The negative values were multiplied by -1 for the fit as in Shi and Eskew (2024).

$$C_t = k_{A_\pm} \times \sigma_{A_\pm} \times \frac{1}{\frac{d}{dc} p(C_{ped})} \tag{2}$$

C_{ped} is the contrast of the pedestal, C_t is the threshold contrast, the denominator is the derivation from the estimated perceptual scale at the pedestal contrast level C_{ped}. $\sigma_{A_\pm}$ is the standard deviation from the noise from the MLCM model and $k_{A_\pm}$ is the only free variable that is used for fitting the prediction scale. $k_{A_\pm}$ and $\sigma_{A_\pm}$ can be different for the incremental and decremental perceptual scale. To compare the goodness of fit with the former work by Shi and Eskew (2024) I also calculated the square of the correlation between the predicted and obtained mean thresholds across pedestal contrasts and Root-Mean-Square-Error (RMSE) between the measured test contrasts and the prediction from the perceptual scales.

3 Results

3.1 MLCM results

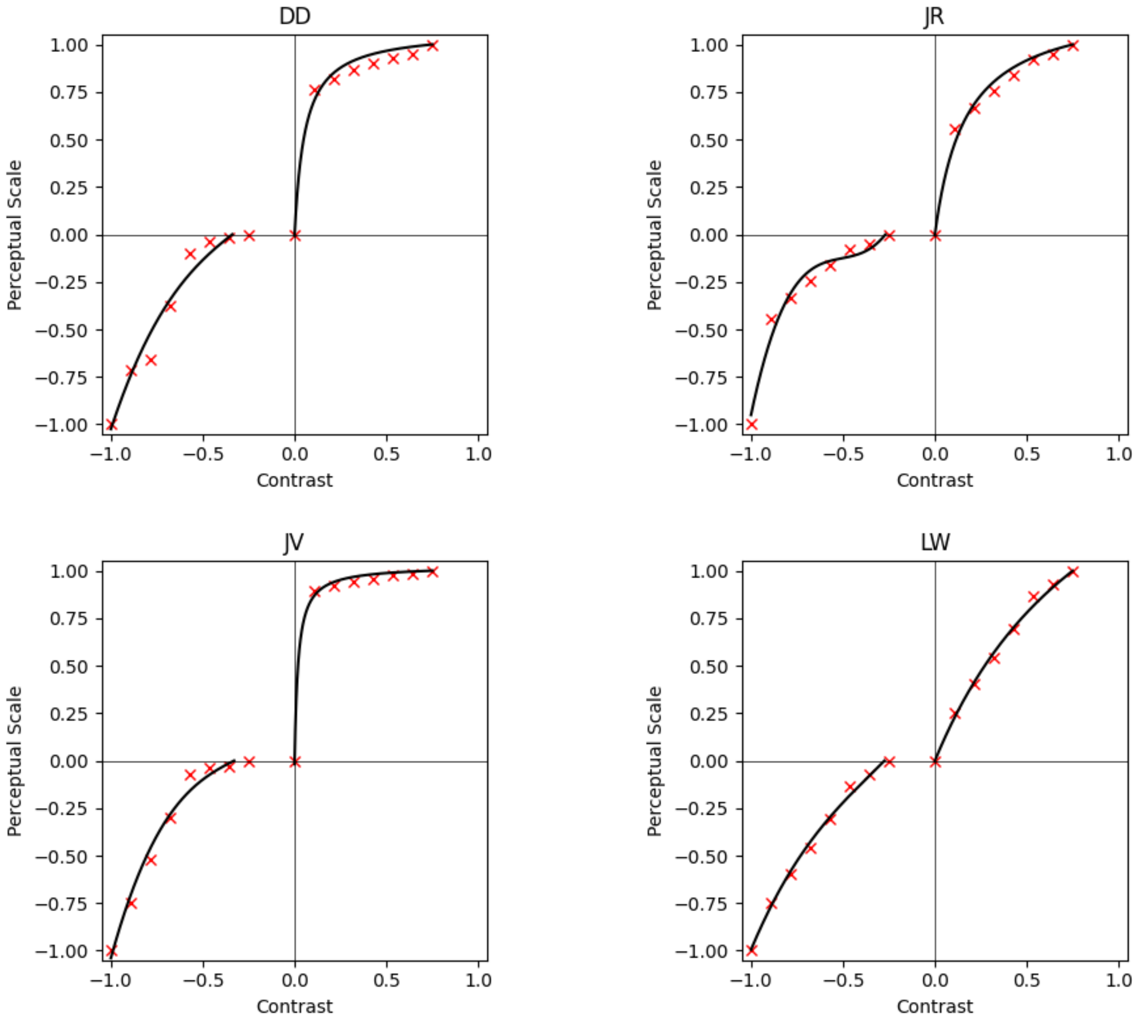


Figure 9: The perceptual scales estimated by MLCM for all four observers. The horizontal axis is the contrast, and the vertical axis is the perceptual scale. The red crosses are the estimated values from the MLCM model individual for each observer. The black curves are the best fits for equation (1a) and equation (1b) for each observer. The perceptual scale for the increments is plotted in the first quadrant, the perceptual scale for the decrements is plotted in the third quadrant.

The fitted results from the MLCM estimation are shown in Figure 9. Every plot reveals the results from one observer. The horizontal axis presents the contrast of the stimuli. In the positive direction, increments occure, and in the negative direction, decrements occure.

The vertical axis displays the perceptual scale. The red crosses are the perceptual intensities estimated by MLCM and the black curves are the fitted curves from 2.6.

The first thing to notice, are the different shapes between perceptual scales from the observers. All of them look visually different but share some key similarities as well.

One similarity is the concave shape of the incremental function. This shape is prominently shown by the observers DD and JV, and JR. LW also shows a concave curve for the increments. However, this shape tends more to a linear function than the other three observers.

The perceptual scales for the decrements also share similarities between the four observers. All four scales show a concave shape. The fitted scale for JR is the only one that shows the prominent reverse S-shape as the perceptual scales in Shi and Eskew (2024). The other three observers share a more linear shape than the scale from JR.

The additive MLCM model was used for JR and LW and the full MLCM model was used for DD and JV.

3.2 Pedestal discrimination results

In the second experiment, I investigated pedestal discrimination on different thresholds, test direction and surroundings (black or white). The complete results are printed in the appendix at 7.1. The estimated offset between the threshold and pedestal is plotted in Figure 11. The increment thresholds increased as the contrast increased for all observers. The thresholds for the decrements are different between the four observers. For LW the thresholds stayed nearly constant. For DD, JR and JV, the thresholds stayed more identical until the contrast reaches -0.5. The next thresholds after this points were lower. As the contrast was lowered further, the thresholds start to increase.

In (Appendix, 7.1) are the psychometric functions for all estimated thresholds for one context plotted. A example for only one pedestal is seen in Figure (10). For all four observers it is shown that the difference between the pedestal and estimated threshold is higher than the difference between the pedestal and the next pedestal with higher contrast. This is

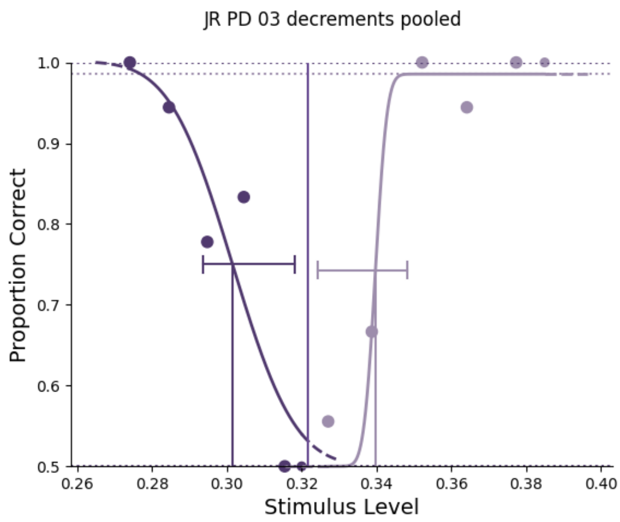


Figure 10: Example result from pedestal discrimination for observer JR. On the horizontal axis is the contrast and on the vertical axis the proportion correct. The vertical line in the middle is the pedestal used in the threshold estimations. To both sides was a psychometric function fitted, and thresholds estimated.

most prominent on the black surrounding and indicates an unclear distinction between the thresholds.

3.3 Prediction results

The model predictions are plotted over the pedestal discrimination in 11. In all cases did the prediction model fit the data well. This can be seen in the visual similarity between the data-points and the plotted curves. However, the difference between the data-points and curves vary between observers. These differences can be found again in the different goodness of fit. The goodness of fit (R^2 and RMSE) is acceptable for all observers and better for A_+ then $A-$. A_+ R^2. One exception is the R^2 $A-$ score from JR. The negative score hints to a insufficient fit. The k_{A_+} value is for JR and LW higher as the k_{A_-} value, for DD and JV it is the other way around. This difference can come from the chosen model from MLCM (JR and LW used the additive model, DD and JV used the full model)

Condition	Observers DD	JR	JV	LW
A+ perceptual scale (1a)				
m_+	0.056	0.167	0.019	0.976
A+ prediction (2)				
k_{A_+}	0.393	0.894	0.271	0.540
σ_{A_+}	0.037	0.046	0.029	0.160
A- perceptual scale (1b)				
b	1.348	6.079	2.035	1.009
d	-1.049	-9.214	-1.783	-1.000
e	1.000	4.844	1.000	1.285
f	-0.271	-0.754	-0.208	-0.297
A- prediction (2)				
k_{A_-}	0.672	0.137	0.697	0.152
σ_{A_-}	0.046	0.046	0.029	0.160

Table 2: The fitting results for the perceptual scales and the corresponding prediction models.

Observer	DD	JR	JV	LW
A_+ R^2	0.691	0.692	0.687	0.419
A_- R^2	0.264	-0.600	0.414	0.241
A_+ RMSE	0.039	0.038	0.045	0.052
A_- RMSE	0.009	0.020	0.014	0.014

Table 3: The goodness of fit estimation for all four observers and both perceptual scales.

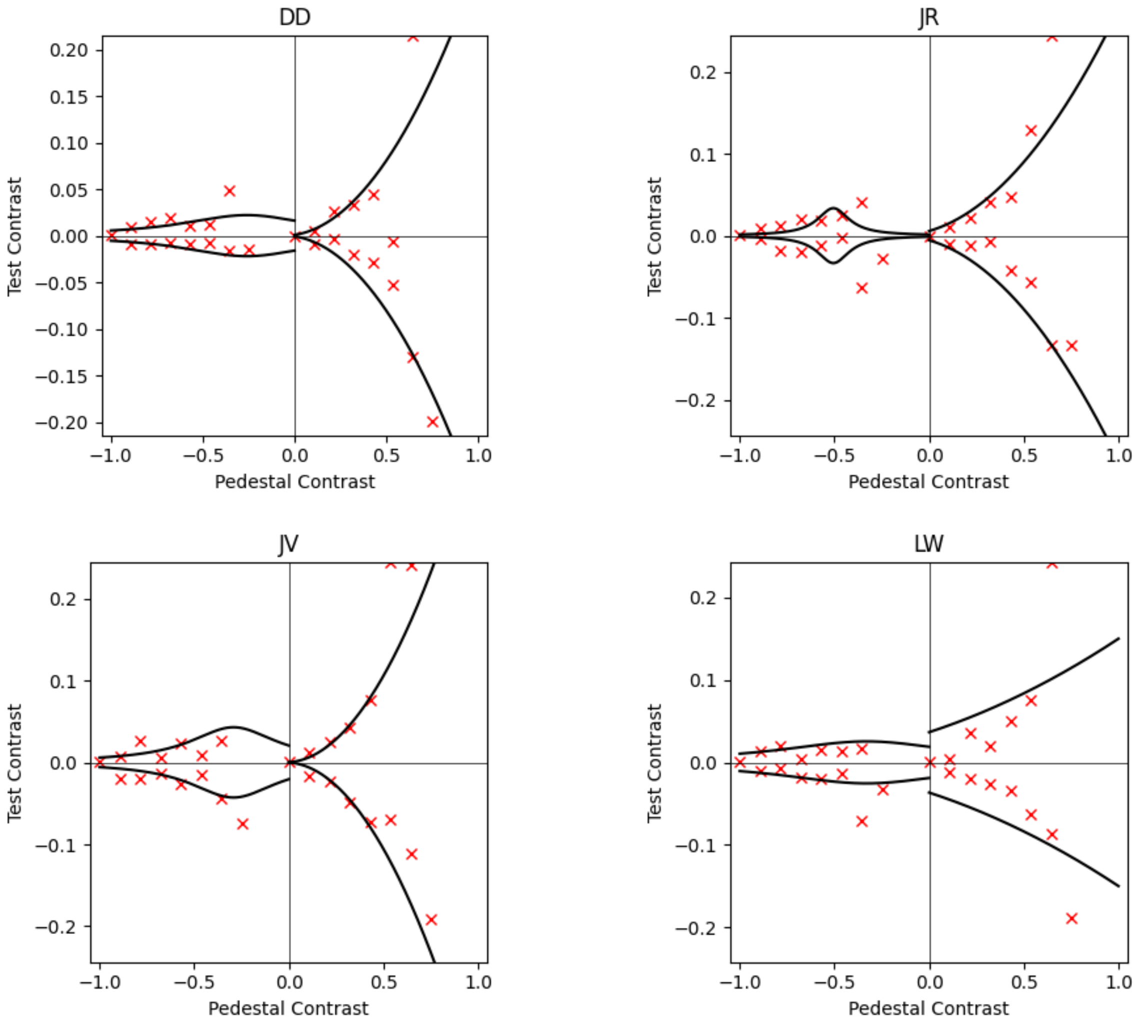


Figure 11: The combined end results for all four observers. The horizontal axis is the pedestal contrast and the vertical axis is the test contrast (difference between estimated threshold contrast and pedestal contrast). The red crosses are threshold contrast results from the pedestal discrimination experiment. The black curves are the estimated prediction from the MLCM experiment.

4 Discussion

The influence of simultaneous brightness contrast on the relationship between sensitivity and perceived intensity varies across observers. The estimated perceptual scales predicted the thresholds well, even in the presence of the simultaneous brightness contrast. However, the prediction varied across observers, with some showing a better predictability then others.

Shi and Eskew (2024) demonstrated a relationship between the perceived intensity and sensitivity. While the main findings from there study align with those in this study, they differ in details. The difference between the measured thresholds and predictions was visibly smaller in Shi and Eskew (2024), and the goodness of fit (R^2 and RMSE) was better for all observers. Additionally, Shi and Eskew (2024) results showed greater consistency between all observers, with better and more constant fits. In contrast, my study shows more variability between the four participants.

There are several factors that may account for these differences between the studies. One explanation is the influence of the simultaneous brightness contrast. This optical illusion is known to affect observers differently. This may explain the difference between predicted and measured thresholds in this study, and consequently, the difference between the present findings and the former findings from Shi and Eskew (2024).

4.1 Limitations

This study has several technical and design limitations that influenced the results. First, the linear spacing of the pedestals lead to a to narrow spacing for incremental high contrast stimuli. This is hinted by the thresholds in (Appendix 7.1) which are estimated after the next pedestal. This poor discriminability may have affected perceptual scale estimation with MLCM and, consequently, the accuracy of threshold predictions derived from these scales.

The decision to include only eight pedestals was made to keep the overall time needed for the experiment manageable for each participant. Also, the choice to sample pedestals linearly resulted in fewer pedestals as used in Shi and Eskew (2024) as higher spacing was needed to discriminate incremental pedestals on black surrounding. This lower number of pedestals may have limited the capability to fit suitable perceptual scales due to insufficient data coverage. Future work should examine whether nonlinear spacing and additional pedestals improve the fit of perceptual curves.

To keep experimental duration manageable, only five repetitions of the 2.4 experiment were conducted. However, MLCM estimates become more stable with additional repetitions. With only five repetitions, this stability may not have been achieved, and the perceptual scale estimates could change with additional data collection.

4.2 Technical difficulties

Shi and Eskew (2024) constraint there perceptual scales by fixing the lowest pedestals (lowest luminance of the gray square) to 0 and the highest pedestals (highest luminance of the gray square) to 1. I used scipy (Virtanen et al., 2020) curve fit for the fitting in this study, which has no support to recreate the exact constraints Shi and Eskew (2024). As a workaround, I applied bounds in the fitting process (Table 1) which where estimated from the results from Shi and Eskew (2024). This difference in the fitting process may have affected the estimated perceptual scales and therefore the threshold predictions.

Additional technical difficulties arouse with the psignifit module, which contained an error in the negative-normal sigmoid function at the time of writing this paper. To work around this error, I inverted the datasets that needed this sigmoid and calculated the thresholds using the normal sigmoid instead. Afterwards, the results were reversed. This process wasn't optimal and may had an influence on the decremental thresholds for both surroundings.

4.3 Open questions

Previous studies have found that effect from the simultaneous brightness contrast appears more intense when presented briefly (Kaneko and Murakami, 2012, Kaneko et al., 2018). Future work could examine whether brief presentations shows a stronger effect on the relation between perceived intensity and sensitivity. However, this approach may complicate pedestal selection, as higher test contrasts would be required due to the stronger effect.

Future studies should also explore a different pedestal spacing scheme to provide more and better data points for the MLCM evaluation. Additionally, collecting more repetitions of both the perceptual scaling and pedestal discrimination tasks. This could clarify if more data results in a clearer influence from the simultaneous brightness contrast on the relation between perceived intensity and sensitivity.

5 Conclusion

This study investigated how the simultaneous brightness contrast influences the relationship between perceived intensity and sensitivity. Two experiments were conducted to estimated perceptual scales and discrimination thresholds for different observers. These perceptual scales were then tested for their ability to predict thresholds using the model from Shi and Eskew (2024). Results showed different prediction accuracy across observers, indicating that the simultaneous brightness contrast affects individuals differently. This varies from the more consistent results reported by Shi and Eskew (2024).

6 References

Anthropic. (2025). Claude 4.5 sonnet [large language model]. https://claude.ai/new

Chichilnisky, E. J., & Wandell, B. A. (1996). Seeing gray through the on and off pathways. *Visual Neuroscience*, *13*(3), 591–596. https://doi.org/10.1017/S0952523800008270

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture [PMID: 18271869]. *Psychological Science*, *19*(2), 196–204. https://doi.org/10.1111/j.1467-9280.2008.02067.x

Kaneko, S., & Murakami, I. (2012). Flashed stimulation produces strong simultaneous brightness and color contrast. *Journal of Vision*, *12*(12), 1–1. https://doi.org/10.1167/12.12.1

Kaneko, S., Murakami, I., Kuriki, I., & Peterzell, D. H. (2018). Individual variability in simultaneous contrast for color and brightness: Small sample factor analyses reveal separate induction processes for short and long flashes [PMID: 30263104]. *i-Perception*, *9*(5), 2041669518800507. https://doi.org/10.1177/2041669518800507

Kingdom, F. A., & Prins, N. (2016). Chapter 1 - introduction and aims. In F. A. Kingdom & N. Prins (Eds.), *Psychophysics (second edition)* (Second Edition, pp. 1–9). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-407156-8.00001-3

Knoblauch, K., & Maloney, L. T. (2012). *Modeling psychophysical data in r*. Springer New York, NY.

Lu, Z.-L., & Dosher, B. (2013, November). *Visual psychophysics: From laboratory to theory*. The MIT Press. https://doi.org/10.7551/mitpress/9780262019453.001.0001

Maloney, L. T., & Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision*, *3*(8), 5–5. https://doi.org/10.1167/3.8.5

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rudd, M. E., & Zemach, I. K. (2004). Quantitative properties of achromatic color induction: An edge integration analysis. *Vision Research*, *44*(10), 971–981. https://doi.org/https://doi.org/10.1016/j.visres.2003.12.004

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

Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and accurate bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research*, *122*, 105–123. https://doi.org/https://doi.org/10.1016/j.visres.2016.02.002

Shi, Y., & Eskew, J., Rhea T. (2024). Asymmetries between achromatic increments and decrements: Perceptual scales and discrimination thresholds. *Journal of Vision*, *24*(4), 10–10. https://doi.org/10.1167/jov.24.4.10

Shooner, C., & Mullen, K. T. (2022). Linking perceived to physical contrast: Comparing results from discrimination and difference-scaling experiments. *Journal of Vision*, *22*(1), 13–13. https://doi.org/10.1167/jov.22.1.13

Vincent, J., Maertens, M., & Aguilar, G. (2024). What fechner could not do: Separating perceptual encoding and decoding with difference scaling. *Journal of Vision*, *24*(5), 5–5. https://doi.org/10.1167/jov.24.5.5

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Westheimer, G. (1985). The oscilloscopic view: Retinal illuminance and contrast of point and line targets. *Vision Research*, *25*(8), 1097–1103. https://doi.org/https://doi.org/10.1016/0042-6989(85)90098-7

Zito, T., Künstle, D.-E., Aguilar, G., Berkes, P., & Schwetlick, L. (2025, January). *Psignifit 4.3* (Version v4.3). Zenodo. https://doi.org/10.5281/zenodo.14765353

7 Apendix

7.1 Pedestal discrimination result plots

On the following pages are the full pedestal discrimination results for the observers. All pedestals, that were estimated with the same surrounding, are combined into one plot. The plot shows the different pedestals with different colors. For the pedestals are in color similar psychometric functions plotted that were fitted through the results from the pedestal discrimination. Each page displays the plots for one context and one observer, on the top the unpooled and on the bottom the pooled plot.

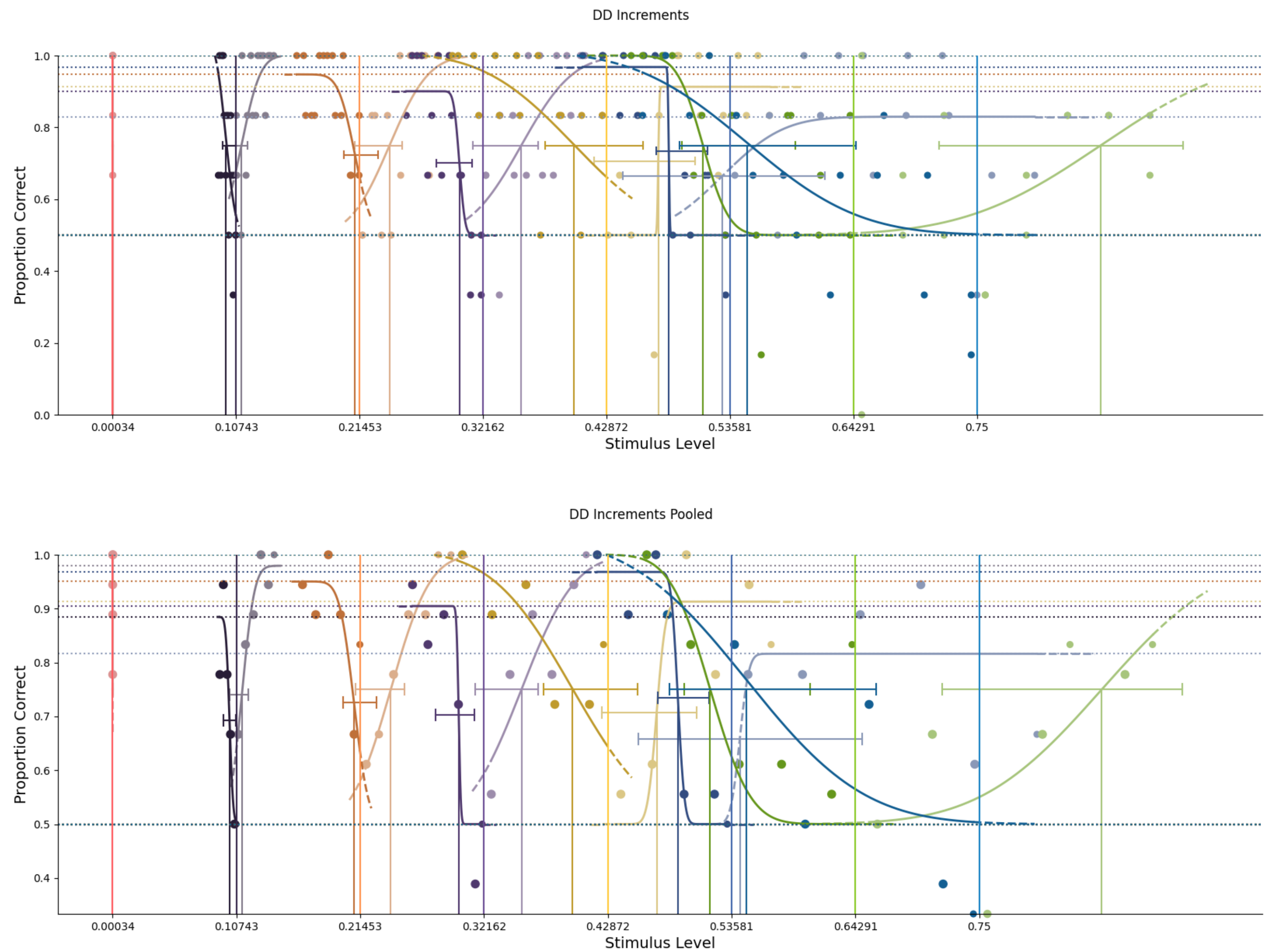


Figure 12: Pedestal discrimination results for observer DD, increments

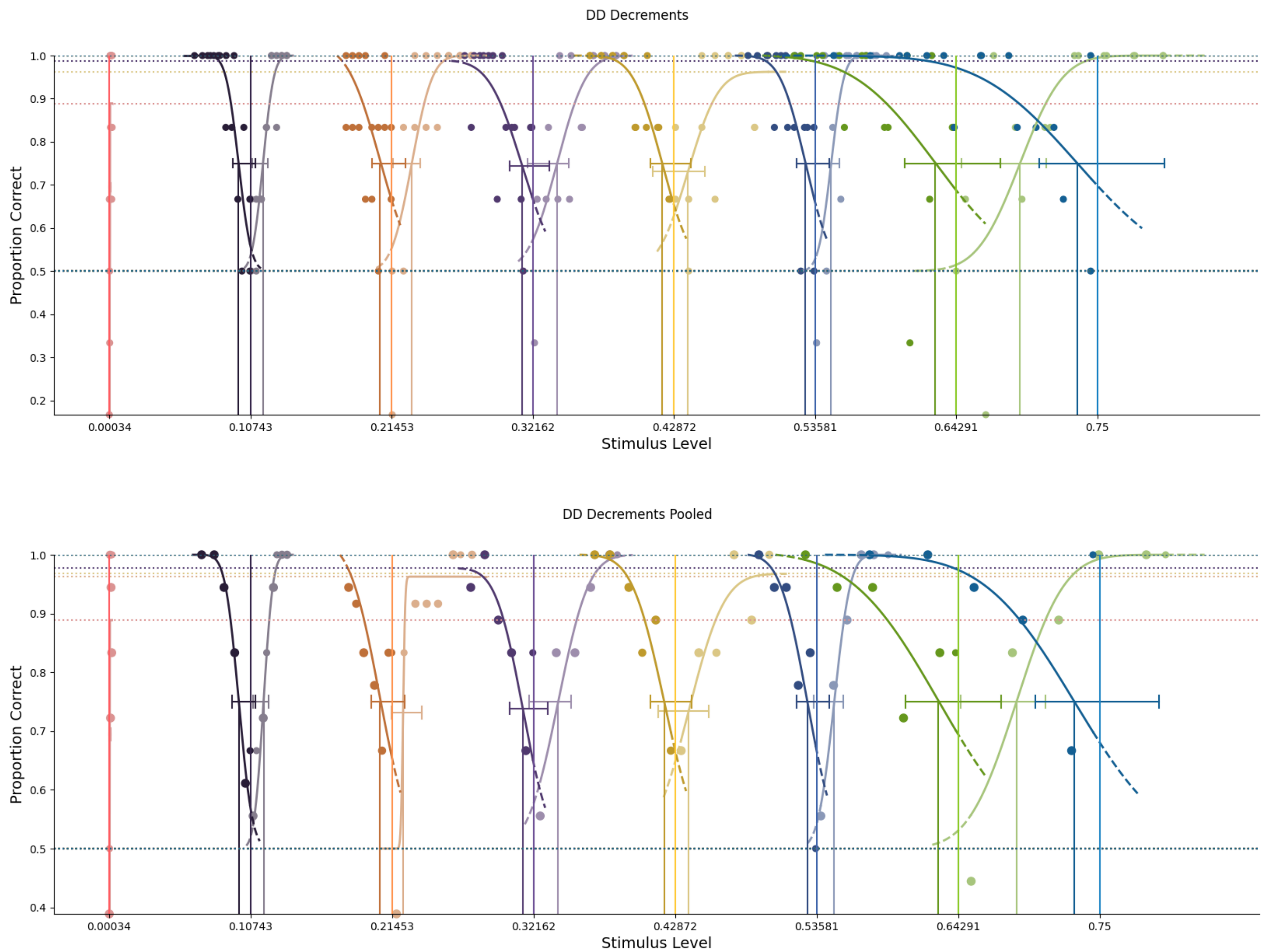


Figure 13: Pedestal discrimination results for observer DD, decrements

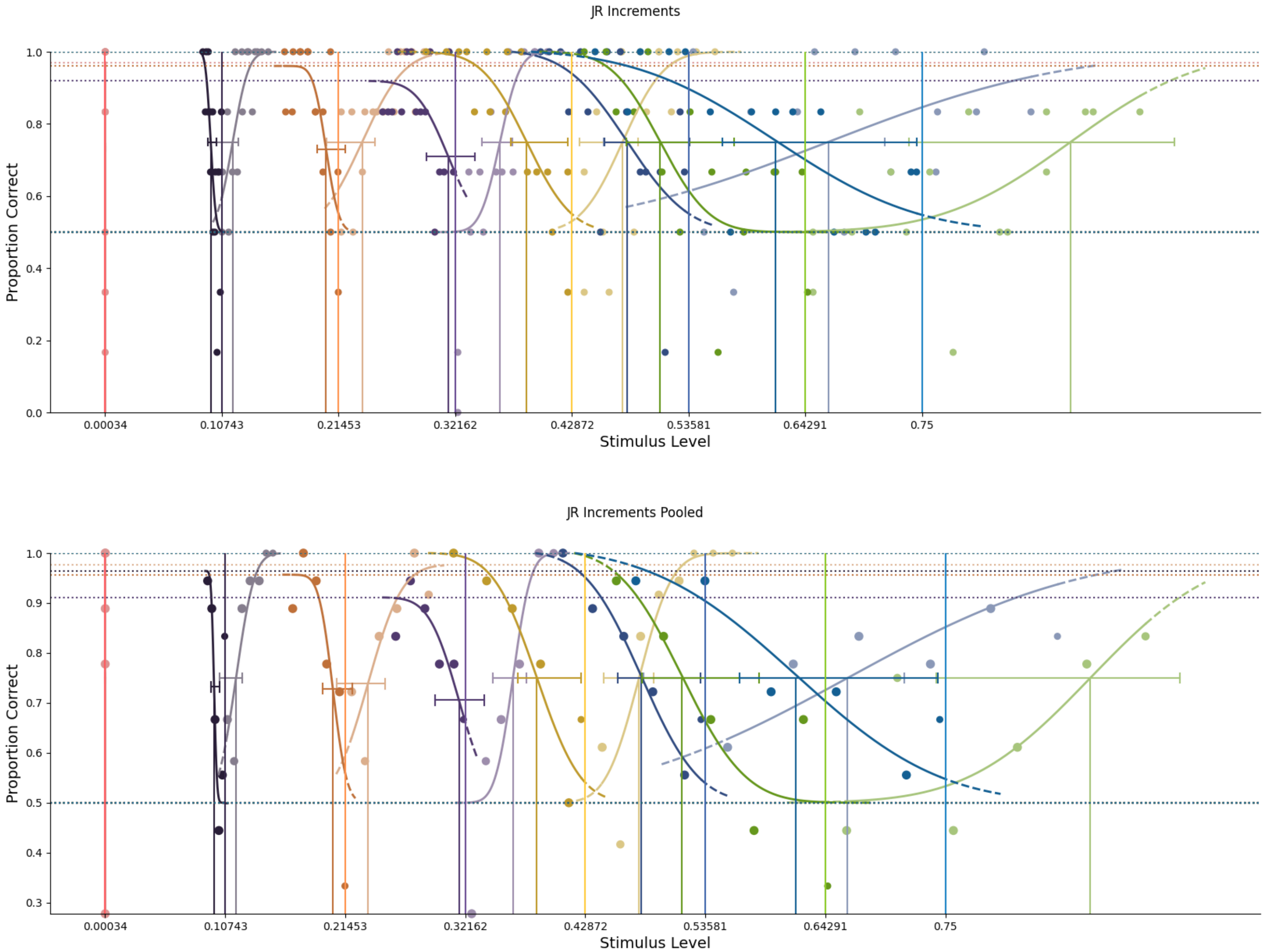


Figure 14: Pedestal discrimination results for observer JR, increments

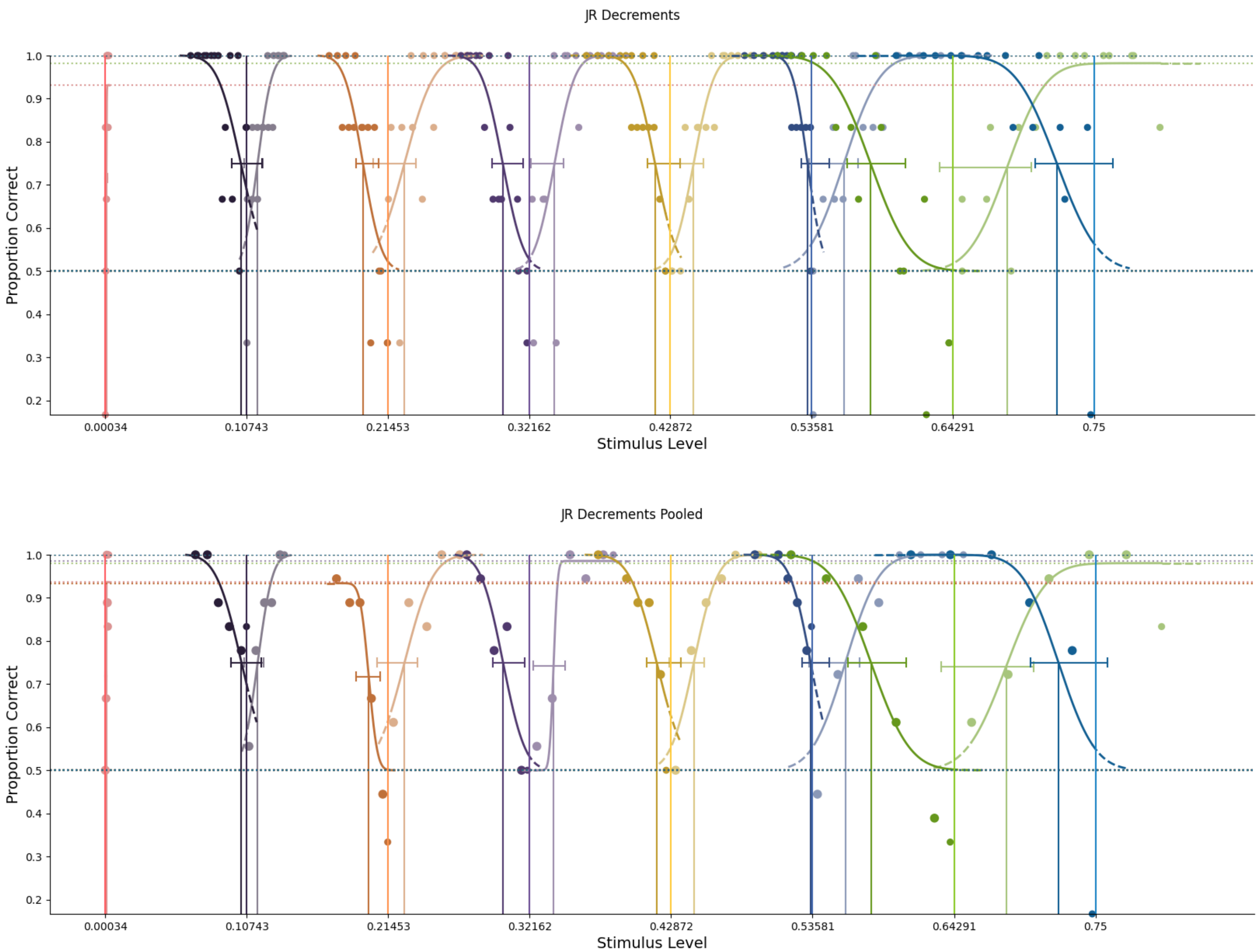


Figure 15: Pedestal discrimination results for observer JR, decrements

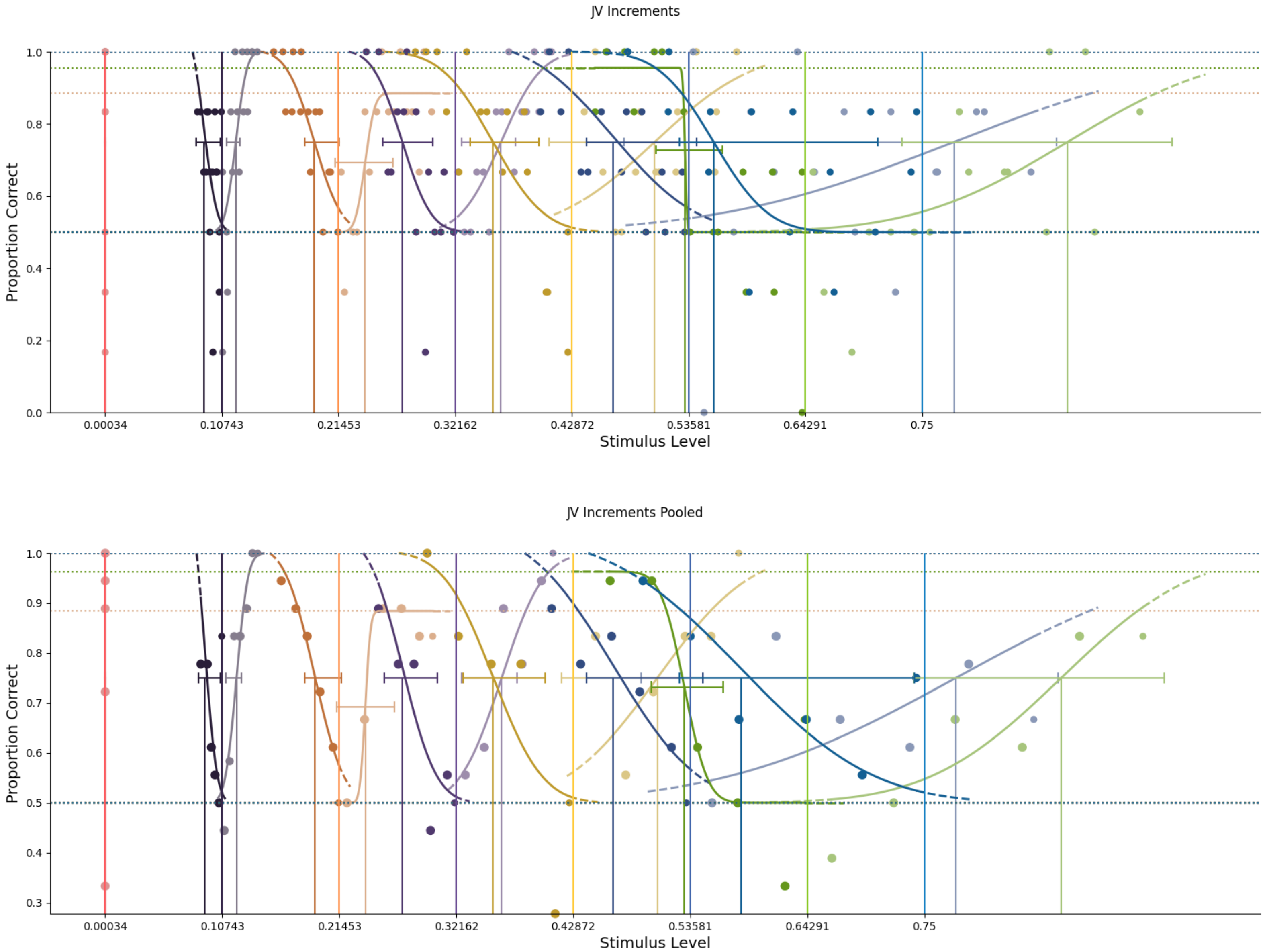


Figure 16: Pedestal discrimination results for observer JV, increments

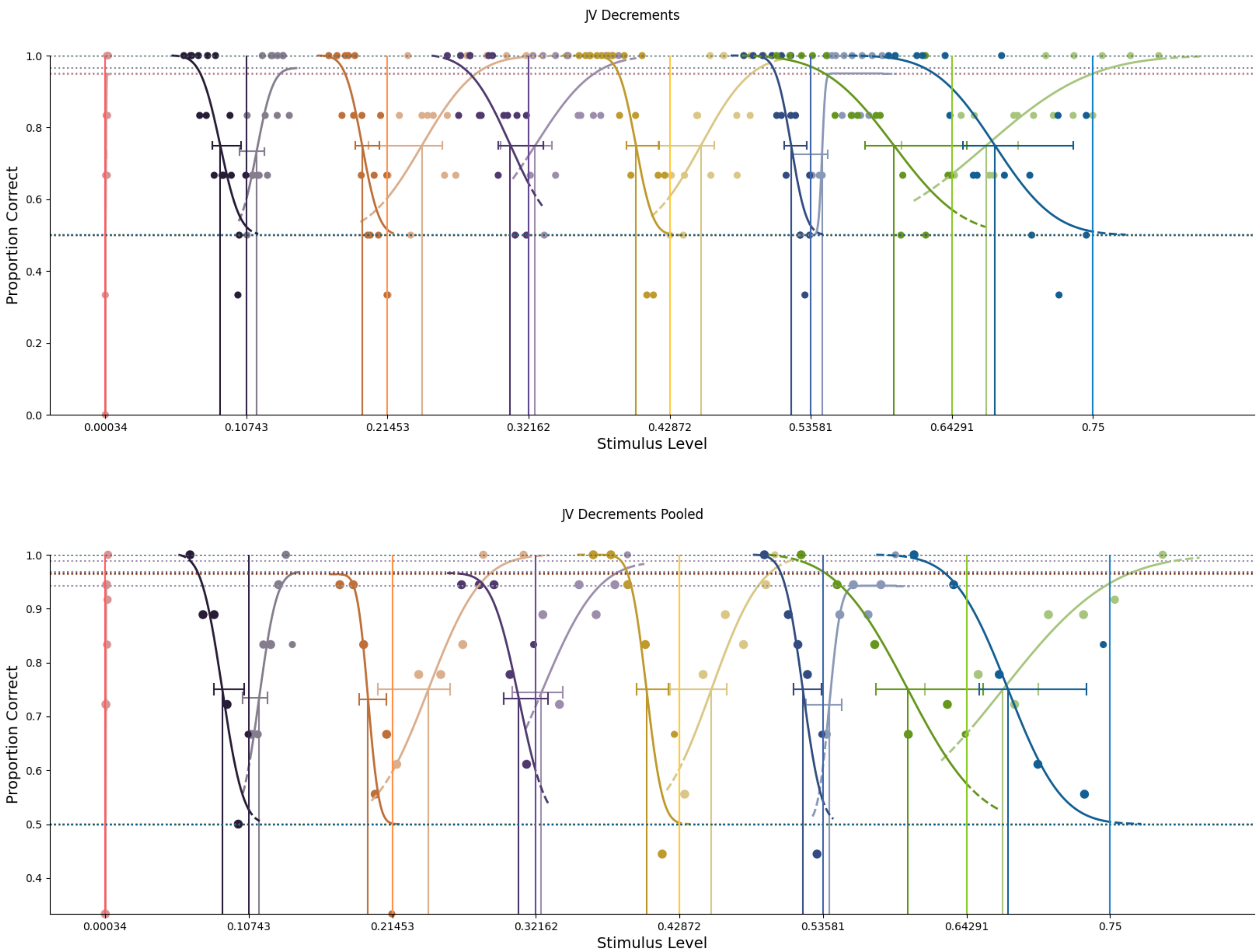
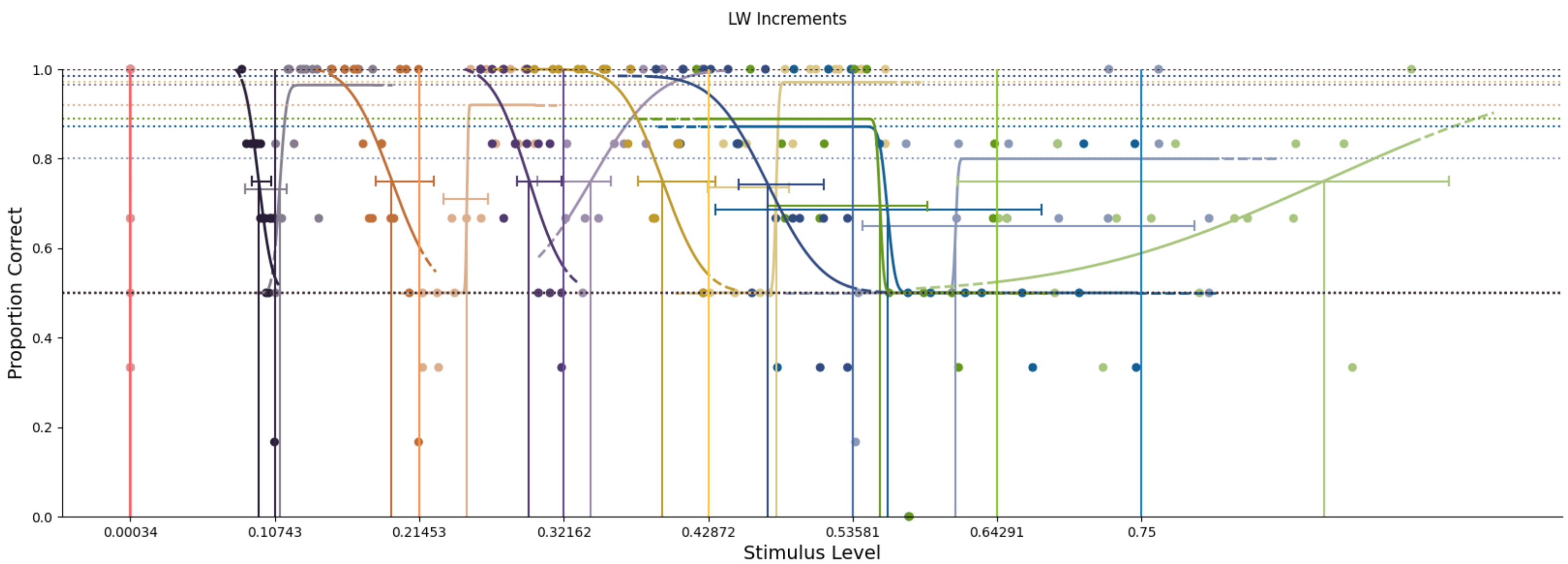


Figure 17: Pedestal discrimination results for observer JV, decrements



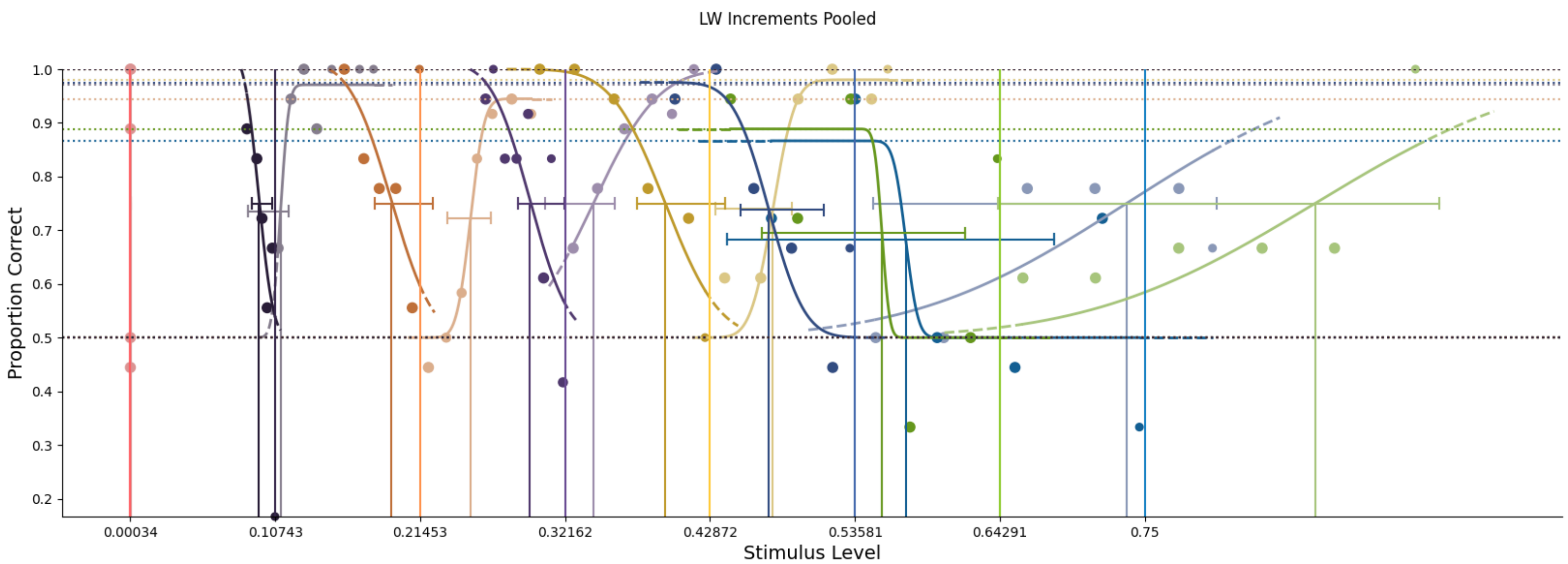


Figure 18: Pedestal discrimination results for observer LW, increments

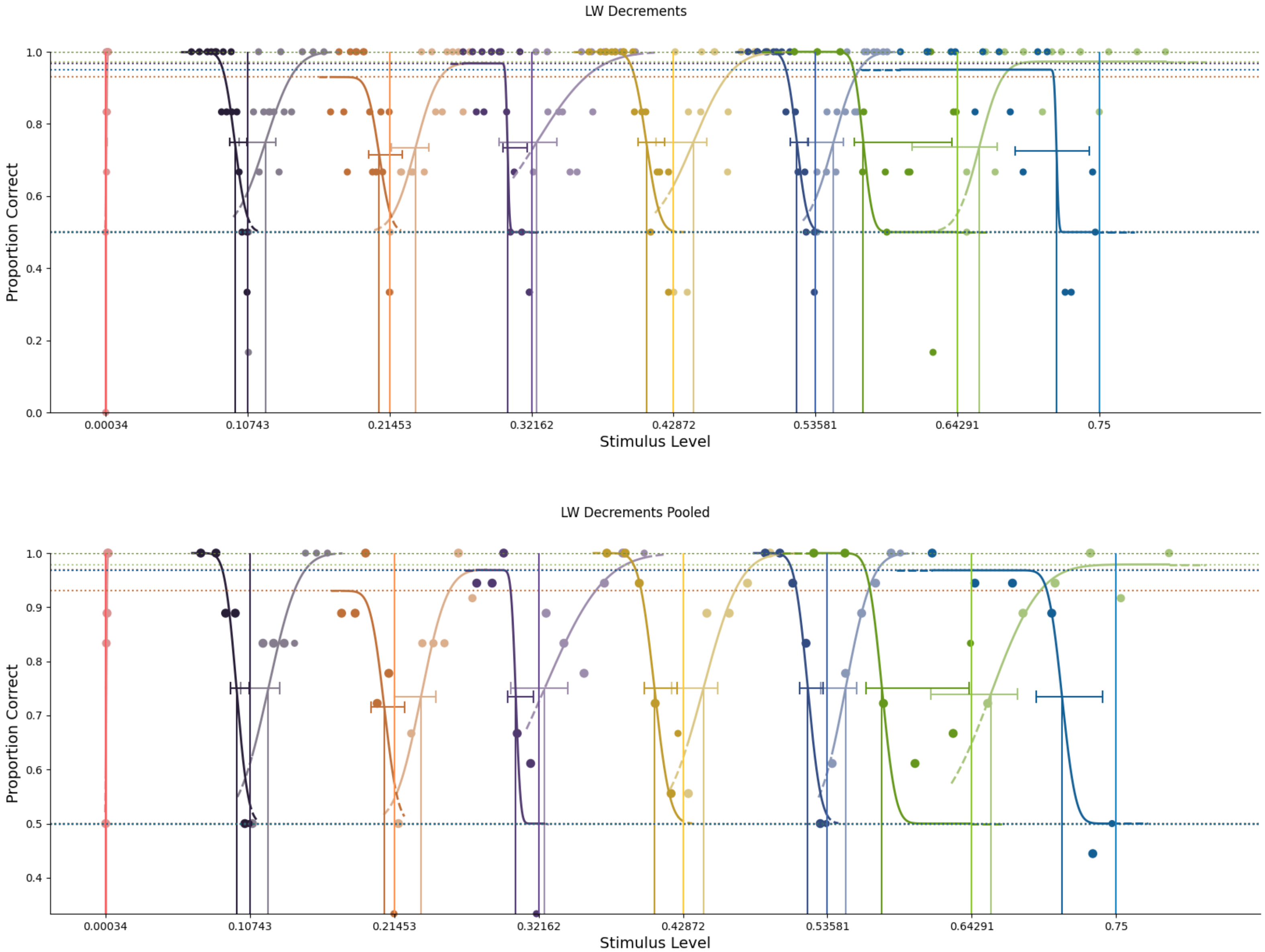


Figure 19: Pedestal discrimination results for observer LW, decrements

7.2 Luminance calculation

In different parts of this thesis are luminance values given. I want to clarify here that these values weren't physically measured and only a result from the fitting of a LUT table that is used for the mapping in the current setup. I estimated a function through the given data-points in this table and used this function to calculate the different luminances.

$$l(x) = 506.63596917 \times x + 0.61565511 \tag{3}$$

The equation used to calculate the physical luminances from the values used in the program. l is the resulting luminance and x is the stimulus intensity used in the program.

7.3 Settings table pedestal discrimination

In the following section are all the settings for the pedestal discrimination experiments. All values are the luminance in cd/m^2. Each table shows the pedestal luminance, the minimum and maximum selected test luminance for the first and second run for the pedestal discrimination. The min and max luminances are the values between the eight test contrasts that were linearly sampled. The tables are labeled with the following schematic: The two characters used for the observer's codes, the surroundings of the stimuli (black or white), and the direction of the test luminances. For every observer are individual tables for the contrast increment and decrements, pedestal and test combinations.

Table 4: DD, black surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	0.8498	0.7956	0.8498
55.0436	55.3323	71.5447	55.1297	70.0248
109.3043	110.5557	149.06	110.5557	139.9405
163.5599	164.7657	206.3099	164.7657	208.3364
217.8206	223.5355	284.3318	218.4691	289.3982
272.0763	274.1991	405.9244	274.1991	380.5926
326.337	329.929	456.588	329.929	456.588

Table 5: DD, black surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	47.7328	54.8257	47.2262	54.8257
109.3043	84.7172	109.0358	81.6774	109.1371
163.5599	132.341	162.7392	129.8078	162.7392
217.8206	150.0733	215.9359	152.6064	215.9359
272.0763	207.3231	270.146	215.9359	270.146
326.337	233.6682	324.8627	228.6018	324.8627
380.5926	243.8009	378.0595	233.6682	378.0595

Table 6: DD, white surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	1.8822	0.7956	1.8822
55.0436	55.3323	69.0115	55.3323	69.0115
109.3043	109.5424	139.9405	109.5424	139.9405
163.5599	165.2723	195.6705	164.2591	195.6705
217.8206	218.4691	253.9336	218.4691	253.9336
272.0763	272.6792	299.5309	272.6792	294.4645
326.337	329.929	405.9244	326.3826	405.9244

Table 7: DD, white surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	33.547	54.8257	38.6134	54.8257
109.3043	91.8101	109.1371	91.8101	109.1877
163.5599	137.4074	162.7392	139.9405	163.2458
217.8206	185.5378	215.9359	185.5378	216.4426
272.0763	246.3341	271.6659	253.9336	271.7166
326.337	264.0664	324.8627	266.5995	325.3693
380.5926	304.5972	378.0595	279.2654	378.0595

Table 8: JR, black surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	0.8498	0.7956	0.8498
55.0436	55.3323	71.5447	55.1297	76.6111
109.3043	110.5557	149.06	110.5557	145.0069
163.5599	164.7657	206.3099	164.7657	200.7369
217.8206	223.5355	284.3318	223.5355	264.0664
272.0763	274.1991	405.9244	279.2654	431.2562
326.337	329.929	456.588	329.929	481.9198

Table 9: JR, black surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	47.7328	54.8257	46.2129	54.3191
109.3043	84.7172	109.0358	84.2106	109.1877
163.5599	132.341	162.7392	129.8078	160.7126
217.8206	150.0733	215.9359	165.2723	215.9359
272.0763	207.3231	270.146	203.27	268.1194
326.337	233.6682	324.8627	223.5355	327.3959
380.5926	243.8009	378.0595	228.6018	375.5263

Table 10: JR, white surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	1.8822	0.7956	2.1356
55.0436	55.3323	69.0115	55.1803	69.5181
109.3043	109.5424	139.9405	110.5557	139.4339
163.5599	165.2723	195.6705	164.7657	194.1506
217.8206	218.4691	253.9336	218.4691	241.2677
272.0763	272.6792	299.5309	272.6792	329.929
326.337	329.929	405.9244	329.929	395.7917

Table 11: JR, white surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	33.547	54.8257	36.0802	54.9777
109.3043	91.8101	109.1371	86.7438	109.1371
163.5599	137.4074	162.7392	139.9405	162.7392
217.8206	185.5378	215.9359	200.7369	215.9359
272.0763	246.3341	271.6659	259.0	271.6659
326.337	264.0664	324.8627	259.0	324.8627
380.5926	304.5972	378.0595	309.6636	379.326

Table 12: JV, black surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	0.8498	0.7956	0.8498
55.0436	55.3323	71.5447	55.3323	69.0115
109.3043	110.5557	149.06	112.0756	152.6064
163.5599	164.7657	206.3099	167.8055	208.3364
217.8206	223.5355	284.3318	228.6018	294.4645
272.0763	274.1991	405.9244	279.2654	431.2562
326.337	329.929	456.588	334.9954	481.9198

Table 13: JV, black surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	47.7328	54.8257	43.6797	53.8124
109.3043	84.7172	109.0358	79.1442	109.0358
163.5599	132.341	162.7392	122.2083	162.7392
217.8206	150.0733	215.9359	145.0069	215.9359
272.0763	207.3231	270.146	203.27	269.1327
326.337	233.6682	324.8627	228.6018	324.8627
380.5926	243.8009	378.0595	243.8009	375.5263

Table 14: JV, white surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	1.8822	0.7956	2.0342
55.0436	55.3323	69.0115	55.3323	71.5447
109.3043	109.5424	139.9405	109.5424	162.7392
163.5599	165.2723	195.6705	164.2591	198.2037
217.8206	218.4691	253.9336	217.9625	253.9336
272.0763	272.6792	299.5309	272.1725	294.4645
326.337	329.929	405.9244	327.3959	380.5926

Table 15: JV, white surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	33.547	54.8257	31.0138	52.2925
109.3043	91.8101	109.1371	86.7438	109.0358
163.5599	137.4074	162.7392	132.341	162.7392
217.8206	185.5378	215.9359	183.0046	213.4028
272.0763	246.3341	271.6659	259.0	271.6659
326.337	264.0664	324.8627	259.0	325.8759
380.5926	304.5972	378.0595	302.0641	378.0595

Table 16: LW, black surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	0.8498	0.7956	0.8498
55.0436	55.3323	71.5447	55.3323	91.8101
109.3043	110.5557	149.06	110.5557	152.6064
163.5599	164.7657	206.3099	164.2591	215.4293
217.8206	223.5355	284.3318	218.4691	284.3318
272.0763	274.1991	405.9244	273.1858	405.9244
326.337	329.929	456.588	326.8892	481.9198

Table 17: LW, black surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	47.7328	54.8257	42.6664	54.927
109.3043	84.7172	109.0358	76.6111	109.0358
163.5599	132.341	162.7392	132.341	162.7392
217.8206	150.0733	215.9359	152.6064	215.9359
272.0763	207.3231	270.146	198.2037	270.146
326.337	233.6682	324.8627	213.4028	325.3693
380.5926	243.8009	378.0595	228.6018	378.5661

Table 18: LW, white surrounding, increment tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
0.7879	0.7956	1.8822	0.7956	1.8316
55.0436	55.3323	69.0115	55.3323	84.2106
109.3043	109.5424	139.9405	109.5424	137.4074
163.5599	165.2723	195.6705	164.2591	203.27
217.8206	218.4691	253.9336	217.9625	253.9336
272.0763	272.6792	299.5309	272.6792	298.011
326.337	329.929	405.9244	329.929	380.5926

Table 19: LW, white surrounding, decrement tests

	1. Run		2. Run	
Pedestal lum	min	max	min	max
55.0436	33.547	54.8257	41.1465	54.8257
109.3043	91.8101	109.1371	86.7438	109.1877
163.5599	137.4074	162.7392	142.4737	162.7392
217.8206	185.5378	215.9359	192.1241	215.9359
272.0763	246.3341	271.6659	249.8806	271.6659
326.337	264.0664	324.8627	264.0664	325.8759
380.5926	304.5972	378.0595	314.73	379.0727

List of Figures

List of Tables