

Research Team: Deena C. Elul, David H. Brainard

Research Project: Identifying individual cone fundamentals using Rayleigh matching

Experiment: First archival data collection and analysis

Background/Experimental question:

This experiment follows on our [previously pre-registered](#) pilot experiment, and the broad background is the same as described there.

As described in the pre-registration for the pilot experiment, we have developed a forced-choice Rayleigh matching paradigm with multiple reference wavelengths for use in human subjects. It employs a procedure similar to our earlier simulations and uses much of the same experimental code, but it displays lights using an actual digital light synthesis engine under computer control (OneLight Spectra). The pilot experiment provided an initial evaluation of this procedure, where we had two subjects conduct four Rayleigh matches each at six different reference wavelengths. We found that each subject's Rayleigh matches were generally consistent across experimental sessions, which suggests that our experimental results provide a reliable metric which could potentially be used to estimate individual differences. However, when we attempted to fit individual difference parameters to subjects' match data, we found that many parameters were set to extreme values or to the limits allowed by the fitting program. Further research is needed to explore whether these extreme parameter values represent subjects' underlying cone fundamentals, or whether they reflect the influence of noise on our fitting procedure. In addition, our initial experiment has led us to refine some aspects of the experimental procedure.

In the current experiment, we aim to evaluate forced-choice Rayleigh matching for a larger number of subjects, as well as to incorporate model comparison into the data analysis procedures. We will assess whether Rayleigh matches are consistent within subjects and analyze differences between subjects. We will also conduct a more careful analysis of different ways of fitting individual difference parameters to subjects' Rayleigh matching data, comparing the performance of several individual difference models with varying levels of constraint using a cross-validation approach. We will also explore the precision of our approach by performing bootstrapping to estimate the confidence intervals of the best-fitting model for each subject. Additionally, we will attempt to validate our procedure by comparing the Rayleigh matches which individual subjects make on the OneLight to the matches they make using a commercial anomaloscope, as well as through genetic analysis of subjects' L and M opsins. Taken together, these

different analyses will provide a fuller evaluation of our approach. Although we are likely to further refine experimental and analysis methods after we complete this experiment, we expect to include the results of the current experiment in an eventual publication.

Methods/Approach:

Subjects

Three subjects will participate in the experiment, drawn from the research group. All subjects will be over age 18 and will have no known color deficiencies. Prior to the experiment, potential subjects will be screened for color deficiencies using Ishihara plates, and for visual acuity using a Snellen eye chart. They will be excluded if they make any errors on the Ishihara plate test, or if their best-corrected visual acuity is worse than 20/40 in the right eye. The experimental protocols are approved by the UPenn IRB, and subjects will provide informed consent.

Sessions

It is expected that subjects will complete the experiment in three sessions—two sessions of Rayleigh matching on the OneLight Spectra digital light synthesis engine, and one session of Rayleigh matching on a commercial anomaloscope (Oculus-HMC). Each session will last 1-2 hours.

For the OneLight sessions, subjects will complete four Rayleigh matches for each of six reference wavelengths (see *Stimuli*). It is expected that for a given reference wavelength, subjects will complete two interleaved matches during the first experimental session, and two during the second experimental session. For each subject and each session, the different reference wavelengths will be tested in a pseudorandom order: a vector containing all six reference wavelengths will be shuffled by using Matlab's `randperm()` function to generate a new indexing, and the Rayleigh matches for that session will be conducted in the order specified by the shuffled vector. If subjects do not complete all their scheduled matches during the time allotted for the session, the remaining matches will be tested either in future scheduled sessions or in additional sessions, following the order specified by the shuffled vector.

Only the right eye will be tested. For the OneLight sessions, subjects will view the stimuli through the OneLight Psychophysics large-field reticle in a darkened room. For anomaloscope sessions, subjects will view the stimuli through an eyepiece with a 2° field, and the room lights will be on. Prior to testing on either device, subjects will remove all corrective lenses and adjust the eyepiece to bring the reticle into as good focus as possible.

OneLight Rayleigh Matching

1. Rayleigh Matching Procedure

Subjects will conduct a series of forced-choice Rayleigh matches using the OneLight Spectra digital light synthesis engine. They will be shown two fields in alternation: a primary mixture and a reference light. The wavelengths of these lights will be fixed for the duration of a given Rayleigh match. Each Rayleigh match will consist of a series of forced-choice trials, where the red: green ratio and reference intensity are adjusted based on subject responses.

On each trial, the two fields will be shown one after the other for 500 ms each, followed by a neutral field shown for 1 s. This stimulus sequence will repeat until the subject responds. Subjects will be prompted to make forced-choice judgements of the redness and brightness of the second field relative to the first field (e.g. “is the second field redder than the first field?”), providing separate redness and brightness judgements on each trial. Stimulus presentation will be repeated in the pattern specified above until the subject responds using a button box. Subject responses will then be used to adjust the lights for subsequent trials, following a staircase procedure: the redness judgement will be used to adjust the red: green ratio, while the brightness judgement will be used to adjust the reference intensity. Separate staircases will be maintained for the red: green ratio and the reference intensity, and the program will automatically record a subject match and terminate once both staircases have converged. In practice, the two fields should appear very similar to the subject at the match point.

2. Stimuli

Stimuli will be presented using the OneLight Spectra digital light synthesis engine, which generates narrowband spectra with width of approximately ~15-20nm FWHM around a specified center wavelength. The primary mixture will be an additive mixture of spectra of primaries at 670 nm and 560 nm center wavelength. The reference light's center wavelength will vary throughout the different Rayleigh matches in the experimental session. Each subject will complete Rayleigh matches with the reference light centered at 570 nm, 584 nm, 598 nm, 612 nm, 626 nm, and 640 nm.

The relative intensities of the primaries at their maximum settings will be adjusted so that the matches of subjects with normal color vision are not set at the limits of available settings. Based on pilot experiments and simulations, we have chosen different scale factors for matches at different reference wavelengths. The 670 nm primary will be used at its maximum intensity. The 560 nm primary in the primary mixture will be scaled to $0.02 \times$ the maximum height that the OneLight can produce for

reference wavelengths shorter than 620 nm, and to $0.004 \times$ the maximum height for reference wavelengths longer than 620 nm. The reference wavelength will be scaled to $0.1 \times$ the maximum height the OneLight can produce for reference wavelengths shorter than 630 nm, and to $0.25 \times$ its maximum height for reference wavelengths longer than 630 nm.

The neutral field will be generated from a spectral power distribution with a uniform amplitude of 0.001 across all wavelengths tested.

3. Staircase Parameters

Red: green ratio and reference intensity will be adjusted using separate staircase procedures, and a Rayleigh match will be recorded once both staircases have converged.

101 settings will be available for both reference intensity and red: green ratio, evenly spaced between 0 and 1, once the effective maximum for each primary has been determined as above. The actual number of device steps may not be 101 for some cases, due to the precision limits of the experimental device. The reference intensity setting will be applied as a multiplicative scale factor to reduce the amplitude of the reference light from its maximal value. The red: green ratio setting will be applied as a multiplicative scale factor to the amplitude of the 670 nm peak, while the 560 nm peak will be scaled by $1 -$ the red: green ratio setting. Thus, a setting of 0 will represent the greenest possible light, while a setting of 1 will represent the reddest possible light.

At the start of a given Rayleigh match, the red: green ratio and reference intensity will be set to random values. They will then be modified based on subject responses to bring the two fields into agreement. For example, if a subject responds that the (second) reference field is brighter than the primary mixture field, the reference intensity will be reduced. If a subject responds that the reference field is redder than the primary mixture field, the red: green ratio will be increased. The exact magnitude of the increase or decrease will depend on the step size at that point in the experiment. Three step sizes will be available, representing changes of 20, 5, and 1 position(s) in the settings array. At the start of a given Rayleigh match, the largest step size will be used. The step size will be reduced on the first trial where both the redness and brightness staircases have reversed at least once, continuing in this manner until the smallest step size is reached.

The program will record a Rayleigh match once both staircases have undergone four reversals at the smallest step size. The match will be represented as the average of the last four reversals for each staircase.

4. Match Interleaving

Rayleigh matches will be interleaved such that at any point during the experiment, subjects will make judgements for two different matches in alternating trials. This interleaved procedure will help prevent subjects from manipulating their performance to attain a certain number of reversals on a given staircase. The two Rayleigh matches in each interleaved pair will use the same set of primary mixture and reference wavelengths, but one will have the reference light shown first on each trial and one will have the primary mixture shown first. The interleaving will take place through sequential “super-trials” consisting of one trial for each of the two Rayleigh matches (where each trial consists of both a redness and a brightness judgement). In a given super-trial, a pseudo-random process will be used to determine which of the two Rayleigh matches should be presented to the subject first. The staircase settings and subject responses for the two Rayleigh matches will be maintained independently. Once both redness and brightness staircases have converged for one of the two Rayleigh matches, trials for the remaining Rayleigh match will be presented sequentially without interruption until the match is complete.

5. Post-session Calibration

At the end of every session, we will display the lights that subjects identified as matches and measure their spectral properties. For each match, we will display the lights corresponding to the last four reversals in the red: green ratio staircase on the OneLight, measure them using a PR670 spectroradiometer, and average the four spectral power distributions. We will repeat this procedure for the lights corresponding to the last four reversals of the reference intensity staircase. We will also measure the neutral light and the dark spd in each session.

Anomaloscope Testing

1. Rayleigh Matching Procedure

In addition to Rayleigh matching on the OneLight, subjects will complete Rayleigh matches on an Oculus-HMC commercial anomaloscope for validation purposes. This device shows a split circular field, where the top half is the primary mixture and the bottom half is the reference light. During testing, the split field is displayed for 5 s followed by a white light for 3 s in a repeating sequence. Using rotating dials, subjects can adjust the red: green ratio of the primary mixture and the intensity of the reference light to make a match, and they can also respond “match” or “no match” using button presses. Subjects can adjust the lights or respond at any point during the procedure, and there is no time limit for making matches.

a. Automated Procedure

We will use an automated anomaloscope exam to test for normal color vision. This procedure automatically samples points around the normal match point and allows subjects to adjust the reference intensity but not the red: green ratio. Subjects respond “match” if they can get the reference light to match the primary mixture at the given test point, and “no match” if they cannot. For this test, we will use the following device settings: Color Test = Rayleigh, Eye = Right, Matching Range = Absolute, Selected Examination = Normal.

b. Self-Adjusted Procedure

Following the automated procedure, we will conduct a self-adjusted procedure to determine subjects’ preferred Rayleigh match settings more precisely. The experimenter will set the anomaloscope to the average match midpoint of 40/15, where 40 represents the red: green ratio setting (on a scale of 0-73) and 15 represents the reference intensity setting (on a scale of 0-45). Subjects will then adjust both the red: green ratio and the reference intensity until the fields appear identical to them, at which point they will respond “match” and the experimenter will end the trial. This procedure will be repeated four times and will be run with the following device settings: Color Test = Rayleigh, Eye = Right, Matching Range = Absolute, Selected Examination = Manual.

2. Stimuli

The spectra used by the Oculus-HMC anomaloscope are fixed properties of the device. The primary mixture consists of two narrowband lights at 666 nm and 549 nm peak power, while the reference light has a peak at 589 nm. All spectra have a ~10 nm FWHM around their specified center wavelengths. The intensities of the primaries at their maximum settings is a property of the anomaloscope and is not immediately available to us.

3. Data Extraction

The Oculus-HMC anomaloscope is not designed to provide precise light settings data, as it is tailored for clinical diagnosis of color deficiencies. Subjects’ anomaloscope Rayleigh match settings will be obtained by digitizing the output plots produced by the anomaloscope, which plot “match” and “no match” points on a graph of reference intensity vs red: green ratio. We will likely use an online tool such as WebPlotDigitizer to extract the data from the plots.

Genetic Analysis of L and M Photopigments

At the beginning of the first experimental session, each subject will provide a saliva sample which we hope to use for genetic analysis of their L and M opsins. If a subject does not wish to provide genetic information, they will not be excluded from participation in the psychophysical portions of the experiment. Details of the opsin analysis procedures will be provided in an addendum to this pre-registration document.

Data Analysis

Cone Fundamental Fitting

1. Basic Procedure

We will fit the Asano et al. model of cone individual differences to each subject's OneLight Rayleigh match data to obtain estimated L and M cone fundamentals. The Asano et al. model accounts for how cone fundamentals are affected by eight individual difference parameters (in addition to age and field size): lens density, macular pigment density, and lambda max and optical density for L, M, and S cones. In all our fits, we will fix macular pigment density and S cone parameters to their average values, as these are believed to have minimal effect on light reaching the photoreceptor at longer wavelengths. A subset of the remaining parameters will be allowed to vary (see *Comparative Analysis and Cross-Validation* below); we will restrict these variable parameters to within 3 standard deviations (for density parameters) or 4 standard deviations (for lambda max parameters) of their means using Asano et al.'s estimates of population-level variability. When computing cone fundamentals, we will also include the actual subject age and a large (10°) field size to match the large field.

The fitting program will take in the radiometer-measured primary mixture and reference spectra which subjects identified as matches, and it will fit cone fundamentals to minimize error using an opponent contrast metric. The program will first compute cone excitations for each pair of spectra and convert these to an opponent contrast representation, using the following matrix equation:

$$\begin{bmatrix} LUM \text{ contrast} \\ RG \text{ contrast} \\ BY \text{ contrast} \end{bmatrix} = \begin{bmatrix} LUM \text{ weight} & 0 & 0 \\ 0 & RG \text{ weight} & 0 \\ 0 & 0 & BY \text{ weight} \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1 & -1 & 0 \\ -0.5 & -0.5 & 1 \end{bmatrix} \begin{bmatrix} L \text{ cone contrast} \\ M \text{ cone contrast} \\ S \text{ cone contrast} \end{bmatrix}$$

We will set 'LUM weight,' 'RG weight,' and 'BY weight' to 40.3908, 205.7353, and 62.9590, respectively. In our simulations, these values were found to bring the opponent contrast space in best agreement with the CIELAB uniform color space (CIE, 1977; Brainard, 2003), which was designed so that a unit step in any direction of the space approximates a near-threshold perceptual difference. After converting to

opponent contrast, the fitting program will calculate the vector length of luminance and red/green contrast and take the RMSE across all match pairs. The program will then adjust the free parameters to minimize the RMSE, using Matlab's `fmincon` routine with an 'active-set' algorithm.

2. Comparative Analysis and Cross-Validation

For each subject's data, we plan to conduct several versions of the basic fitting procedure described above, with varying levels of constraint. In the most unconstrained version, five parameters will be allowed to vary about their means (with the limits described above): lens density, L optical density, M optical density, L lambda max, and M lambda max. We will also conduct a fit where both lambda max parameters are constrained to their average values, and a fit where both optical density parameters are constrained to their average values. Finally, we will conduct a fit where the two optical density parameters are constrained to be equal, while the lambda max parameters are allowed to vary independently. We will also conduct a version of each of the four fits described above which locks lens density to its average value. We will compare the parameter fits obtained for each subject using each of these procedures, using the standard observer as a baseline. We will also perform cross-validation for each individual subject, analyzing which fitting procedure is best at describing the underlying structure of the subject's Rayleigh matching performance. In addition to the parameter fits, we will cross-validate the fit of the standard cone fundamentals, given each subject's age.

The cross-validation procedure will obtain the fit values on a reduced data set that leaves one match out of the four for each reference wavelength, and then evaluate the fit on the fourth match of each reference wavelength. This will be repeated four times, leaving out each of the four matches for each reference wavelength once. Across reference wavelengths, how the left-out matches are paired in each cross-validation run will be chosen randomly, and will be held constant across all nine models tested. To reduce variability due to random pairing of left-out matches across reference wavelengths, we will repeat the cross-validation procedure 10 times (with different draws on the leftout data each time) and average the cross-validated error for each model across the 10 runs. The best fitting model will be the one that leads to the lowest average cross-validated error, and we will use the parameters for this model for each subject as our estimate of that subject's parameters. The best fitting model may vary across subjects. Following the cross-validation procedure, we will use bootstrapping to estimate confidence intervals on the parameters from each subject's best cross-validated model.

Cone Excitation Visualization

We will visualize subjects' matches at different reference wavelengths using a cone contrast plot. We will compute the L and M cone excitations associated with the radiometer-measured primary mixture and reference spectra which subjects identified as matches. We will plot these in a figure with $L+M$ on the y axis and $(L-M)/(L+M)$ on the x axis. The distance between primary mixture and reference light points within each match pair will allow us to see whether observers successfully matched the lights to produce similar cone excitations. The spread of points for different match pairs at a given reference wavelength will characterize the consistency of subjects' performance.

For each subject, several versions of this figure will be produced: one using the standard cone fundamentals to compute excitations, and others using the various fitted cone fundamentals described above. This will allow us to see whether the fitted cone fundamentals reduce the distance between primary mixture and reference light points as compared to the standard cone fundamentals, indicating improved explanation of subject performance.

Anomaloscope Data Comparison

Because we do not have spectral data available for the anomaloscope, our comparison to this data will be qualitative. We are interested in whether the relative positions of matches across subjects across the two measurement devices is consistent.

Works Cited

- Asano, Y., Fairchild, M. D., & Blondé, L. (2016). Individual Colorimetric Observer Model. *PLoS One*, 11(2), e0145671. Green
- Asano, Y., Fairchild, M. D., Blondé, L., & Morvan, P. (2016). Color matching experiment for highlighting interobserver variability. *Color Research & Application*, 41(5):530, 530-539.
- Brainard, D. H. (2003). Color appearance and color difference specification. In S. K. Shevell (Ed.), *The Science of Color* (2 ed., pp. 191-216). Oxford: Optical Society of America; Elsevier Ltd.
- CIE. (1977). CIE recommendations on uniform color spaces, color-difference equations, and metric color terms. *Color Research and Applications*, 2(1), 5-6.
- He, J. C., & Shevell, S. K. (1994). Individual differences in cone photopigments of normal trichromats measured by dual Rayleigh-type color matches. *Vision Research*, 34(3), 367-376.
- Jacobs, G. H., Neitz, J., & Krogh, K. (1996). Electroretinogram flicker photometry and its applications. *Journal of the Optical Society of America A*, 13(3), 641-648.

- Moreland, J. D., & Kerr, J. (1979). Optimization of a Rayleigh-type equation for the detection of tritanomaly. *Vision Res*, 19(12), 1369-1375.
- Moreland, J. D. (2004). Moreland match revisited. *Visual Neuroscience*, 21(3), 471-476.
- Neitz, J., & Jacobs, G. H. (1986). Polymorphism of the long-wavelength cone in normal human colour vision. *Nature*, 323(6089), 623-625.
- Neitz, J., & Neitz, M. (2011). The genetics of normal and defective color vision. *Vision Research*, 51(7), 633-651.
- Sanocki, E., Lindsey, D. T., Winderickx, J., Teller, D. Y., Deeb, S. S., & Motulsky, A. G. (1993). Serine/alanine amino acid polymorphism of the L and M-cone pigments: Effects on Rayleigh matches among deuteranopes, protanopes and color normal observers. *Vision Research*, 33(15), 2139-2152.
- Stockman, A. & Brainard, D. H. (2010). Color vision mechanisms. In the OSA Handbook of Optics (3rd edition, M. Bass, ed). McGraw-Hill, New York, pp. 11.1-11.104
- Thomas P. B. Mollon J. D. (2004). Modelling the Rayleigh match. *Visual Neuroscience* 21, 477–482.