Applying Resampling and Visualization Methods in Factor Analysis to Model Human Spatial Vision

Seung Hyun Min[1] and Alexandre Reynaud[2,3]

[1]School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, China
[2]McGill Vision Research, Department of Ophthalmology and Visual Sciences, McGill University, Montreal, Canada
[3]Brain Repair and Integrative Neurosciences Program, Research Institute of the McGill University Health Center, Montreal, Canada

Correspondence: Seung Hyun Min, School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, No. 270, Xueyuan West Road, Lucheng District, Wenzhou, Zhejiang 325000, China;
seung.min@eye.ac.cn.
Alexandre Reynaud, McGill Vision Research, Department of Ophthalmology and Visual Sciences, McGill University, 1650 Cedar Ave., Rm L11.403, Montreal H3A1A1, Canada;
alexandre.reynaud@mcgill.ca.

Received: August 16, 2023
Accepted: December 19, 2023
Published: January 5, 2024

Citation: Min SH, Reynaud A. Applying resampling and visualization methods in factor analysis to model human spatial vision. *Invest Ophthalmol Vis Sci.* 2024;65(1):17.
https://doi.org/10.1167/iovs.65.1.17

Purpose. Studies have reported different numbers of spatial frequency channels for chromatic and achromatic vision. To resolve the difference, we performed factor analysis, a multivariate modeling technique, on population data of achromatic and chromatic sensitivity. In addition, we included resampling and visualization methods to evaluate models from factor analysis. These routines are complex but widely useful. Therefore we have archived our analysis routines by building *smCSF*, an open-source software package in R (https://smin95.github.io/dataviz/).

Methods. Data of 103 normally-sighted adults were analyzed. They included blue-yellow, red-green, and achromatic contrast sensitivity. To obtain the confidence interval of relevant statistical parameters, factor analysis was performed using a resampling method. Then exploratory models were developed. We then performed model selections by fitting them against the empirical data and quantifying the quality of the fits.

Results. During the exploratory factor analysis, different statistical tests supported different factor models. These could partially be reasons for why there have been conflicting reports. However, after the confirmatory analysis, we found that a model that included two spatial channels was adequate to approximate the chromatic sensitivity data, whereas that with three channels was so for the achromatic sensitivity data.

Conclusions. Our findings provide novel insights about the spatial channels for chromatic and achromatic contrast sensitivity from population data. Also, the analysis and visualization routines have been archived in a computational package to boost the transparency and replicability of science.

Keywords: contrast sensitivity, factor analysis, color vision, computational modeling, spatial vision

Factor analysis is a century-old method that aims to describe the relationships between variables from a high dimensional dataset.[1] It involves retaining the least possible number of dimensions to approximate the empirical data. This dimension is often referred to as a *factor*. For example, in psychology, general intelligence (or *g*) is an intangible quantity or value that shapes the human mind across different levels; this is an example of a model that has one latent (or implicit) factor that affects various response variables, such as mathematical aptitude, musicianship and reading comprehension. In the case of human spatial vision, a factor can govern how perception can co-vary across a certain class of stimulus but not another. In vision research, the term *channel* was initially used to vaguely refer a neurophysiological mechanism that underlies a particular visual function, such as contrast sensitivity.[2] However, physiological studies in cats and monkeys began to support the existence of channels for visual perception,[3] demonstrating that the term was not merely an abstract reference to what could occur in the visual system but an actuality. In this article, we will use the term *channel* to refer to a common mechanism that subserves a particular visual function, and employ factor analysis to gain more insights about how these channels operate in human visual system.

Humans' window of visibility can be illustrated as the contrast sensitivity function (CSF). It is often measured by logarithmically modulating the contrast and spatial frequency of a sinusoidal grating.[4] Contrast refers to the difference between brightest and darkest parts of an image. Spatial frequency indicates whether an image is fine or coarse (high or low spatial frequency). Typically, the CSF has an asymmetrical and inverted U-shape with a peak at a medium spatial frequency. In the past, studies using adaptation and masking paradigms have attempted to estimate the number of contrast spatial channels in the human visual system.[5,6] However, studies using different psychophysical methods have reported different numbers of spatial channels.[7–10] Factor analysis is an alternative solution to indirectly derive the spatial channel by evaluating the sensitivity data's covariance,[11] which can cause a joint variability of sensitivity across a range of spatial frequency but not others.

Studies using factor analysis have reported the existence of multiple spatial frequency channels (i.e., factors) that are responsible for contrast sensitivity at different ranges of spatial frequency in both infants and adults.[11–20] They have aimed to confirm the findings of psychophysical studies that directly analyzed data (but not covariance).[21,22] Specifically, Peterzell and his colleagues[14,15,19] have shown that there is one channel above 1 c/deg (Wilson A channel) and one below 1 c/deg (Wilson B channel) with a possibly third factor that accounts for covariance of sensitivity driven by optics at higher frequencies (>2.25 c/deg).[17,23] However, factor analysis can yield discrepant results depending on differences in methodologies[24] or sample sizes because it can be easily distorted by skewed values or outliers.[25–27] For example, the range of tested spatial frequency from the sensitivity data (i.e., methods for data collection) and different criteria for factor retention can both introduce different results.[24] Sometimes, the standard criteria must be adjusted depending on the local dataset.[12,15,28]

Correctly estimating the number of factors is important. However, differences in criteria for factor retention and sample sizes can yield different numbers of factors.[24,29] First, the likelihood ratio test can support one model over another with a statistical significance, such as the p-value. However, it tends to overestimate the number of factors.[30] Another method is the Guttman criterion,[31] which selects factors with eigenvalues larger than 1. Eigenvalue refers to a proportion of variance that is explained by a factor. The higher the eigenvalue for a factor, the larger the factor explains the variance of the data. This method, however, has been shown to be problematic because the exact eigenvalue can vary depending on the sample.[25] For instance, a factor's eigenvalue can exceed 1 in one sample but not in another, resulting in different number of factors depending on the sample. In vision research, factors that show systematic patterns in their loadings are qualitatively retained.[15,19,28] Also, a method that has become popular across fields is parallel analysis, which includes factors when the associated eigenvalue for each factor from empirical data exceeds that from randomly generated data.[25] In other words, the empirical eigenvalue must be higher than the random eigenvalue. However, statistical value such as factor loadings and eigenvalues themselves can vary from sample-to-sample and have a range of variability both in large and small samples. Therefore calculating the range of uncertainty, such as the confidence interval of eigenvalues, can be useful to compare those from a null distribution. Therefore these methods without a resampling method to compute the range of uncertainty are all prone to the variability of the eigenvalue or factor loadings, and can support different models depending on the local dataset.

These tools are exploratory measures of factor analysis that can be used to develop a potential model (i.e., factor model). However, their support does not *confirm* whether the fit of the factor model is adequate. The standard approach for confirmatory factor analysis requires a sample size of at least 200.[27] In cases of questionnaires from psychometrics, a large sample size can be achieved and clear conclusions from a confirmatory factor analysis can be formulated because one test can record responses for multiple response variables at once.[32] However, in ophthalmology and vision research where one datum is a summarized data of multiple trials, there is no systematic and uniform confirmatory approach to determine the number of factors in a model because meeting the requirements for a standard confirmatory factor analysis is not feasible. Fortunately, contrast sensitivity data can be visualized, and the fit of the factor model can be visually inspected against the experimental data. With modern tools, the fitted data from the factor model can be plotted and stochastic (random) methods can be implemented without consuming excessive computer memory. In other words, in vision research and ophthalmology, because of contemporary tools for factor analysis and availability of diverse software packages, it has now become possible to confirm whether a given factor model is appropriate even if the data do not meet the standard requirements for confirmatory factor analysis.

The aim of our study was twofold. First, we wanted to resolve the conflicting reports that suggest different numbers of spatial frequency channels in humans. To do so, we initially used standard routines for factor analysis to derive a factor model. These tests supported different models. So, we used a series of statistical tests for model comparisons, such as the likelihood ratio test and examination of R^2 for the model fit. Moreover, we applied a statistical method that generates a range of error for eigenvalues from bootstrap resampling to derive the most appropriate factor model. Bootstrapping is a stochastic procedure of resampling data with replacement,[33] which is repeated numerous times. It can be used to generate the range of error, such as the confidence interval, of statistical values that can be only collected once from the raw data. Also, as a confirmatory measure, we used a visualization method to inspect whether the fit of the selected factor model was faithful to the experimental data. Second, to promote reproducibility and replicability of research practices in factor analysis for vision scientists and ophthalmologists, we compiled our analysis and visualization routines for contrast sensitivity data in an R package *smCSF* (https://smin95.github.io/dataviz/: Chapters 13–16).

Methods

This study used two published datasets, both of which have been uploaded online (https://mvr.mcgill.ca/AlexR/data_en.html). Datasets from 62 normally sighted observers of Reynaud et al.,[34] and 51 from Kim et al.[35] were analyzed separately. In both studies, the quick Contrast Sensitivity Function (qCSF) method was used to measure the sensitivity at various spatial frequencies.[36,37] For each qCSF test, there were 100 trials, and each test would take about eight minutes.

Data From the Study of Kim et al.[35]

Data from the study of Kim et al.[35] includes contrast sensitivity of 51 adults with normal vision for red-green, blue-yellow, and achromatic noise patterns. Noise gratings were generated in the space domain by filtering white noise by an oriented Gabor filter and were presented in a 5° Gaussian window. Each measurement was performed twice, and the values were then averaged across the two repetitions. The color sensitivity data were obtained from 0.25 c/deg to 2.54 c/deg, whereas the achromatic data were obtained from 0.25 c/deg to 9.57 c/deg.

Data From the Study of Reynaud et al.[34]

This is a dataset of contrast sensitivity for noise gratings (same properties, but presented in a 10° window) from

both dominant and non-dominant eyes of 52 adults.[34] In our computational study, only the data of the dominant eye's sensitivity was used because a preliminary analysis, using the Kaiser-Meyer-Olkin test, revealed that the dataset was more apt for factor analysis than that from the non-dominant eye. The range of tested spatial frequencies was from 1 c/deg to 14.16 c/deg.

Exploratory Factor Analysis

Factor analysis was performed using the software R with the *psych* package's *fa* function.[38] Factor solutions underwent the varimax rotation,[39] which maximizes the variance of the squared loadings and the interpretability of results from the solutions.

Various methods were performed to estimate the number of significant factors. First, we performed a traditional statistical test for model comparisons, notably the χ^2 likelihood ratio test. It tests whether a model with more factors can fit the model better than a simpler model with fewer factors with a statistical significance. Next, the simple scree test was used, a qualitative method that has been used in vision research[14,19] and involves locating where the plot "breaks" into a flat line. The breaking point is noted as the boundary between significant and insignificant factors. Third, factors with loadings showing systematic patterns were kept. This qualitative method has been used in previous psychophysical studies.[19,28] For example, if a three-factor model is used to derive the loading scores from factor analysis on a dataset that supposedly has two factors, then the loading scores for the third factor should not show meaningful pattern with high values.

Furthermore, parallel analysis was performed by computing the eigenvalues from the empirical and randomly generated data matrix, both of which have the same size.[25] The 95% confidence interval of the eigenvalues was obtained using the resampling method (see the Appendix) to counter against the potential variability of eigenvalues. Then, the eigenvalues and their confidence intervals from empirical and random data were visualized as the scree plot.[40,41] The factors whose confidence intervals of the eigenvalue from the raw data do not overlap with those from factors from the random data are deemed to be statistically significant. These steps are compiled in the functions of the *smCSF* package (see the Appendix).

Generating Predicted Values From the Model

To confirm whether the model from exploratory factor analysis was appropriate, we plotted the fitted values from the model against the empirical data. The fitted values were computed from weights, which could be calculated as following:

$$\beta = \mathrm{X}^{+}\mathrm{y},$$

where β is the weight of each latent factor at each (each spatial frequency) variable, X^{+} is the Moore-Penrose pseudo inverse of the loadings for each factor, and y is the matrix of the raw data (contrast sensitivity).[42,43] By performing a matrix multiplication between the coefficient weights β and the matrix of the loadings, we computed the following fitted value:

$$\hat{\mathrm{y}} = \mathrm{X}\,\beta,$$

where X refers to the loading scores from factor analysis and $\hat{\mathrm{y}}$ refers to the fitted value of the sensitivity data. To assess the quality of the fit, we calculated the coefficient of determination (R^2) between the fitted values from the factor model and the experimental data.

Calculating the Tuning of Each Spatial Channel

The spatial frequency tuning of each factor channel was estimated using the loadings from factor analysis using the formula below[11,19,44]:

$$\text{Channel contrast sensitivity of channel}_{\text{in}} = \text{Mean contrast sensitivity}/\text{abs}\left(1/\text{factorloadings}_{\text{in}}\right)^{1/Q}$$

where the statistical factor loadings are from factor i at spatial frequency n. The value of Q was set to 4 based on earlier works.[13,19] The derived sensitivity values were then used to derive the qCSF parameters (see the Appendix) to obtain the truncated log-parabola model of CSF so that the smooth tuning function of each spatial channel could be computed across the tested range of spatial frequency. This analysis comes along with an assumption of the combination rule, pointing that different spatial channels can *combine* to influence contrast sensitivity.[44] In addition, the derived sensitivity values were fitted against Wilson's model[45] (Equation 1 in the previous study) to estimate the tuning of each statistical factor.

Results

Data of Kim et al.[35]: Chromatic and Achromatic Sensitivity

Combined Factor Analysis. A factor analysis using the combined dataset of all stimuli, including achromatic, red-green and blue-yellow sensitivity data across all spatial frequencies, was performed to examine whether the spatial channels responsible for chromatic and achromatic vision operate independently.[16] These datasets were combined as if the three sensitivity functions were part of a single matrix. Specifically, the three datasets were concatenated one after the other on the spatial frequency dimension. According to the Kaiser-Meyer-Olkin test (most values > 0.6) and Bartlett's test for sphericity ($P < 0.05$), which is a test that compares the identity matrix to the observed correlation matrix, combined dataset was appropriate for factor analysis.

The scree plot was visualized to deduce the number of factors that underlay the variance of the combined dataset. According to the scree plot from Figure 1, there were five factors whose eigenvalues were larger than 1, passing the Guttman criterion. Also, parallel analysis revealed that five factors would be sufficient for the combined dataset. On the other hand, the simple scree test revealed that seven factors were meaningful (see Fig. 1; red and brown points). Also, six factors were found to be significant because they showed loadings with systematic patterns.

In this article, loadings from five factors are listed in the Table because more tests supported the five-factor model as shown above. Factor loadings describe the correlation between a specific spatial frequency (i.e., response variable) and the factor of interest (ex. Factor 1). For instance, Factor 1 has high loadings for the red-green sensitivity data, indicat-

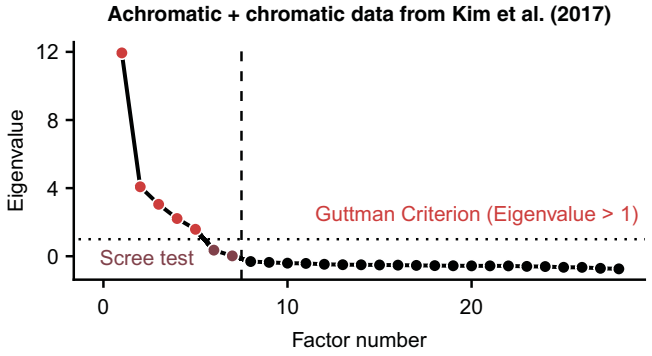

Figure 1. Scree plot of combined data of achromatic and chromatic sensitivity from the data of Kim et al.[35] Five factors (*red points*) had eigenvalues that are larger than 1, passing the Guttman Criterion test. Seven factors passed the simple scree test (*red* and *brown points*), a qualitative method to identify the factor where the eigenvalue decreases to near zero. Eigenvalue refers the measure of variance in the data that is explained by a particular factor.

Table. Loading Scores of the Combined Dataset Above 0.5 With the Exception of Fourth Factor's 0.453 (*)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
0.25 c/d – BY					0.913
0.35 c/d – BY					0.916
0.48 c/d – BY					0.869
0.68 c/d – BY				0.602	0.751
0.94 c/d – BY				0.782	0.576
1.31 c/d – BY				0.902	
1.83 c/d – BY				0.958	
2.54 c/d – BY				0.962	
0.25 c/d – RG	0.804				
0.35 c/d – RG	0.881				
0.48 c/d – RG	0.940				
0.68 c/d – RG	0.945				
0.94 c/d – RG	0.890				
1.31 c/d – RG	0.771				
1.83 c/d – RG	0.601				
2.54 c/d – RG				0.453*	
0.25 c/d – Ach		0.633			
0.35 c/d – Ach		0.783			
0.48 c/d – Ach		0.881			
0.68 c/d – Ach		0.935			
0.94 c/d – Ach		0.935			
1.31 c/d – Ach		0.851			
1.83 c/d – Ach		0.680	0.704		
2.54 c/d – Ach			0.860		
3.54 c/d – Ach			0.937		
4.93 c/d – Ach			0.936		
6.87 c/d – Ach			0.850		
9.57 c/d – Ach			0.709		
Proportion of variance	0.205	0.191	0.175	0.165	0.140
Cumulative variance	0.205	0.396	0.571	0.735	0.876

Different factors can possibly represent different spatial channels.

ing that the data from red-green sensitivity are highly correlated with each another but not to those from other stimuli types. In other words, the red-green sensitivity seems to be processed independently by a separate mechanism. In addition, if a pair of two spatial frequencies is strongly correlated, then their loadings will be comparable. For example, loadings from 0.25 c/deg to 0.94 c/deg in the red-green sensitivity data are similar but not those from high spatial frequencies. This implies that there is covariance among the data at lower spatial frequencies but less so at higher spatial frequencies.

Fortunately, the results from the five-factor solution are highly interpretable. The first factor is concentrated on the red-green contrast sensitivity data except at 2.54 c/deg. It accounts for 20.5% of the total variance, which is not largely different from those of other four factors, indicating that there is no single factor that underlies all data. Factor 2 is mainly loaded onto achromatic sensitivity data at low spatial frequencies (0.25 to 1.83 c/deg). It accounts for 19.1% of the variance. Factor 3 accounts for 17.5% of the variance and mainly loads to achromatic sensitivity data at mid-to-high spatial frequencies (1.83 c/deg to 9.57 c/deg). We can infer from these results that there are separate channels for low and high spatial frequencies for achromatic sensitivity. However, upon a closer inspection, we can see that the loadings have a wide range (0.6 to 0.9) for achromatic sensitivity from 0.25 c/deg to 1.83 c/deg even if they

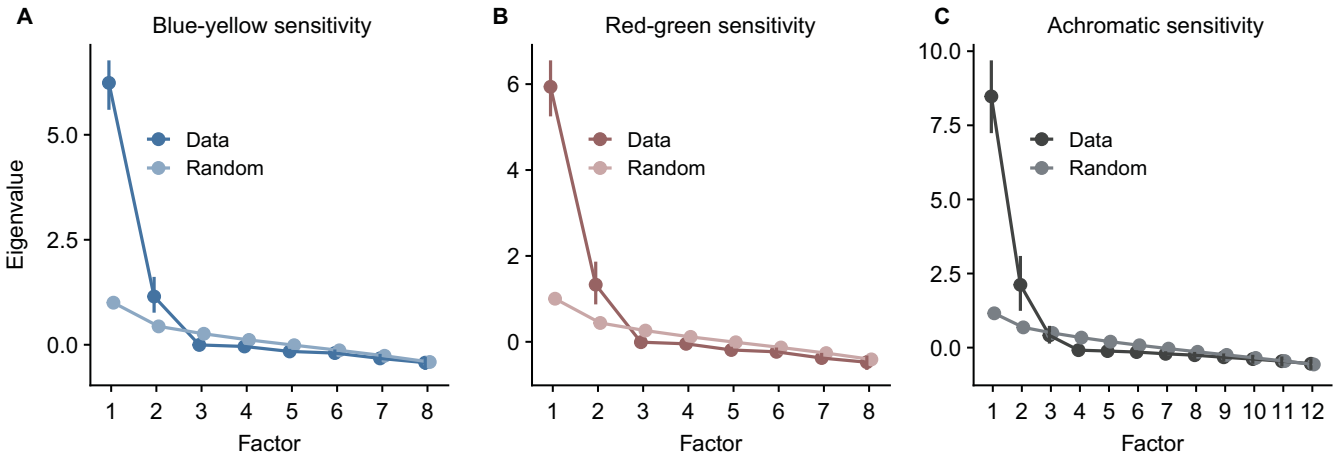

Figure 2. Eigenvalues with their range of error (95% confidence intervals) for the data of contrast sensitivity from blue-yellow (**A**), red-green (**B**), and achromatic stimuli (**C**). Darker shades indicate eigenvalues from the dataset. Lighter shades indicate results from random sampling in parallel analysis. In the parallel analysis, the Data and Random curves intersect at the third factor for all three stimuli, supporting a two-factor model.

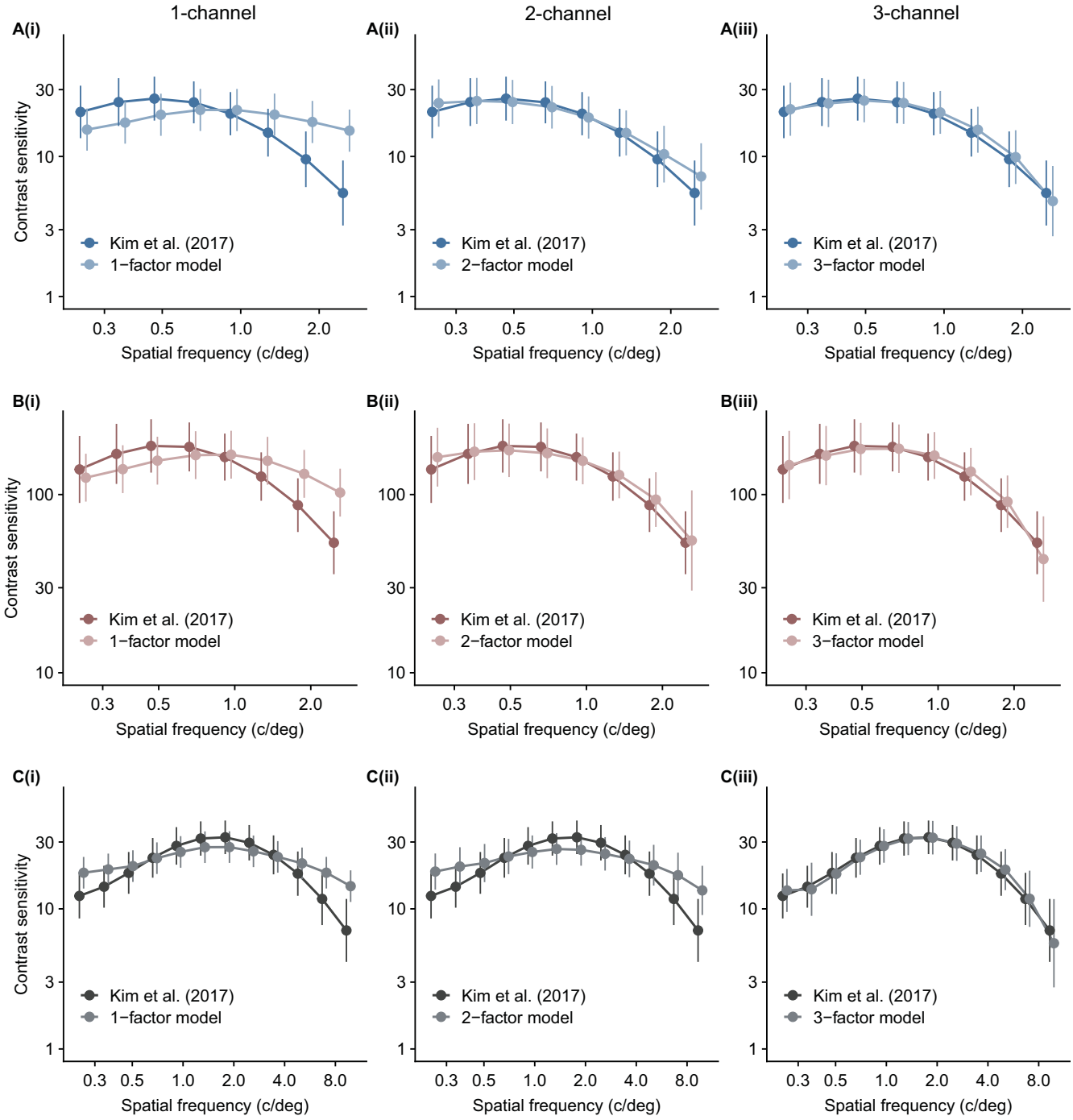

Figure 3. Model fit against the empirical data. The first column represents one-factor models. The second column represents two-factor models. The third column represents three-factor models. (**A**) Blue-yellow sensitivity data between 0.25 c/deg and 2.54 c/deg. (**B**) Red-green sensitivity data between 0.25 c/deg and 2.54 c/deg. (**C**) Achromatic sensitivity data between 0.25 c/deg and 9.57 c/deg. *Error bars* represent standard deviations across subjects.

are all concentrated on the second factor. This wide range of the loadings could indicate that multiple factors could account for the covariance of the achromatic data in this range of spatial frequency, thereby warranting a separate factor analysis. Factor 4 accounts for an additional 16.5% of the variance and is mainly loaded onto blue-yellow sensitivity data at spatial frequencies ranging from 0.68 c/deg to 2.54 c/deg. The range of the loadings is broad (0.6 to 0.96). Last, Factor 5 accounts for 14% of the total variance and is primarily responsible for the variance of blue-yellow sensitivity data at low spatial frequencies (0.25 c/deg to 0.94 c/deg). We can deduce that there are separate channels for low and medium spatial frequencies in contrast sensitivity for blue-yellow stimuli. In short, it seems that no factor is shared among separate classes of the visual stimulus.

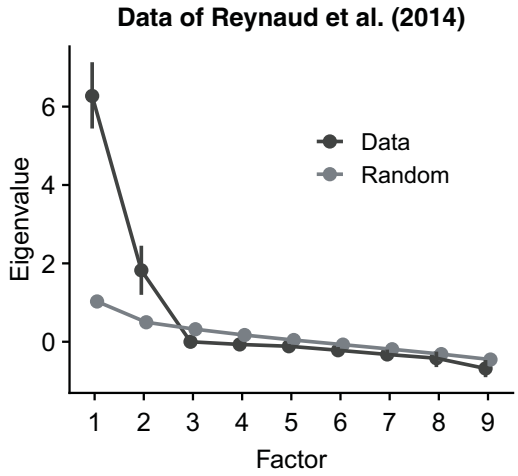

FIGURE 4. Scree plot for the data of Reynaud et al.[34] Eigenvalue with its range of variability (95% confidence intervals) for each factor is visually illustrated. The parallel analysis revealed that two-factor model might be adequate to describe the empirical data.

The cumulative proportion of variance that are explained by the five factors is 87.6%, indicating that no other factors are necessary to be included. It seems that the mechanisms of visual processing for these three classes are unique. Surprisingly, we see that one factor is enough to account for the data of red-green sensitivity. This could open up two possibilities. First, there could be a single spatial-frequency channel for red-green sensitivity. Second, the difference brought by the multiple factors in red-green sensitivity could be more subtle than those from achromatic and blue-yellow sensitivity, thereby making it difficult for the combined factor analysis to detect. The wide range of loadings supports the second possibility. In sections below, we will perform factor analysis for each class of visual stimulus so that we can analyze the data upon closer inspection.

Exploratory Factor Analysis for Each Class of Stimulus. Exploratory factor analysis was separately performed for each class of visual stimulus to develop a preliminary factor model. According to the factor analytic results from blue-yellow data (see Fig. 2A), the confidence intervals of the eigenvalues for the first two factors did not overlap with those of the first two factors from the random data (see Fig. 2A). The scree test supported two factors because the scree plot "broke" between the second and third factor. Also, two factors showed loadings with systematic patterns but not others, supporting a two-factor model. However, the likelihood ratio test supported the three-factor model ($P < 0.05$). In sum, most methods supported the two-factor model but with the exception from the likelihood ratio test.

Next, factor analytic results were used to plot the eigenvalues of factors and evaluate their loadings from red-green sensitivity data. The scree plot broke between the second and third factor, revealing that a two-factor model is adequate (see Fig. 2B). Also, the confidence intervals of the eigenvalue associated with the two factors did not overlap between data and randomly generated data confirming this observation (Fig. 2B). Furthermore, loadings from two factors exhibited systematic relationships, whereas those from others were random. However, the likelihood ratio test supported the three-factor model ($P < 0.05$). As was the case with the blue-yellow sensitivity data, the analyses yielded contradictory results regarding which model is most appropriate.

Finally, the achromatic sensitivity data were examined in a similar fashion. The breaking point of the scree plot was between the third and fourth factors, supporting the three-factor model (Fig. 2C). Loadings from three factors showed systematic patterns but not from subsequent factors. However, parallel analysis and the confidence intervals of the eigenvalues from resampling revealed that two factors were significant (Fig. 2C). Also, the likelihood ratio test supported the three-factor model. The factor analytic results from the achromatic data demonstrate a strong example of inconsistent results from different methods that estimate the number of factors.

Together, in light of our findings, factor analysis yielded mixed results regarding which model best describes the data across all three classes of the visual stimuli. For instance, analyses that assessed the quality of the model based on eigenvalues supported the two-factor model, whereas the likelihood ratio test supported the three-factor model with a statistical significance. These results demonstrate that even with a sample size of 51 with multiple experimental conditions, where each datum represents a summarized value of 100 trials, developing an appropriate model using exploratory factor analysis is not a straightforward process that leads to a clear, *unequivocal* result.

Are Spatial Channels Necessary To Preserve the Shape of the CSF?. In this section, we examined whether the presence of increasingly numerous channels could determine the shape of the CSF. To do so, we generated a model fit (i.e., prediction curve) based on the factor scores of each observer. The matrix of the factor loadings and the matrix of the raw data were multiplied (see Methods) to obtain the fitted values. Then, we plotted the average and its standard deviation of the fitted data for each model and stimulus category (see Fig. 3). To begin with, both chromatic and achromatic sensitivity curves seem to require multiple spatial channels to capture the typical asymmetry and the inverted U-shape of the CSF along the log x-axis. We could observe from the first column (i) of Figure 3 that, when there is only one spatial frequency channel in the model, this fitted function does not capture the asymmetrical and inverted-U shape of the CSF, illustrating that one-factor model is inadequate.

However, for chromatic sensitivity (rows A and B in Fig. 3), the fitted functions from two- and three-factor models capture the true asymmetrical and inverted-U shape of the empirical curve. The third factor only explained 2% and 5%, respectively, of the total variance for blue-yellow and red-green sensitivity, suggesting that the third factor does not significantly improve the fit of the model. In addition, the distributions of R^2 for both three-factor models ranged from 0.95 to 1; this range is very narrow and high, demonstrating that the three-factor model is possibly overfitting the data.

However, there seems to be a considerable difference in the quality of the model fit between two- and three-factor models in the achromatic data. In fact, the first two factors of the three-factor model for the achromatic data explained 75% of the variance, and the third factor explained 25% of the total variance, demonstrating its importance. This would argue for the existence of 3 channels in the larger spatial frequency range which was tested for

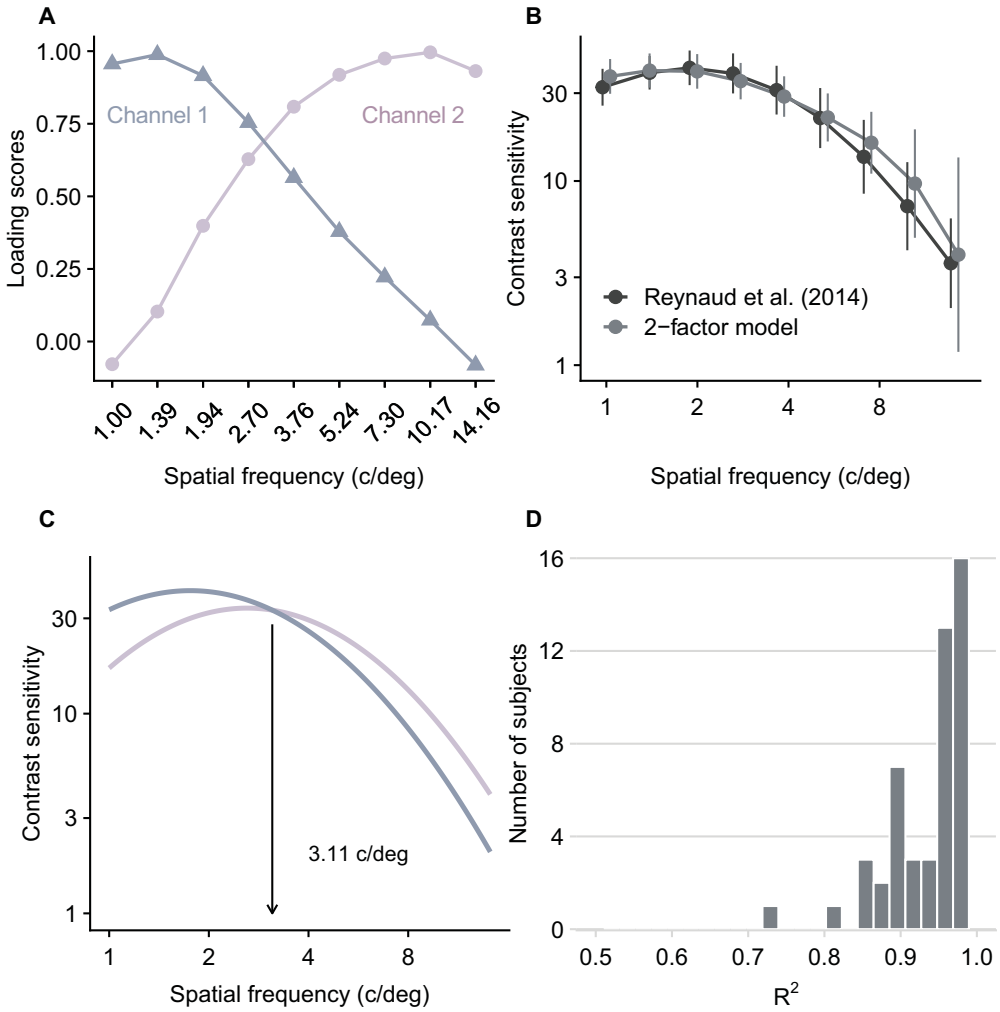

Figure 5. Evaluation of the two-factor model for the data of Reynaud et al.[34] from 1 c/deg. (**A**) Loading scores of the two factors from the model. They intersect at between 2.7 c/deg and 3.76 c/deg. (**B**) Model fit against the empirical data. The model fit is faithful to the empirical data. Error bars represent standard deviations across subjects. (**C**) Spatial tuning of the two channels. The point of intersection is 3.11 c/deg between the two spatial channels. (**D**) Evaluation of the model fit based on R^2 for each subject. This histogram demonstrates that there is an excellent fit of the model against the data.

achromatic sensitivity (0.25 c/deg to 9.57 c/deg), in accordance with the previous likelihood ratio test but in contradiction with parallel analyses and the resampling method we previously performed. So, to reappraise the number of channels for achromatic spatial vision, we decided to replicate the analysis using a different dataset of achromatic sensitivity from the study of Reynaud et al.,[34] but with a different range of spatial frequency from 1 c/deg to 14.16 c/deg.

Data of Reynaud et al.[34]: Achromatic Sensitivity Above 1 c/deg

Factor analysis using the achromatic data of Kim et al.[35] revealed mixed results regarding which model is most appropriate. Therefore, we analyzed the achromatic data of the study by Reynaud et al.,[34] who tested 52 normally-sighted adults; this dataset contains tested spatial frequencies between 1 c/deg and 14.16 c/deg. Using the dataset would enable us to examine whether the inclusion of only one spatial channel in the factor model at a frequency range above 1 c/deg is adequate. First, the scree test (Fig. 4) had a breaking point between the second and third factors, indicating that a two-factor model is sufficient. The confidence intervals of the eigenvalues from two factors were also significant. Parallel analysis also supported a two-factor model. The likelihood ratio test endorsed the two-factor model. Furthermore, loadings from two factors showed meaningful patterns. To summarize, the different tests unanimously supported the two-factor model for achromatic vision above 1 c/deg.

Several observations can be made from Figure 5, where the loadings of the two factors, fit of the two-factor model and the distribution of R^2 across 52 observers are shown. First, the loading scores (see Fig. 5A) demonstrate that the two spatial channels intersect at about 3 c/deg, which is approximately where the peak of the achromatic CSF is located. We also see that the two-factor model can faithfully reproduce the achromatic curve at a range of spatial frequency above 1 c/deg (Fig. 5B). The spatial tunings of the two channels are shown in Figure 5C, indicating that the

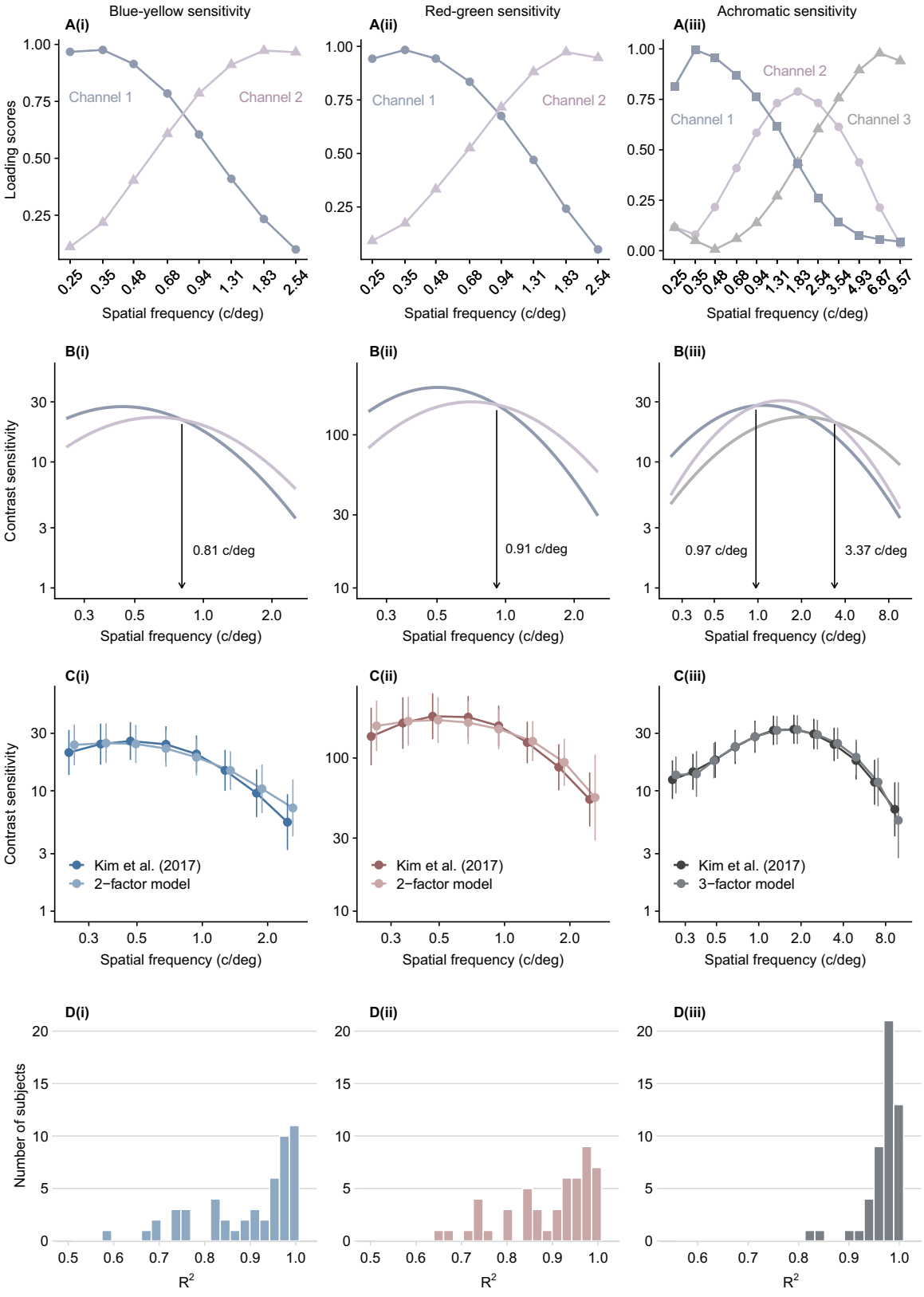

Figure 6. Evaluation of each model against each class of empirical data. The first column represents the blue-yellow data and its two-factor model. The second column shows the red-green data and its two-factor model. The third column shows the achromatic data and its three-factor model. (**A**) Loading scores of each model. Each plot represents each factor that is included in the factor model. (**B**) Contrast sensitivity of each spatial channel based on the equation from Peterzell and Teller.[15] (**C**) Model fit against the empirical data. All these models capture the true, asymmetrical shape of the CSF. *Error bars* represent standard deviations across subjects. (**D**) Evaluation of the model based on R^2. The histograms show that the current models can reliably approximate the data for each subject.

Investigative Ophthalmology & Visual Science

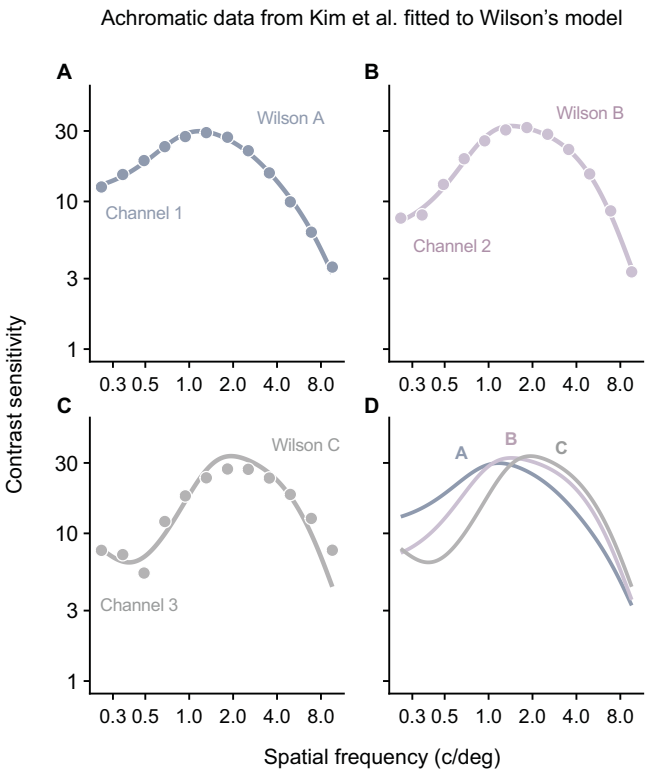

Figure 7. Achromatic data from the study of Kim et al.[35] and each factor's loadings were used to compute the tuning of each statistical factor, which was then fitted to Wilson's model.[45] The points represent extrapolated tuning of each factor (see Methods for details). The solid lines represent the fitted Wilson's model, specifically the first three channels (**A, B,** and **C**) of the model.[45] (**A**) The extrapolated tuning function (*points*) of the first factor was fitted against the Wilson's model, yielding a Wilson A channel (*solid line*). (**B**) The extrapolated tuning function (points) of the second factor was fitted against the Wilson's model, yielding a Wilson B channel (*solid line*). (**C**) The extrapolated tuning function (*points*) of the third factor was fitted against the Wilson's model, yielding a Wilson C channel (*solid line*). (**D**) Wilson's three spatial channels are plotted together.

intersection between the two spatial channels is at about 3.11 c/deg. The distribution of R^2 (mostly above 0.8) also indicates that the model-fit is appropriate (Fig. 5C). However, there remains a possibility that a third spatial channel might operate. To address this issue, we included a third factor in the model. However, it contributed to only 4% of total variance, whereas the two primary factors explained 96% of variance. Therefore we can conclude that there are two frequency-tuned statistical factors beyond 1 c/deg for achromatic sensitivity.

Revisiting the Data of Kim et al.[35] After Model Selections

Now that we have established that there are most likely two statistical factors above 1 c/d for achromatic sensitivity, we can perform a confirmatory analysis of the data of Kim et al.[35] with the three channels that are necessary for achromatic sensitivity in the range 0.25 c/deg to 9.57 c/deg. For chromatic stimuli, together with the visualization and resampling methods, our results reveal that a two-factor model adequately describes the contrast sensitivity for chromatic stimulus, be it red-green or blue-yellow. The loading scores of the two spatial channels across the two chromatic classes intersect at about 0.7 to 1 c/deg [Fig. 6A(i) and (ii)]. To precisely capture the point of intersection, we plotted the tuning of each spatial channel [Fig. 6B(i) and (ii)] by fitting the data using the loading scores with the qCSF model. The intersection of spatial frequency between the two channels were 0.81 and 0.91 for blue-yellow and red-green stimuli, respectively. Indeed, the points of intersection are close to where the peaks of the sensitivity functions are shown in Figure 6C(i) and (ii). Also, the shape of the model CSF accurately fits the data shown in Figure 6C(i) and (ii). Additionally, the distribution of the coefficient of determination (R^2) of the two-factor model for these two stimuli types also shows that the model faithfully describes the data shown in Figure 6D(i) and (ii). Because the range of 0.25 c/deg and 2.54 c/deg encompasses entire shape of the CSF, there might be no need for another one beyond the spatial two channels. Therefore our results show that the likelihood ratio test, which supported the three-factor model, overestimated the number of factors in our chromatic data. For the blue-yellow data, the fit from one-factor model (Fig. 3) and the tuning of the second channel [pink line in Fig. 6B(i)] have similar shapes and ranges of sensitivity, indicating that one-factor model from Figure 3 might represent the second spatial channel. However, for the red-green data, the fit from one-factor model (Fig. 3) has a higher range of sensitivity than the second spatial channel [pink line in Fig. 6B(ii)], suggesting that the one-factor model was merely a coarse attempt of the factor analysis to summarize the original data rather than a direct representation of the second spatial channel.

With regard to the achromatic data, the loading scores of the three-factor model for the achromatic data show that there is one spatial channel prior to the peak of the CSF, one at the peak, and one beyond the peak [Figs. 6A(iii) and 6C(iii)]. In fact, the intersections of the loading scores represent the boundaries of the three spatial channels, which were found to be at 0.97 and 3.37 c/deg [Fig. 6B(iii)]. Interestingly, the first intersection point near 1 c/d is also where the achromatic peak begins, whereas the second intersection point after 3 c/d represents the starting point of the function's decline. Also, the intersection of the second and third spatial channels is similarly located to that of the two channels from Reynaud et al.[34] (see Fig. 5). Also, the histogram of R^2 indicates that the three-factor model has an excellent model-fit to the empirical data [Fig. 6D(iii)]. Together, these findings suggest that there are three spatial-frequency-tuned statistical factors in the range of 0.25 c/deg to 9.54 c/deg for achromatic contrast sensitivity.

Finally, we compared our statistical factors to the channels that were determined behaviorally in a masking experiment by Wilson and Gelb.[45] We fitted the derived sensitivity values from each statistical factor against Wilson's spatial model, which assumes that there are separate mechanisms for processing contrast sensitivity across different ranges of spatial frequency (Fig. 7). The lowest spatial frequency range is represented by the Wilson A channel (Fig. 7A). The mid and higher ranges are represented by Wilson B and C channels, respectively (Figs. 7B, 7C). The agreement between our extrapolated sensitivity from each statistical factor and the Wilson channels demonstrate that Wilson's model can reliably capture the tuning functions of the three statistical factors. The three fits resemble faithfully to those in previous studies[11] (compare Fig. 7D vs. Fig. 3 in Peterzell's report[11]). For instance, the peak of Wilson A channel is

close to 1 c/deg, whereas that of Wilson B channel is above 1 c/deg. Our findings from factor analysis also support Wilson's model and earlier studies[11,44] that describe contrast sensitivity data using covariance.

Discussion

Earlier, we alluded to the fact that factor analytic results could be different depending on the methodologies of the analysis method and data collection, as well as differences in local samples. Here, we demonstrated with our data that results from factor analysis using empirical data can be equivocal because different statistical tests and methods of model selections supported different models, requiring us to apply a resampling method in the exploratory approach and visualization in the confirmatory measure. The reason why exploratory measures were not able to collectively support a single model could be due to the limited number of response variables in our dataset, such as spatial frequency. In fact, a latent model from exploratory factor analysis can be tested using confirmatory factor analysis, which requires a sample size of at least 200.[27,32] Confirmatory factor analysis is often used in psychometrics that rely on questionnaires for data analysis, where each question can be considered a response variable and a cluster of questions can be part of a latent factor.[32] In such a case, there can be as many as 50 or more questions with a sample size above 300. This is impractical for psychophysical experiments, along with the fact that data at each response variable could be a summary measure of many trials.[46] However, as opposed to in psychometrics, the model from psychophysical studies can be fitted against experimental data visually, enabling us to estimate the number of factors clearly.

Our results from factor analysis using resampling and visualization methods revealed that two spatial channels were necessary to model blue-yellow and red-green sensitivity, whereas three channels were required to model achromatic sensitivity. Also, they showed that blue-yellow, red-green and achromatic channels were independent. Indeed, our results agree to findings of previous studies.[15,19] The number of channels was found to be different between achromatic and chromatic vision for two reasons. First, the shapes of achromatic and chromatic data are different. The chromatic sensitivity function has the lowpass shape, peaking (0.5 c/deg) and falling off at a lower spatial frequency,[47] whereas the achromatic sensitivity function has a bandpass shape, peaking at about 2-3 c/deg and falling off at a higher frequency.[47,48] Second, the range of tested spatial frequency for the chromatic sensitivity was narrower than that of the achromatic sensitivity in our dataset, requiring us to consider the possibility that a third spatial channel might still exist for chromatic sensitivity at beyond 3 c/deg (see Fig. 6). However, in light of the difference in the shapes of the chromatic (lowpass) and achromatic (bandpass) sensitivity functions, it is possible that the minimal sensitivity levels at higher spatial frequencies (beyond 4 c/d) for chromatic stimuli might minimize the role of the potential third spatial channel even if it operates, reducing its significance and impact on visual perception.

The shape of the CSF when there was an inadequate number of spatial channels in the factor model was inaccurate (i.e., roughly symmetrical) in all classes of visual stimuli along the logarithmic axis of spatial frequency, failing to approximate the original data. Conversely, the shape of the fitted curve when there was an adequate number of channels in the factor model was similar to that of the empirical data with an asymmetrical and inverted-U shape. For chromatic vision (see Fig. 6), the fits of inadequate models could be a reflection of one spatial channel that is more tuned for higher spatial frequency. However, for achromatic vision, we found that the fit from a one-factor model [gray line in Fig. 3C(i)] was not representative of a single spatial channel. This could because it was a coarse attempt of factor analysis to reduce the original data into one factor rather than to capture a single spatial channel, demonstrating that at times statistical factor is not equivalent to spatial channel.

The loading scores and the tuning functions (Figs. 6, 7) provide additional insights about spatial vision. For chromatic sensitivity, the intersections of the loadings (and the tuning functions) from the factor models were near the peak of the CSFs. This finding indicates that two spatial channels were active at the peak. It supports the combination rule, which assumes that multiple spatial channels combine to determine the shape of the contrast sensitivity function.[12,13,15] For achromatic sensitivity, the loadings from the three-factor model intersected at about 1 and 3 c/deg [see Fig. 6 B(iii)], surrounding the achromatic peak. Following the same interpretation, we can speculate that three spatial channels were active near the peak between 1 and 3 c/deg. Interestingly, the first two channels from our achromatic three-factor model are similarly positioned as Wilson A and B channels from Peterzell et al.,[11,14,15,19,20,44,49] despite some differences in where the spatial channels intersect, which can be attributed to sample-to-sample variability. Interestingly, the tunings of the three statistical factors were faithfully captured by Wilson's model,[45] which yielded three separate channels (see Fig. 7). The third factor is similarly positioned to the covariance factor driven by optics at a higher spatial frequency.[17] Furthermore, two lowest achromatic channels (blue and pink lines in Fig. 6B) show similar tunings to those of the two chromatic channels. This similarity in their tuning does not necessarily support an interdependence of these channels[15] because we observed no dependent relationships between them. Instead, it would suggest that vision of luminance modulations (i.e., achromatic contrast) can simply benefit from an additional high spatial frequency channel compared to color vision, thereby supporting the classical view that achromatic vision processes finer details.[50] Conclusions from our factor analysis confirm the findings from primate studies, namely that achromatic and chromatic vision are processed independently[50–55] and that there is a form of multiplexing of retinal ganglion cells resulting in the dependency of spatial frequency[56] that results in similar tunings of the achromatic and chromatic channels. Midget ganglion cells relay color information and spatial resolution to the parvocellular layers of the lateral geniculate nucleus. For example, if the parvocellular pathway is perturbed, color sensitivity and achromatic sensitivity at high spatial frequency are compromised.[50,51]

We realized that there has been a shortage of computational tools for vision research and ophthalmology. For this reason, we created the *smCSF* package, which is for researchers and clinicians who wish to plot elegant contrast sensitivity functions and perform standardized data analysis routines that have been used in the last decade with minimal coding (code examples and documentation in https://smin95.github.io/dataviz: Chapters 13–16). It also

allows the user to easily compute the qCSF parameters for all subjects across experimental conditions, groups and repetitions in a single line of code if the data frame has an appropriate structure. In addition, the package provides functions for fitting the sensitivity curve from raw data; examples are shown in Chapters 13 and 14 of the documentation. Finally, it provides some additional methods for performing factor analysis (Chapter 16). To our knowledge, this library is the first R package for analyzing and visualizing contrast sensitivity data. It has been built to simplify the process of data visualization and analysis of contrast sensitivity data within a single software environment. We hope that the methods and the tool we have introduced will be useful for the researchers and clinicians in the fields of vision science and ophthalmology.

Acknowledgments

The authors thank Fang Hou for providing his insights on modeling psychophysical data. His comments greatly facilitated the development of the R package *smCSF*.

Supported by a National Natural Science Foundation of China grant (32350410414) and a National Foreign Expert Project fund (no. QN2022016002L) to S.M., and a Starting Fund from the Research Institute of the McGill University Health Center to A.R.

Disclosure: **S.H. Min**, None; **A. Reynaud**, None

References

1. Spearman C. "General intelligence" objectively determined and measured. In: Jenkins JJ, Paterson DG, eds. *Studies in Individual Differences: The Search for Intelligence*. Norwalk, CT: Appleton-Century-Crofts; 1961:59–73.
2. Campbell F. Why do we measure contrast sensitivity? *Behav Brain Res*. 1983;10:87–97.
3. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*. 1962;160:106.
4. Campbell FW, Robson JG. Application of Fourier analysis to the visibility of gratings. *J Physiol*. 1968;197:551.
5. De Valois RL, De Valois KK. Spatial vision. *Annual review of psychology*. 1980;31(1):309–341.
6. Graham NVS. *Visual Pattern Analyzers*. Oxford, UK: Oxford University Press; 1989.
7. Greenlee MW, Magnussen S, Nordby K. Spatial vision of the achromat: spatial frequency and orientation-specific adaptation. *J Physiol*. 1988;395:661–678.
8. Tolhurst D. Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *J Physiol*. 1973;231:385.
9. Stromeyer III C, Klein S, Dawson B, Spillmann L. Low spatial-frequency channels in human vision: adaptation and masking. *Vis Res*. 1982;22:225–233.
10. Watson AB, Robson JG. Discrimination at threshold: labelled detectors in human vision. *Vis Res*. 1981;21:1115–1122.
11. Peterzell DH. Discovering sensory processes using individual differences: a review and factor analytic manifesto. *Electronic Imaging*. 2016;28:1–11.
12. Peterzell DH, Chang SK, Teller DY. Spatial frequency tuned covariance channels for red–green and luminance-modulated gratings: psychophysical data from human infants. *Vis Res*. 2000;40:431–444.
13. Sekuler R, Wilson HR, Owsley C. Structural modeling of spatial vision. *Vis Res*. 1984;24:689–700.
14. Peterzell DH, Teller DY. Individual differences in contrast sensitivity functions: the lowest spatial frequency channels. *Vis Res*. 1996;36:3077–3085.
15. Peterzell DH, Teller DY. Spatial frequency tuned covariance channels for red–green and luminance-modulated gratings: psychophysical data from human adults. *Vis Res*. 2000;40:417–430.
16. Dobkins KR, Gunther KL, Peterzell DH. What covariance mechanisms underlie green/red equiluminance, luminance contrast sensitivity and chromatic (green/red) contrast sensitivity? *Vis Res*. 2000;40:613–628.
17. Elliott S, Peterzell D. Individual differences in contrast sensitivity functions with and without adaptive optics: direct estimates of optical and neural processes in young and elderly adults using factor analysis. *J Vis*. 2017;17:791–791.
18. Peterzell D, Schefrin B, Tregear S, Werner JS. Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In: *Vision Science and Its Applications*. Washington, DC: Optica Publishing Group; 2000:FC2.
19. Peterzell DH, Werner JS, Kaplan PS. Individual differences in contrast sensitivity functions: the first four months of life in humans. *Vis Res*. 1993;33:381–396.
20. Peterzell DH, Kelly JP. Spatial frequency channels revealed by individual differences in contrast sensitivity functions: Visual evoked potentials from adults and infants. In: *Vision Science and Its Applications*. Washington, DC: Optica Publishing Group; 1996:ThA-3.
21. Wilson HR. The perception of form: retina to striate cortex. In: Spillmann L, Werner JS, eds. *Visual Perception: The Neurophysiological Foundations*. Philadelphia: Elsevier; 1990:231–272.
22. Wilson H. Theories of infant visual development. In: *Early Visual Development: Normal And Abnormal*. New York: Oxford University Press; 1993:560–572.
23. Ng CJ, Sabesan R, Barbot A, Banks MS, Yoon G. Suprathreshold contrast perception is altered by long-term adaptation to habitual optical blur. *Invest Ophthalmol Vis Sci*. 2022;63(11):6–6.
24. Reynaud A, Min SH. Spatial frequency channels depend on stimulus bandwidth in normal and amblyopic vision: an exploratory factor analysis. *Front Comput Neurosci*. 2023;17:1241455.
25. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organ Res Methods*. 2004;7:191–205.
26. Warton DI. *Eco-Stats: Data Analysis in Ecology: From t-Tests to Multivariate Abundances*. Berlin: Springer Nature; 2022.
27. Baker DH. *Research Methods Using R: Advanced Data Analysis in the Behavioural and Biological Sciences*. Oxford, UK: Oxford University Press; 2022.
28. Webster MA, MacLeod DI. Factors underlying individual differences in the color matches of normal observers. *JOSA A*. 1988;5:1722–1735.
29. Costello AB, Osborne J. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract Assess Res Eval*. 2005;10(1):7.
30. Hayashi K, Bentler PM, Yuan KH. On the likelihood ratio test for the number of factors in exploratory factor analysis. *Struct Equation Model*. 2007;14:505–526.
31. Warne RT, Larsen R. Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory factor analysis. *Psychol Test Assess Model*. 2014;56:104.
32. Kyriazos TA. Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*. 2018;9:2207.
33. Mooney CZ, Duval RD, Duvall R. *Bootstrapping: A Nonparametric Approach to Statistical Inference*. Thousand Oaks, CA: Sage Publishing; 1993.

Investigative Ophthalmology & Visual Science

34. Reynaud A, Tang Y, Zhou Y, Hess RF. A normative framework for the study of second-order sensitivity in vision. *J Vis*. 2014;14(9):3–3.
35. Kim YJ, Reynaud A, Hess RF, Mullen KT. A normative data set for the clinical assessment of achromatic and chromatic contrast sensitivity using a qCSF approach. *Invest Ophthalmol Vis Sci*. 2017;58:3628–3636.
36. Lesmes LA, Lu ZL, Baek J, Albright TD. Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method. *J Vis*. 2010;10(3):17–17.
37. Hou F, Huang CB, Lesmes L, et al. qCSF in clinical application: efficient characterization and classification of contrast sensitivity functions in amblyopia. *Invest Ophthalmol Vis Sci*. 2010;51:5365–5377.
38. Revelle W, Revelle MW. Package "psych." *The comprehensive R archive network*. 2015;337(338).
39. Gorsuch RL. *Factor Analysis: Classic Edition*. Routledge; 2014.
40. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychol Methods*. 1999;4(3):272.
41. Ford JK, MacCallum RC, Tait M. The application of exploratory factor analysis in applied psychology: a critical review and analysis. *Personnel Psychol*. 1986;39:291–314.
42. Reynaud A, Takerkart S, Masson GS, Chavane F. Linear model decomposition for voltage-sensitive dye imaging signals: application in awake behaving monkey. *Neuroimage*. 2011;54:1196–1210.
43. Friston KJ, Holmes AP, Worsley KJ, Poline J, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*. 1994;2(4):189–210.
44. Peterzell DH, Kelly JP. Development of spatial frequency tuned "covariance" channels: Individual differences in the electrophysiological (VEP) contrast sensitivity function. *Optom Vis Sci*. 1997;74:800–807.
45. Wilson HR, Gelb DJ. Modified line-element theory for spatial-frequency and width discrimination. *JOSA A*. 1984;1:124–131.
46. Prins N. *Psychophysics: A Practical Introduction*. New York: Academic Press; 2016.
47. Mullen KT. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *J Physiol*. 1985;359:381–400.
48. Kelly D. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *JOSA*. 1983;73:742–750.
49. Peterzell DH, Werner JS, Kaplan PS. Individual differences in contrast sensitivity functions: longitudinal study of 4-, 6- and 8-month-old human infants. *Vis Res*. 1995;35:961–979.
50. Merigan WH. Chromatic and achromatic vision of macaques: role of the P pathway. *J Neurosci*. 1989;9:776–783.
51. Schiller PH, Logothetis NK, Charles ER. Functions of the colour-opponent and broad-band channels of the visual system. *Nature*. 1990;343(6253):68–70.
52. Shapley R. Visual sensitivity and parallel retinocortical channels. *Ann Rev Psychol*. 1990;41:635–658.
53. Wiesel TN, Hubel DH. Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *J Neurophysiol*. 1966;29:1115–1156.
54. Kaplan E, Shapley R. X and Y cells in the lateral geniculate nucleus of macaque monkeys. *J Physiol*. 1982;330:125–143.
55. Derrington AM, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *J Physiol*. 1984;357:241–265.
56. Wool LE, Crook JD, Troy JB, Packer OS, Zaidi Q, Dacey DM. Nonselective wiring accounts for red-green opponency in midget ganglion cells of the primate retina. *J Neurosci*. 2018;38:1520–1540.
57. Wickham H. Data analysis. In: *Ggplot2*. Berlin: Springer; 2016:189–201.

Appendix

qCSF Parameters and *smCSF*

The *smCSF* package fits the CSF based on the maximum likelihood estimates of key parameters of the truncated log-parabola model for human contrast sensitivity data.[35–37] The mathematical model predicts the shape of the CSF, using the equations below:

$$S'(f) = log_{10}(\gamma_{max}) - \kappa \left(\frac{log_{10}(f) - log_{10}(f_{max})}{\beta'/2} \right)^2,$$

$$\text{where } \kappa = log_{10}(2) \text{ and } \beta' = log_{10}(2\beta)$$

$$S(f) = log_{10}(\gamma_{max}) - \delta, \text{ if } f < f_{max}$$

$$\text{and } S'(f) < log_{10}(\gamma_{max}) - \delta$$

$$S(f) = S'(f) \text{ when } f > f_{max}.$$

Various functions from the package, such as *sm_CSF()*, *sm_areaCSF()* and *sm_ribbonCSF()*, can plot the best-fitted model automatically in log scales using a data frame that contains experimental data with R implementations of the equations above. These functions estimate the parameters based on default arguments and fit a best possible model based on maximum likelihood. They have been created as direct extensions of *ggplot2* via *ggproto* objects.[57] The parameters specify the overall characteristic of the sensitivity function, rather than a data point at a specific frequency (Fig. A1).

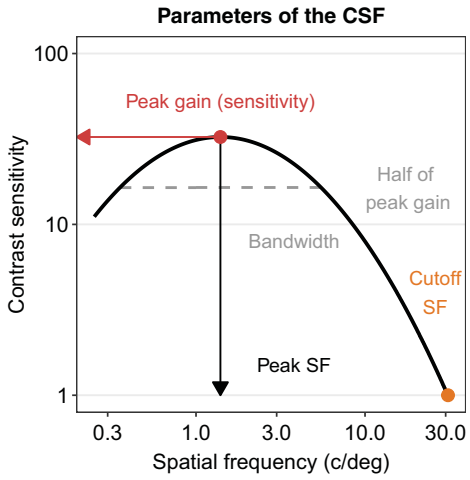

Figure A1. Visual representation of the parameters that define the overall characteristic of the sensitivity across a range of spatial frequency.

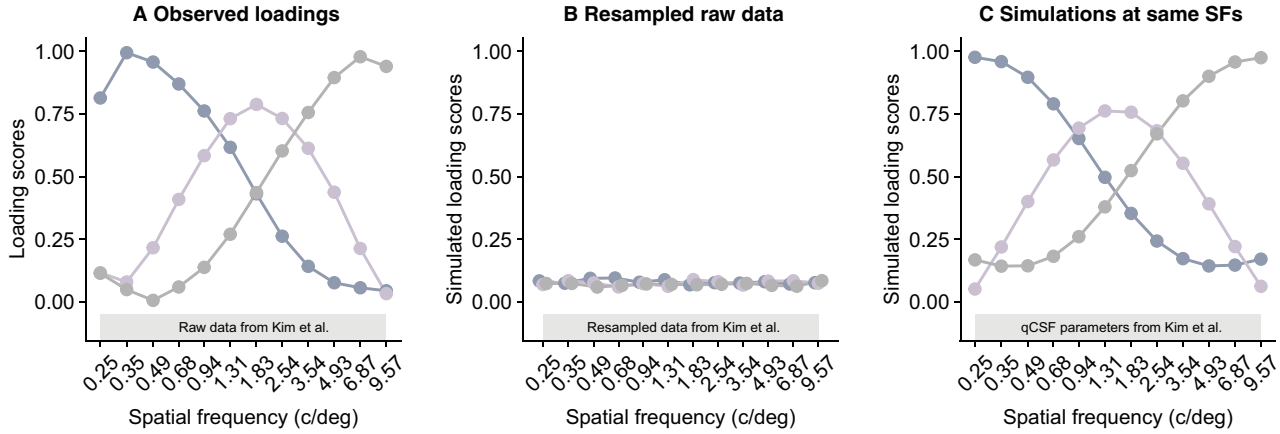

Figure A2. Loading scores from raw data and Monte Carlo simulations. (**A**) Loading scores based on the achromatic sensitivity data (0.25–9.57 c/deg) from Kim et al.[35] (**B**) Mean loading scores based on the resampled data from 1000 simulations. (**C**) Mean loading scores from 1000 simulations based on the factor analysis of the extrapolated data computed from the resampled qCSF parameters at the same tested spatial frequencies.

1. Peak gain (γ_{max}): Peak contrast sensitivity (hence, it is in the y-axis unit).
2. Peak spatial frequency (f_{max}): Spatial frequency where the peak gain is located (hence, it is in the x-axis unit).
3. Bandwidth (β): Width of the CSF at half of the peak gain. The larger the bandwidth, the higher the overall sensitivity across the region of high spatial frequencies.
4. Truncation value (δ): This is a parameter creates a narrow plateau (i.e., flat curve) at the low spatial frequency range and resolves the issue of the CSF's asymmetry. In this model, however, the factor plays a very minor role and the plateau is usually unnoticeable.

The four parameters are often used to evaluate the status of the visual system directly from the contrast sensitivity data. These parameters can be computed across all subjects, groups, and experimental conditions with a single line of code using sm_param_list(). More information about the package is uploaded online (https://smin95.github.io/dataviz: Chapters 13–16), where it will be continuously updated.

Resampling Method for Factor Analysis

It turns out that resampling the sensitivity data themselves does not retain the variability of the data from each observer. This is because it would be analogous to resampling "parts" (from one spatial frequency to another) of the contrast sensitivity function (CSF), rather than its whole, which retains the partial covariance. Breaking it up into "parts" and re-picking them with replacement across observers would remove the partial covariance that can only be found in the whole function. This would defeat the purpose of factor analysis, which aims to identify a factor that causes data to covary with some data but not others. Covariance refers to the joint variability of response variables (ex. sensitivity). Therefore, to retain the partial relationship among variables, we fitted the data with the log truncated model of qCSF,[36] and estimated the model parameters using maximum likelihood. Then, we resampled the parameters with replacement to create a simulated dataset in the form of a full CSF (note that parameters were considered being independent), and computed the eigenvalues with factor analysis. Resampling was performed independently because it is more robust to potential outliers. These steps would be one iteration, and 1000 iterations were performed in total. If we resampled the sensitivity data from multiple observers to create a simulated dataset, the resampled data might still produce a normal CSF with a peak and trough, but it would not retain covariance of the data. A computational demonstration has been posted online (https://smin95.com/dataviz: Chapter 16). Also, factor analytic results from Monte Carlo simulations are shown above to illustrate this point.

To explore how different sampling methods affect factor loadings, we performed Monte Carlo simulations. The loadings from the original raw data are shown in the panel A of Figure A2. If the data were directly resampled, the covariance disappeared, and loading scores from factor analysis reached near 0 (i.e., no correlation or covariance; panel B of Fig. A2). However, if factor analysis was performed using recomputed data from resampled qCSF parameters at the same spatial frequencies as those from the raw data, then the loadings had similar magnitudes, tuning, and intersection points

compared to those from the original data (panel C of Fig. A2).

NonParametric Resampling Methods With *smCSF*

The code fragments below extract the four parameters from the psychophysical data, resample them with replacement independently at a certain sample size, generate the fitted values of sensitivity from the resampled parameters, and then perform factor analysis to compute the eigenvalue for each iteration. The eigenvalues are then summarized as means and upper and lower ends of the 95% confidence intervals. In each code, > is the prompt of R and + denotes a continuation from the previous line. The bolded codes are related to *smCSF*.

```
> library(tidyverse)
> library(smCSF) # devtools::
  install_github('smin95/smCSF')
> library(smplot2) #
  devtools::install_github('smin95/
  smplot2')
```

First, the three libraries have to be loaded to memory. The installation codes can return error if the version of the *devtools* package is not up to date; the user should redownload the *devtools* package if the installation of *smCSF* fails.

```
> data_frame <- read_csv('https://
www.smin95.com/data_ACh.csv') %>%
+ group_by(Subject, SpatialFreq)
  %>%
+ summarise(Sensitivity =
  mean(Sensitivity),
+ Repetition = 'avg')
```

Next, the sample data should be loaded to memory from online. It contains data for achromatic sensitivity. There are two repetitions, and these are averaged. The final output is stored in the variable data_frame.

```
> param_res <- sm_params_list
  (subjects = 'Subject',
+ conditions = 'Repetition',
+ x = 'SpatialFreq', values =
  'Sensitivity', data = data_frame)
```

sm_params_list() from the *smCSF* package allows the users to compute the five parameters for all subjects across different conditions and/or groups with a single line of code. It has several arguments. It returns a list of important outputs. The user must identify the column that has all the subject identifiers within the subjects argument, as well as for the column with identifiers for the conditions in the conditions argument. Likewise, the argument x is for the column with spatial frequencies, values for the column with the contrast sensitivity data (in linear units) and data for the data frame itself that contains the linear data of spatial frequency and contrast sensitivity. Its output is a list with two vectors. The list from sm_params_list() can be used as argument for the function sm_np_boot(), which performs nonparametric simulation by resampling the parameters with replacement and refitting the CSF using the parameters. The output is saved in the variable param_res, which is short for *parameter results*.

```
> nSim <- 1000 # 1000 iterations
> nObs <- 51 # 51 observers in the
  dataset
> boot_res <- sm_np_boot(param_res,
  n = nObs, nSim = nSim)
```

The function sm_np_boot() returns a data frame of the averaged eigenvalues from the simulations as well as their 95% confidence intervals of the resampled and randomly generated data (from parallel analysis) of the CSF parameters. The number of sampled observer (n) should be identical to the observed sample size (nObs) because of the nature of nonparametric bootstrapping. The number of simulations can be specified by using nSim argument, which in this case has been set to 1000. The output from the function can then be directly plotted below (i.e., similar to Fig. 4) in the form of a scree plot with error bars using sm_plot_boot().

```
> sm_plot_boot(boot_res, shapes =
  16) + ylab('Mean eigenvalues')
```

Based on this analysis, the number of retained factors should be the ones that do not have their error bars overlapped with those of the randomly generated data from parallel analysis. However, if the reader is interested in computing the uncertainty of eigenvalues using another model, it is also possible by resampling the relevant parameters of an appropriate model, which can be as simple as a simple regression model with its slope and intercept. Then using the resampled parameters, users can generate the fitted values, and perform factor analysis using the fitted data to compute the eigenvalue from each iteration.