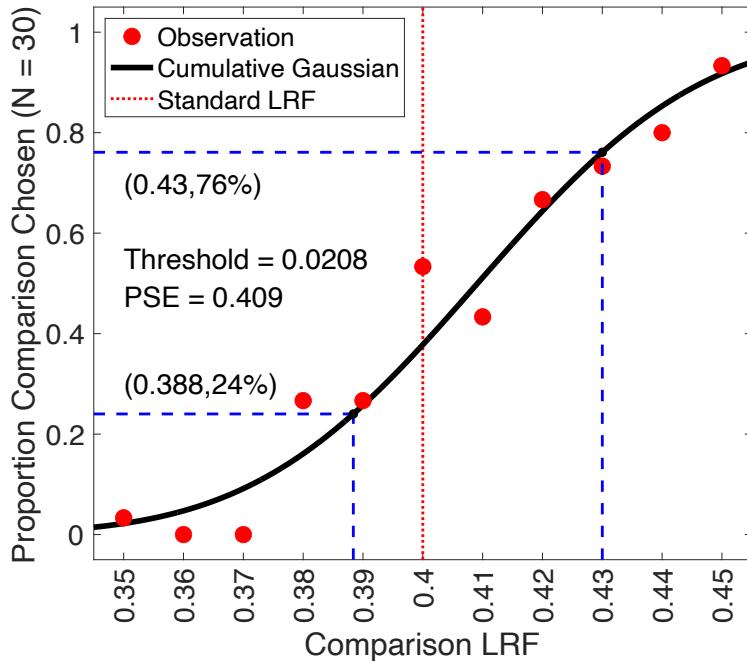


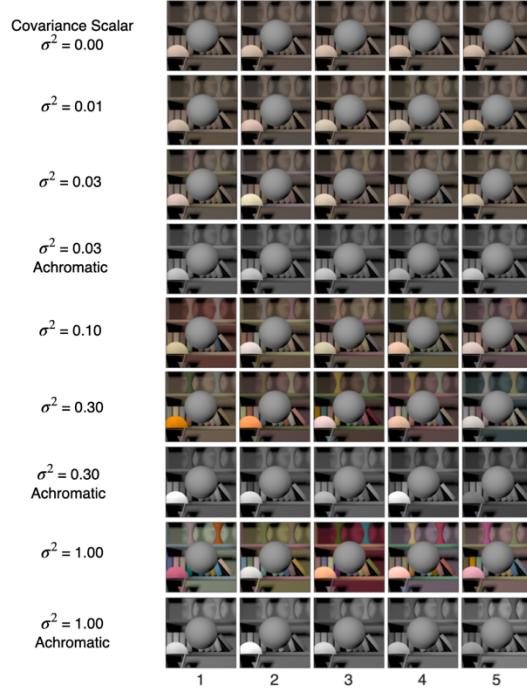
1
2 **Figure 1: (a) Psychophysical task.** (Adapted from Figure 1 in Singh, Burge, & Brainard, 2022) The
3 psychophysical task involved comparing two images, a standard image and a comparison image, on each
4 trial and selecting the image with the lighter target object. The target object was an achromatic sphere at
5 the center of the image. The images were generated computationally by graphically rendering models of
6 3D scenes. They were displayed on a color calibrated monitor. This panel shows examples of standard
7 and comparison images. The reflectance spectrum of the target object was spectrally flat, and the target
8 object appeared gray. The reflectance of the target object in the standard image was held fixed and it
9 changed for the comparison image. In this panel, the target object in the comparison image is lighter. We
10 measured the fraction of times the observers chose the target object in the comparison image to be lighter
11 as a function of the lightness of the target object in the comparison image. Fraction comparison chosen
12 data was used to determine lightness discrimination threshold (Figure 2). We studied how the lightness
13 discrimination thresholds changed as the trial-to-trial variability in the reflectance spectra of the
14 background objects and the intensity of the light sources increased. **(b) Trial sequence:** R_{N-1} indicates the
15 recording of the observer's response for the $(N-1)^{th}$ trial. The N^{th} trial begins 250ms after the completion
16 of the $(N-1)^{th}$ trial (Inter Trial Interval, ITI = 250ms). In the N^{th} trial, the standard and comparison images
17 are presented for 250ms each with a 250ms inter stimulus interval (ISI) in between the two images. The
18 order of the standard and comparison images is chosen in pseudorandom order. The observer records their
19 choice by pressing a button on a gamepad after both images have been presented and removed from the
20 screen. The observers could take as long as they wish before making their choice. The recording of their
21 choice is indicated by R_N in the panel. The next trial begins 250ms after the choice has been recorded.



1

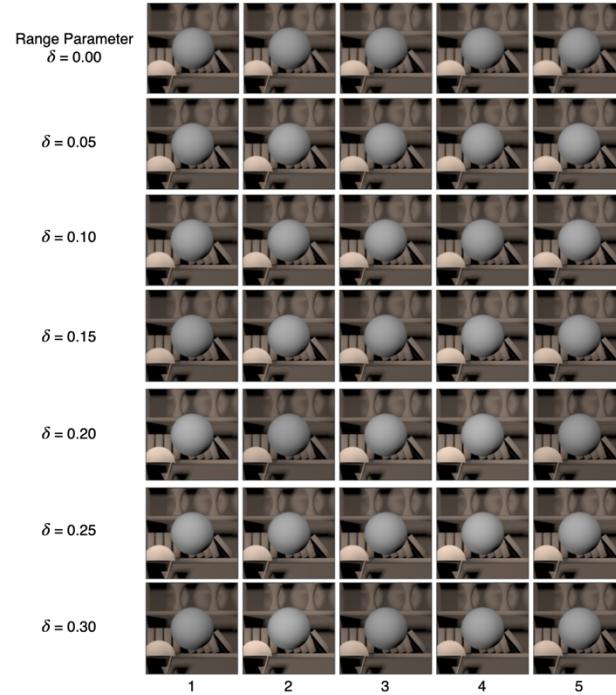
2 **Figure 2: Psychometric function:** We measured the proportion of times the observers selected the target
 3 in the comparison image to be lighter as function of the LRF (lightness reflectance factor) of the target
 4 object. We collected 30 responses for each of the 11 equally spaced values of the comparison image target
 5 object LRF, ranging from 0.35 to 0.45. The target object in the standard image had an LRF of 0.40. The
 6 LRF of the target object in the comparison image was selected in a pseudorandom order. To analyze the
 7 data, we used maximum likelihood methods to fit a cumulative normal function to the proportion
 8 comparison chosen data. We imposed constraints on the guess rate and lapse rate, requiring them to be
 9 equal and within the range of 0 to 0.05. The threshold was determined as the difference between the LRF
 10 values corresponding to a proportion comparison chosen of 0.76 and 0.50, obtained from the cumulative
 11 normal fit. The figure presented here illustrates the data for observer 0003 in the second block of the
 12 background reflectance variation experiment (previously registered as Experiment 6) for the no variation
 13 ($\sigma^2 = 0.00, \delta = 0.00$) condition. The discrimination threshold was measured to be 0.0208. The point of
 14 subjective equality (PSE), which corresponds to a proportion of 0.5 in the comparison task, was found to
 15 be 0.409. The lapse rate for this particular fit was determined to be 0.00.
 16

1



2 **Figure 3: Background object reflectance variation:** We studied two types of variations in the
3 reflectance spectra of background objects in the scene: chromatic variation and achromatic variation. In
4 chromatic variation, the reflectance spectra could take any shape, and the objects varied in their
5 luminance and chromaticity. In achromatic variation, the reflectance spectra were spectrally flat, and the
6 objects appeared gray and varied only in their luminance. The spectra were chosen from a multivariate
7 normal distribution that modeled the statistics of natural reflectance spectra. The covariance matrix of the
8 multivariate normal distribution was multiplied by a scalar to control the variance in the samples. We
9 generated images at six logarithmically spaced values of the covariance scalar for chromatic variation and
10 at three values of the covariance scalar for achromatic variations. The figure shows five typical images for
11 each of these nine conditions. For each condition we generated 1100 images, 100 images at 11 linearly
12 spaced value of target object LRF in the range [0.35, 0.45]. The target object in each image in the figure is
13 at LRF = 0.4.

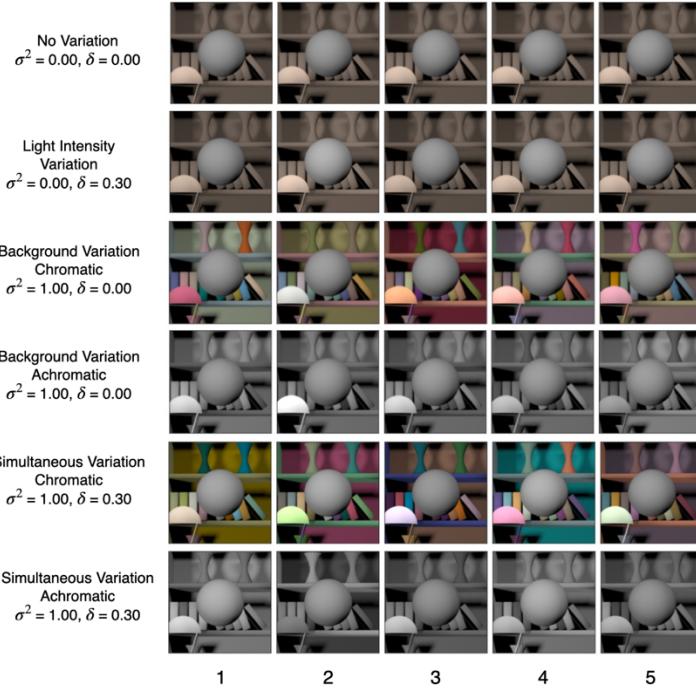
1



2 **Figure 4: Light intensity variation:** The shape of the power spectrum of the light sources in the scene
 3 was chosen to be CIE reference illuminant D65. The intensity of the power spectrum was varied by
 4 multiplying the normalized D65 spectrum with a scalar sampled from a log uniform distribution in the
 5 range $[1 - \delta, 1 + \delta]$. The amount of variation was controlled by changing the value of the range parameter
 6 δ . We generated images at seven linearly spaced values of the range parameter in the range $[0.00, 0.30]$.
 7 For each value of the range parameter, we generated 1100 images, 100 images at each value of the target
 8 object LRF in the range $[0.35, 0.45]$. The figure shows five sample images at each of the seven values of
 9 the range parameter. The target object in each image in the figure has the same LRF of 0.40.

1

2 **Figure 5: Simultaneous variation:** This figure shows five sample images for the six conditions studied
 3 in preregistered Experiment 8. We generated 1100 images for each of these conditions, 100 images at
 4 each value of the target object LRF in the range [0.35, 0.45].



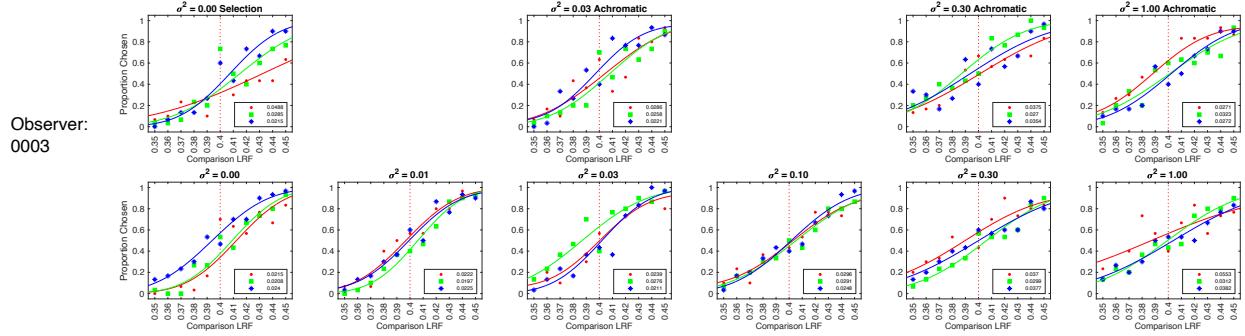
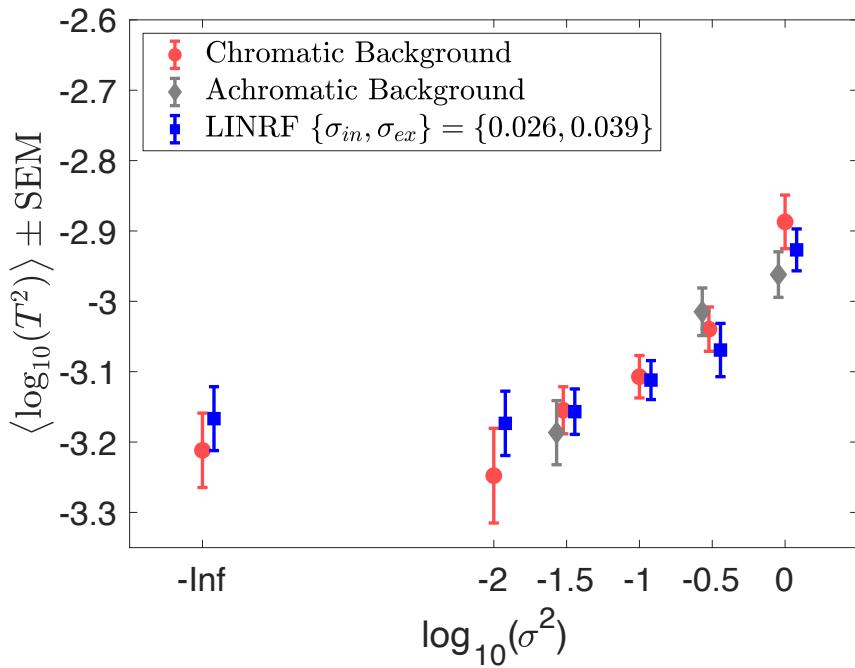


Figure 6: Psychometric functions for observer 0003 for background reflectance variation

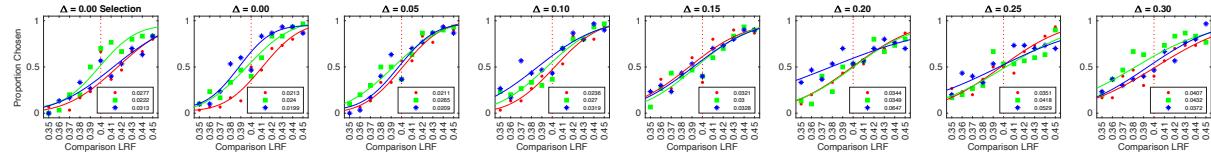
experiment: We measured the proportion comparison chosen data for the nine conditions separately in three blocks for each observer. The figure shows the psychometric function for observer 0003. The psychometric functions for all six observers are shown in Figure S2. A cumulative normal function was fit to the data from each block to determine the discrimination threshold (see Figure 2). The legend provides the estimated lightness discrimination threshold for each block, obtained from the cumulative fit. The first panel in the top row shows the data and thresholds for the selection session. The selection session was a practice session in which the thresholds for the no variation condition was measured three times. An observer was selected for the experiment only if the average of their last two discrimination threshold measurements in the selection session was less than 0.30. The last three panels in the top row show the data for the three achromatic conditions. The bottom row shows the data for the chromatic variation conditions.



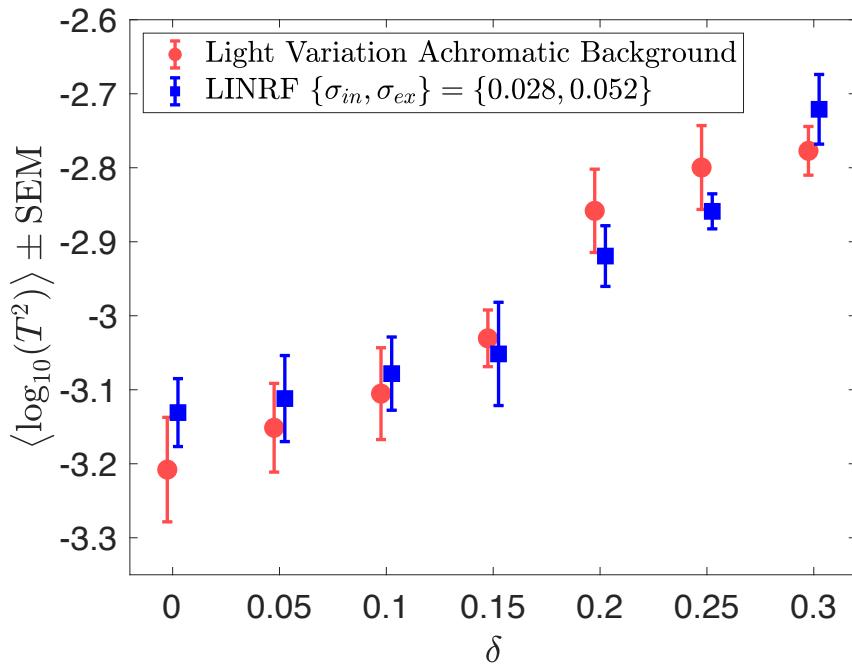
1

2 **Figure 7: Background variation increases lightness discrimination thresholds.** Mean (N = 6) log
 3 squared threshold vs log covariance scalar from human psychophysics for chromatic (red circles) and
 4 achromatic conditions (gray diamonds). The error bars represent +/- 1 SEM taken between observers. The
 5 threshold of the linear receptive field (LINRF) model was estimated by simulation for the six values of
 6 the covariance scalar (blue squares). The blue error bars show +/- 1 standard deviation estimated over 10
 7 independent simulations. The legend shows the parameters of the linear receptive field (LINRF) model
 8 fit. The data has been jittered for ease of viewing. A comparison of the thresholds with the previously
 9 published data in Singh, Burge, Brainard 2022 is shown in Figure S3.

1



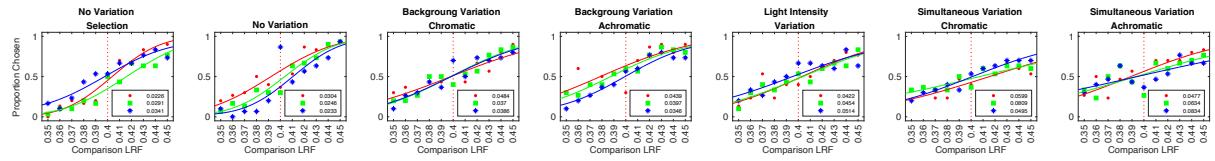
2 **Figure 8: Psychometric functions for observer 0003 for light intensity variation experiment:** Same
3 as Figure 6, but for the light intensity variation experiment. The figure shows the proportion comparison
4 chosen data for the selection session and the seven condition for observer 0003. The psychometric
5 functions for all observers are shown in Figure S4.



1

2 **Figure 9: Light source intensity variation increases lightness discrimination threshold.** Mean (N = 5)
3 log squared threshold vs range parameter from human psychophysics for the seven light source intensity
4 variation conditions (red circles). The error bars represent +/- 1 SEM taken between observers. The
5 threshold of the linear receptive field (LINRF) model was estimated by simulation for the seven values of
6 the range parameters (blue squares). The blue error bars show +/- 1 standard deviation estimated over 10
7 independent simulations. The legend shows the parameters of the LINRF model fit. The data has been
8 jittered for ease of viewing. The data for all six observers is shown in Figure S5.

1



2 **Figure 10: Psychometric functions for observer 0003 for simultaneous variation experiment:** Same
3 as Figure 6 and 8, but for simultaneous variation experiment. The figure shows the proportion comparison
4 chosen data for the selection session and the six condition for observer 0003. The data for all observers
5 are shown in Figure S6.

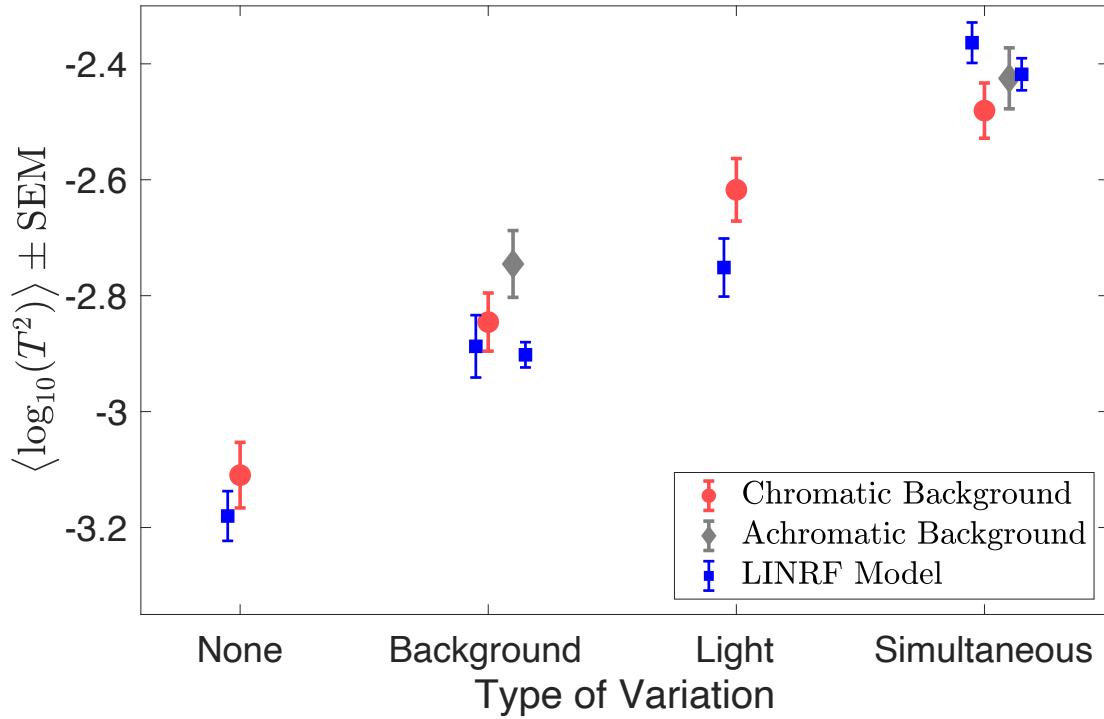
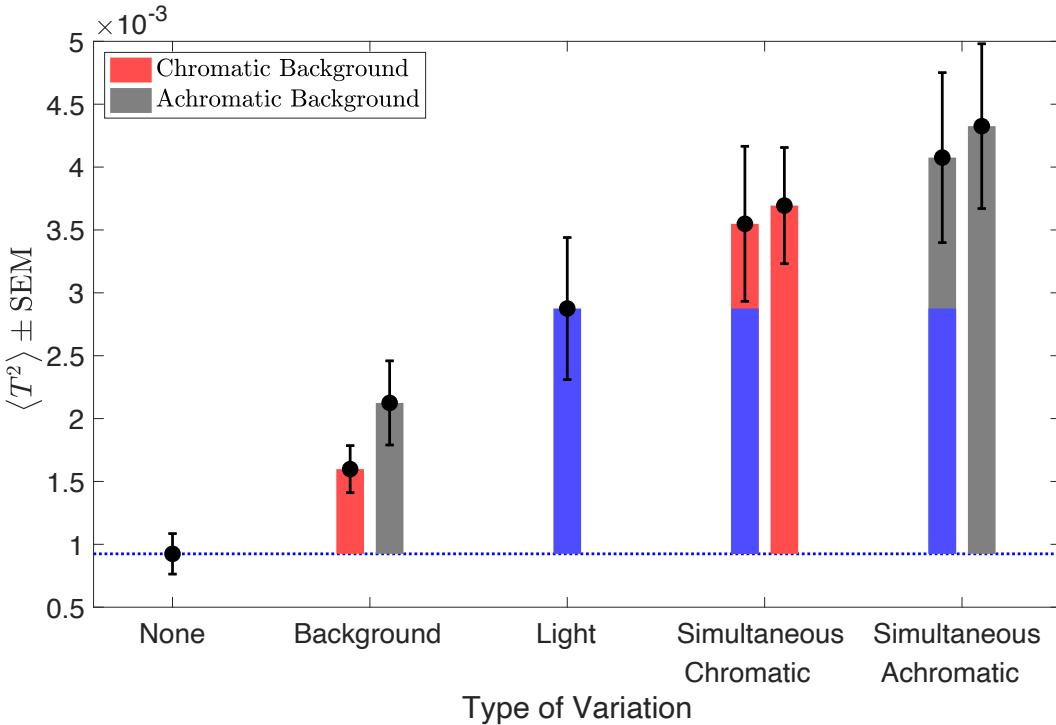
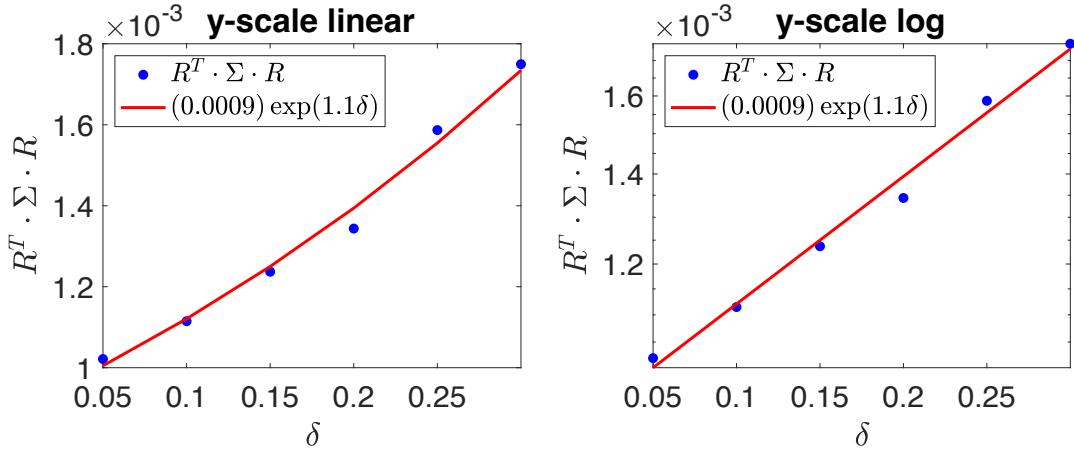


Figure 11: Discrimination thresholds for simultaneous variation of two sources are higher than individual discrimination thresholds. Mean ($N = 6$) log squared threshold for the six conditions in simultaneous variation experiment. The error bars represent ± 1 SEM taken between observers. The data for chromatic (red circles) and achromatic (gray diamonds) conditions have been plotted next to each other for visual comparison. The thresholds of the linear receptive field (LINRF) model (blue squares) were estimated using the parameters of the background variation condition (Figure 7) for the None, Background variation and Simultaneous variation conditions and using the parameters of the light intensity variation condition (Figure 9) for the Light condition. The blue error bars show ± 1 standard deviation estimated over 10 independent simulations. See Figure S7 for LINRF model thresholds with the same set of parameters for all conditions.



1

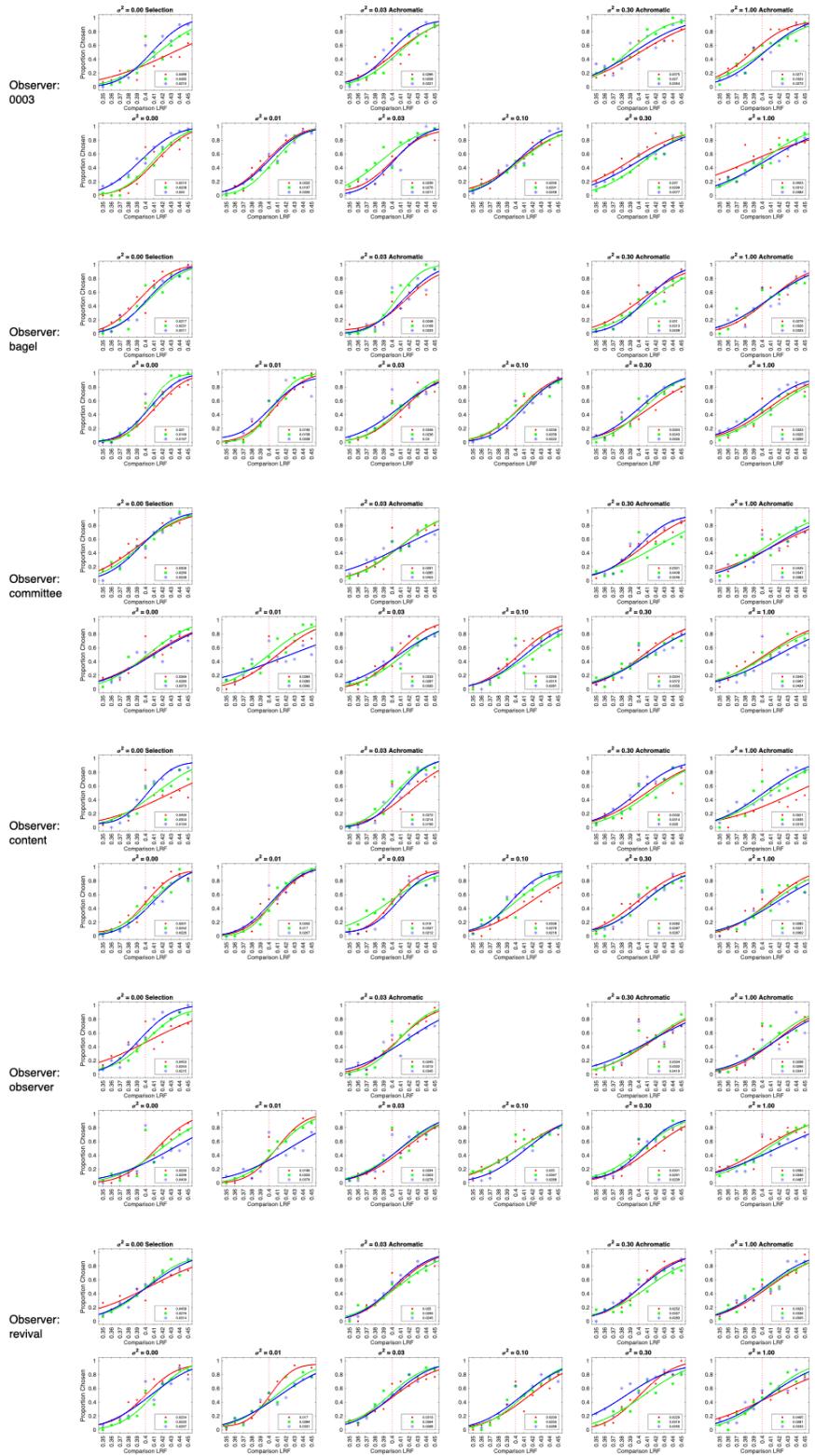
2 **Figure 12: Extrinsic noise of independent variations add linearly for simultaneous variation:** Mean
3 squared thresholds ($N=6$) for the six conditions in simultaneous variation experiment (black circles). The
4 black error bars represent ± 1 SEM taken between observers. The bars (red, gray, blue) represent the
5 increase in squared thresholds compared to the no variation condition (blue dotted line). For the
6 simultaneous variation conditions, the bars on the right (bars with one color, red or gray) represent the
7 increase in measured squared threshold for the simultaneous variation conditions and the bars on the left
8 (stacked bars of two different colors) represent the increase in the sum of the squared threshold of the
9 light intensity variation (blue bar) and the corresponding background variation conditions (red or gray).

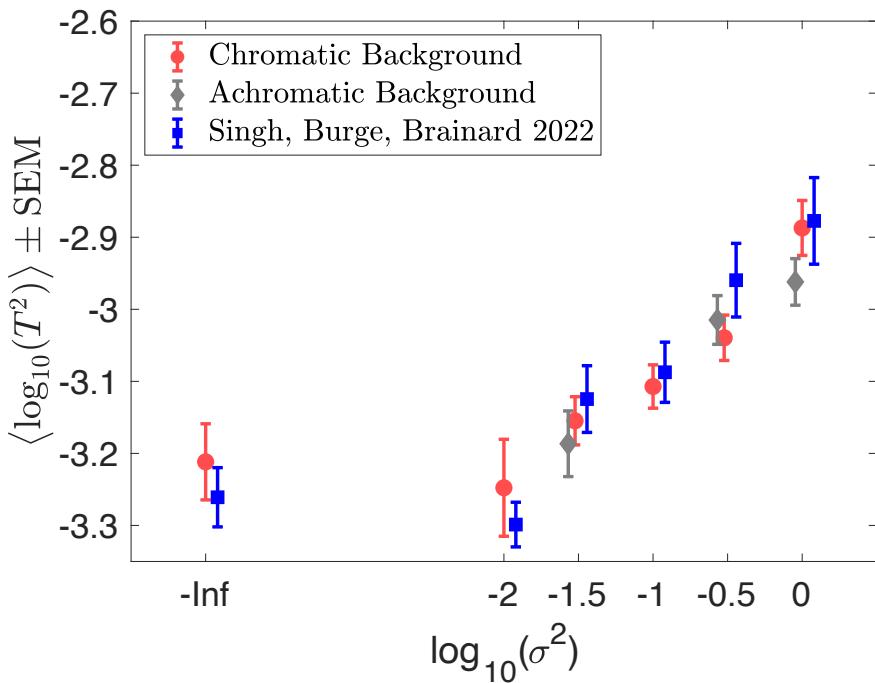


1

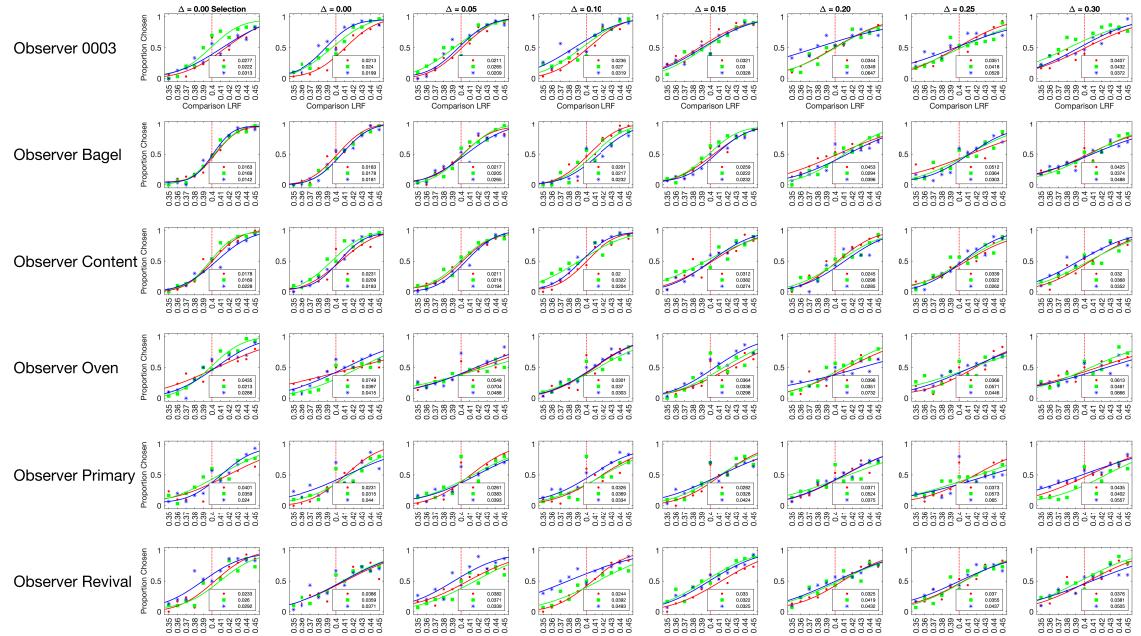
2 **Figure S1: Estimation of extrinsic noise for light intensity variation:** Plot of the variance ($R^T \Sigma R$) as a
3 function of the range parameter δ on a linear (left panel) and logarithmic (right panel) scale. We fit the
4 function with an exponential of the form $A * \exp(B \cdot \delta)$. The variance in the extrinsic noise is estimated
5 as the value of the fit at $\delta = 1$.

2 **Figure S2: Psychometric functions for all observers for background variation experiment.** Same as
3 Figure 6, for all observers retained in background variation experiment.



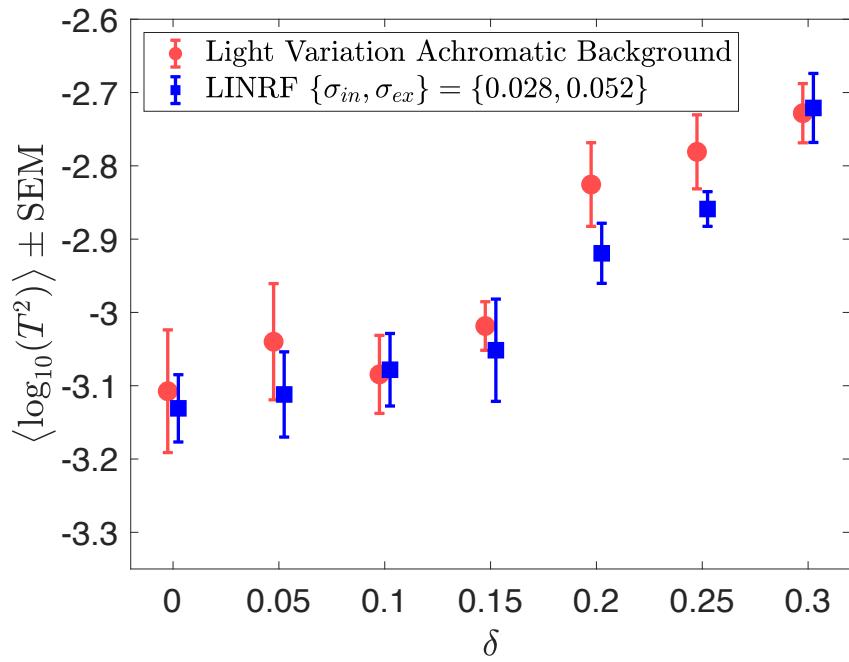


1
2 **Figure S3: Comparison with Singh, Burge, Brainard 2022.** Lightness discrimination thresholds for
3 background variation condition measured in preregistered Experiment 6 and previously reported data
4 from Singh, Burge, Brainard 2022. Preregistered Experiment 6 had both chromatic and achromatic
5 conditions, while the previous experiment (Singh, Burge, Brainard 2022) only had chromatic condition.
6 Singh, Burge, Brainard 2022 made three threshold measurements for each condition for 4 naïve
7 observers.



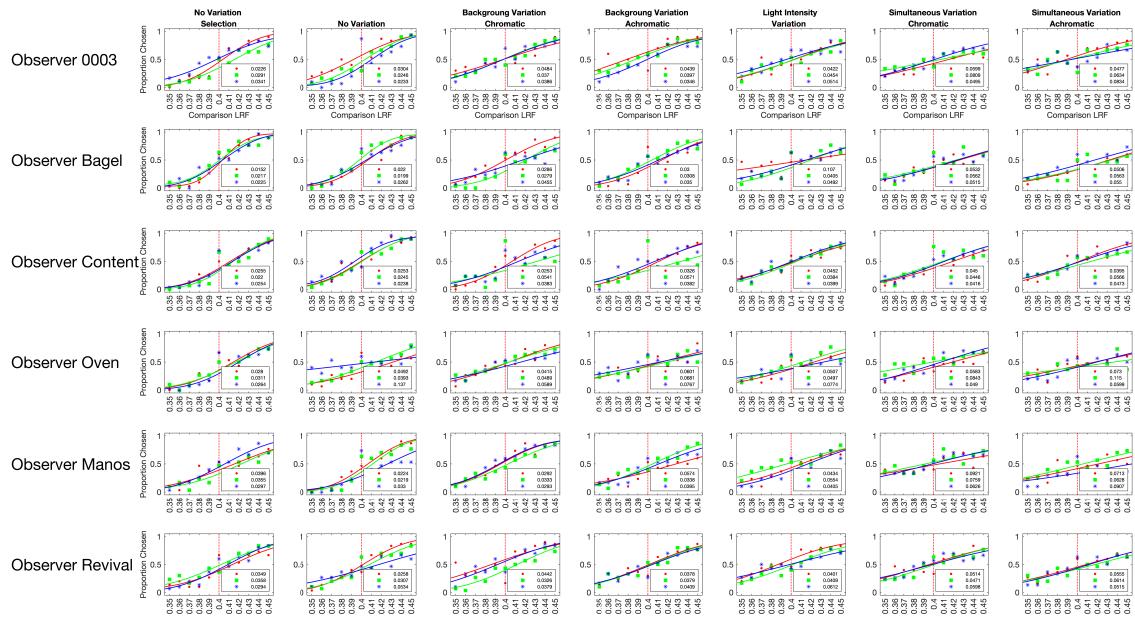
1

2 **Figure S4: Psychometric functions for all observers for light intensity variation experiment.** Same
3 as Figure 8, for all observers retained in light intensity variation experiment.



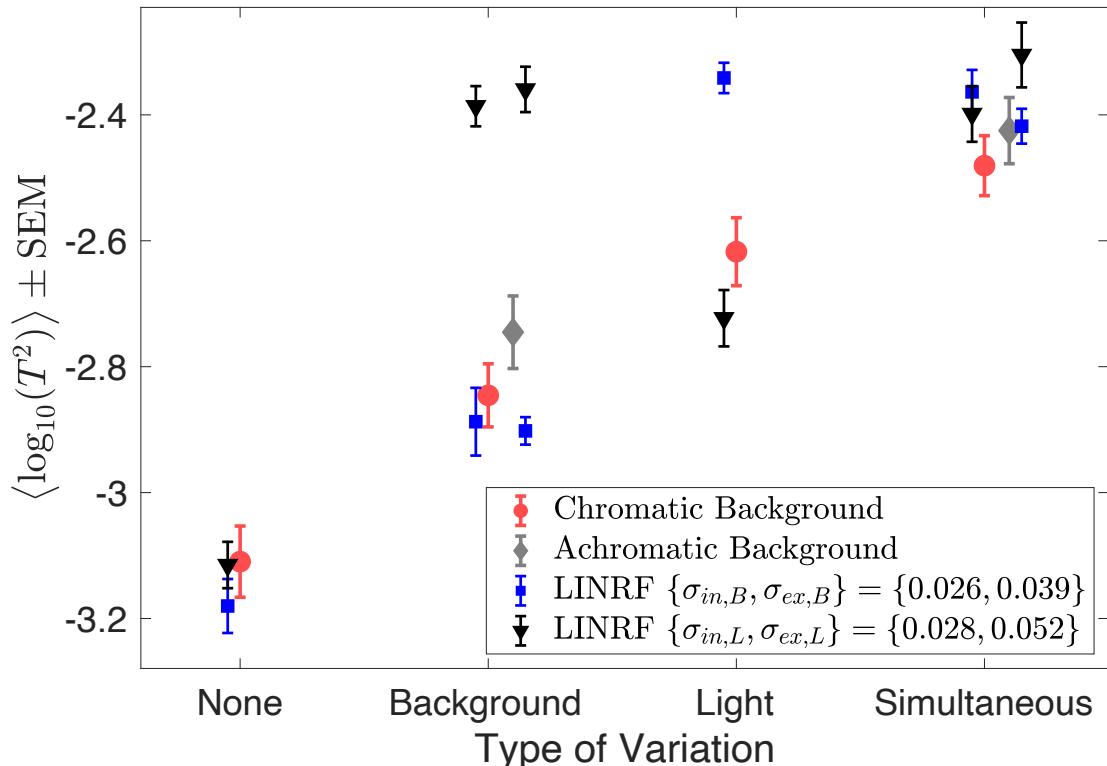
1

2 **Figure S5:** Same as Figure 9, for all six observers retained in light intensity variation experiment. The
3 parameters for the LINRF model are the same as in Figure 9.



1

2 **Figure S6: Psychometric functions for all observers for simultaneous variation experiment.** Same as
3 Figure 10, for all observers retained in simultaneous variation experiment.



1
 2 **Figure S7:** Same as Figure 11, but the thresholds of the linear receptive field (LINRF) model were
 3 estimated using the same set of parameters for all six conditions studied in Experiment 8. Blue square
 4 markers show log squared thresholds estimated using the parameters of the background variation
 5 condition (Experiment 6, Figure 7). Black triangular markers show log squared thresholds estimated using
 6 the parameters of the light intensity variation condition (Experiment 7, Figure 9). The black and blue error
 7 bars show +/- 1 standard deviation estimated over 10 independent simulations. The parameters of the
 8 background variation condition (Experiment 6, blue squares) predict the thresholds of the no variation
 9 condition, the background variation condition, and the simultaneous variation condition quite well, but
 10 fail to predict the threshold of the light source intensity variation condition. Similarly, the parameters of
 11 the light source intensity variation condition (Experiment 7, black triangles) predict the thresholds of the
 12 no variation condition, the light source intensity variation condition, and the simultaneous variation
 13 condition quite well, but fail to predict the threshold of the background variation condition. This could
 14 possibly be because observers in the three experiments were different. Future work would aim at studying
 15 these conditions using the same set of observers.