

1 Title: Characterization of human lightness discrimination thresholds for independent spectral variations

2 Authors: Devin Reynolds, Vijay Singh.

3 Department of Physics, North Carolina Agricultural and Technical State University, Greensboro, NC,
4 USA.

5

6 **ABSTRACT:** The lightness of an object is an intrinsic property that depends on its surface reflectance
7 spectrum. The visual system estimates an object's lightness from the light reflected off its surface. But the
8 reflected light also depends on object extrinsic properties of the scene, such as the light source. For stable
9 perception, the visual system needs to discount the variations due to the object extrinsic properties. We
10 characterize this perceptual stability for variation in two spectral properties of the scene: the reflectance
11 spectra of background objects and the intensity of light sources. We measure human observers' thresholds
12 of discriminating computer-generated images of 3D scenes based on the lightness of a spherical target
13 object in the scene. We measured change in discrimination thresholds as we varied the reflectance spectra
14 of the objects and the intensity of the light sources in the scene, both individually and simultaneously. For
15 small amounts of extrinsic variations, the discrimination thresholds remained nearly constant indicating
16 that the thresholds were dominated by observers' intrinsic representation of lightness. As extrinsic
17 variation increased, it started affecting observers' lightness judgment and the thresholds increased. We
18 estimated that the effects of extrinsic variations were comparable to observers' intrinsic variation in the
19 representation of object lightness. Moreover, for simultaneous variation of these spectral properties, the
20 increase in threshold squared compared to the no-variation condition was a linear sum of the
21 corresponding increase in threshold squared for the individual properties, indicating that the variations
22 from these independent sources combine linearly.

23 **KEYWORDS:** Lightness, Human Psychophysics, Color Vision

24 **PRECIS:** We measure human lightness discrimination thresholds as a function of the amount of variation
25 in object-extrinsic spectral properties of visual scenes. We show that the visual system largely
26 compensates for such variations and that the effect of variation in independent properties combines
27 linearly.

29 **INTRODUCTION**
30 Our visual system provides perceptual representations of distal properties of objects based on the
31 proximal stimuli captured by the eyes. While object properties are intrinsic to the object (its color, shape,
32 etc.), the proximal stimuli also depend on the properties of the scene in which the object lies (object-
33 extrinsic properties such as background objects in the scene, illumination, etc.) as well as the position and
34 pose of the observer. The task of the visual system is to provide stable correlates of object-intrinsic
35 properties under variability of the proximal signal due to object-extrinsic scene properties. This work
36 quantifies the extent to which the visual system provides such stability for the representation of the
37 reflectance of an object under variation in spectral properties of the scene, specifically, variation in the
38 spectra of the background objects and the intensity of light sources in the scene.

39 The perceptual correlate of the diffuse spectral reflectance of an object is its perceived color. For
40 achromatic objects, the analogous perceptual quantity is object lightness. The human visual system is
41 known to provide a relatively stable representation of the color/lightness of an object despite variability in
42 the proximal signal due to changes in the light source, the surface reflectance of other objects in the scene,
43 and the geometry and other properties of the scene (Foster, 2011; Brainard & Radonjic, 2004). The degree
44 to which such stability can be achieved is termed *color/lightness constancy* (Adelson, 2000; Gilchrist,
45 2006). Human color/lightness constancy has been measured using appearance-based approaches and
46 discrimination-based approaches (Olkonen & Ekroll, 2016). Appearance-based approaches involve tasks
47 in which the observer makes a judgment about the appearance of stimuli. This approach includes methods
48 such as color matching, color naming, scaling, and nulling (Foster, 2003). In color matching, observers
49 adjust a test stimulus to match a standard stimulus. Color matching experiments show varying degrees of
50 constancy, with constancy measured between 15%-90% under conditions such as changes of illumination
51 (Arend & Goldstein, 1987; Arend & Spehar, 1993), reflectance (Arend & Spehar, 1993; Patel,
52 Munasinghe, & Murray, 2018), illumination gradients (Arend & Goldstein, 1990; Brainard, Brunt, &
53 Speigle, 1997), and illumination and simulated reflectance (Rutherford & Brainard, 2002). Color naming
54 is a more direct and arguably natural method to measure color constancy where observers are asked to
55 categorize stimuli based on their color (Troost & De Weert, 1991). This method has been used with real
56 (Uchikawa, Uchikawa, & Boynton, 1989; Olkkonen, Witzel, Hansen, & Gegenfurtner, 2010) and
57 simulated stimuli (Olkonen, Hansen, & Gegenfurtner, 2009) to measure constancy. Color naming
58 methods have the limitation that there is a vast number of possible discernible colors (Linhares, Pinto, &
59 Nascimento, 2008), but typically observers are asked to name from a small set of color names (Speigle &
60 Brainard, 1996; Smithson & Zaidi, 2004; Hansen, Walter, & Gegenfurtner, 2007) which might provide an
61 overestimate of the measured constancy. In color scaling methods, observers view a stimulus and provide
62 a rating on a scale for a set of colors, thus allowing for a finer level of comparison for measuring
63 constancy (Luo, et al., 1991; Schultz, Doerschner, & Maloney, 2006). Scaling methods can also be used
64 to measure changes in stimuli, where observers provide a rating of the change between stimuli (Ennis &
65 Doerschner, 2019). Nulling or achromatic adjustment methods involve changing a test stimulus such that
66 it appears achromatic (Arend, 1993; Brainard, 1998; Delahunt & Brainard, 2004). This method has the
67 limitation that it provides data only for achromatic/gray stimuli and additional assumptions about the
68 observers' criterion need to be made for color appearances (Speigle & Brainard, 1996).

69 Discrimination-based approaches provide an objective method to measure color constancy (Bramwell &
70 Hurlbert, 1996; Reeves, Amano, & Foster, 2008). In these experiments, observers discriminate stimuli as
71 to whether they are the same or different from each other. The stimuli are varied in some relevant
72 parameter space to measure the threshold for discriminating changes in the parameter (Craven & Foster,
73 1992; Pearce, Crichton, Mackiewicz, Finlayson, & Hurlbert, 2014; Aston, Radonjic, Brainard, &
74 Hurlbert, 2019). Recently, Singh et al. (Singh, Burge, & Brainard, 2022) developed an *equivalent noise*
75 paradigm that relates thresholds of discrimination to the variability in observers' intrinsic representation
76 of object properties (e.g., its lightness) and the variability due to object extrinsic properties of the scene.
77 They measured human lightness discrimination thresholds as a function of the amount of variability in the

78 spectra of background objects in a scene. They related the discrimination thresholds to the variance in
79 observers' internal perceptual representation of lightness and the variance in the spectrally induced
80 extrinsic variability. A comparison of the strength of intrinsic and extrinsic variability provided a measure
81 of the degree of constancy in the object-intrinsic property due to the variability in object extrinsic
82 property.

83 This equivalent noise paradigm can also be used to compare the effect of different sources of variabilities.
84 The variance of multiple extrinsic properties can be characterized relative to the variance of intrinsic
85 variability. These in turn can be compared to each other to measure their relative effects. It can also be
86 used to characterize how the effect of multiple sources of variability combines when presented
87 simultaneously.

88 In this work, we use this paradigm to compare the variation in two spectral properties of the scene to
89 human observers' representation of lightness. The spectral variations we study are the surface reflectance
90 of the background objects in the scene and the intensity of the light sources in the scene. We measure
91 human observers' threshold of discriminating two images based on the lightness of an achromatic target
92 object in the images. We measure how these discrimination thresholds change as we increase the
93 variability in the reflectance spectra of the background objects and the intensity of the light sources. We
94 measure discrimination thresholds for individual and simultaneous variations of these properties. We use
95 the equivalent noise paradigm to relate the thresholds to the variance of observers' intrinsic noise and
96 extrinsic variability. These variances allow one to compare the relative effect of these spectral variations.
97 A comparison of variances of individual and simultaneous variation conditions can provide information
98 about the combination rules of multiple sources of variation.

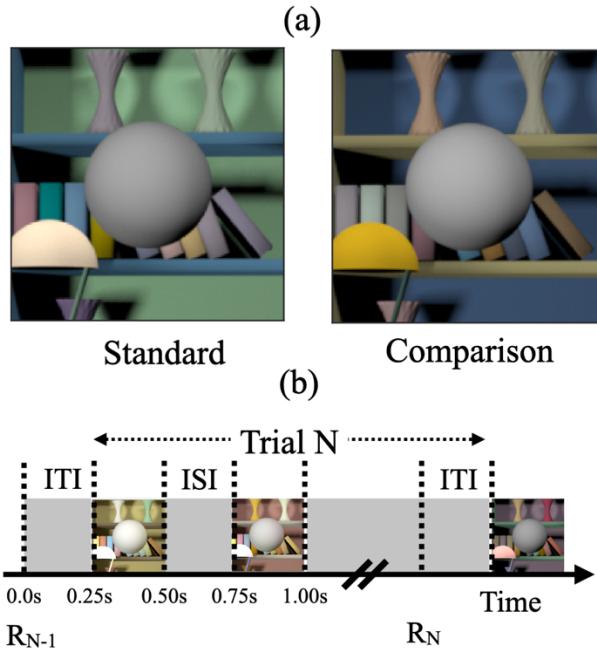
99 We show that as the variability in the extrinsic sources increases, initially for a small amount of variation,
100 the thresholds remain nearly constant. In this region, the thresholds are determined primarily by the
101 variation in observers' internal representation of lightness. As the variability increases further, the
102 discrimination thresholds increase. The increase in thresholds can be accounted for by a model based on
103 signal detection theory. This model shows that the effect of extrinsic variation is within a factor of two
104 compared to the variability in the intrinsic representation of lightness. This indicates that the visual
105 system provides a large degree of lightness constancy under object extrinsic scene variation. By
106 comparing the increase in thresholds of the individual and simultaneous variation condition from the no
107 extrinsic variation condition, we show that the effects of individual sources combine linearly under
108 simultaneous variation.

109 The paper is organized as follows. Section 2, Experimental Methods, provides the details of the
110 experimental methods, stimuli used, and model fitting. Section 3, Results, provides the results of three
111 experiments: variation in background reflectance spectra, variation in light source intensity, and
112 simultaneous variation in these two properties. Section 4, Discussion, provides a summary of the results
113 and remarks.

114 **2 EXPERIMENTAL METHODS**

115 **Overview**

116 We followed the methodology of a previous study (Singh, Burge, & Brainard, 2022) that measured
117 human lightness discrimination thresholds under variability of the reflectance spectra of background
118 objects in the scene. The current study employs similar experimental methods but uses different stimuli.
119 This section outlines the methodology, emphasizing the changes compared to the previous work.



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. The proportion comparison
12 chosen data was used to determine the lightness discrimination threshold (Figure 2). We studied how the
13 lightness 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 Nth trial begins 250ms after the completion
16 of the (N-1)th trial (Inter Trial Interval, ITI = 250ms). In the Nth 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.
22

120 Similar to the previous work, we used a two-alternative forced-choice (2AFC) procedure to measure
121 thresholds (Figure 1). On each trial of the experiment, observers viewed pairs of images of computer-
122 generated 3D scenes displayed on a color-calibrated monitor. These pairs consisted of a standard image
123 and a comparison image. The images were presented sequentially for 250ms, with a 250ms inter-stimulus
124 interval (ITI) between them. Both images contained a centrally located achromatic sphere as the target
125 object. Observers were required to indicate the image in which the target object was lighter. Between
126 trials, we manipulated the luminous reflectance factor (LRF) of the target object in the comparison image.
127 The LRF is defined as the ratio of the luminance of a surface under a reference illuminant to the
128 luminance of the reference illuminant itself (American Society for Testing and Materials, 2017). The
129 reference illuminant was chosen as the CIE D65 standard illuminant. The order of the standard and the
130 comparison image was chosen in a pseudorandom order. We plotted the psychometric functions of the
131 observers by collecting data on the proportion of times the observer chose the comparison image to be
132 lighter (see Figure 2 for an example). To estimate observers' discrimination thresholds, we fit the
133 proportion-comparison-chosen data with a cumulative normal function. We defined the threshold (T) as
134 the difference between the LRF of the target object for which the cumulative normal fit was equal to 0.76
135 and 0.50. This corresponds to a d-prime of 1 in a two-interval task.

136 We measured the effect of variation in two types of object-extrinsic scene properties on human lightness
137 discrimination thresholds: variation in the reflectance spectra of the background objects in the scene and
138 variation in the intensity of the light sources in the scene. We performed three experiments. These
139 experiments were preregistered (see below Preregistration).

140 (1) *Background reflectance variation*: In this experiment, we measured human lightness discrimination
141 thresholds as a function of the amount of variation in the background objects while the spectra of the light
142 sources were kept fixed.

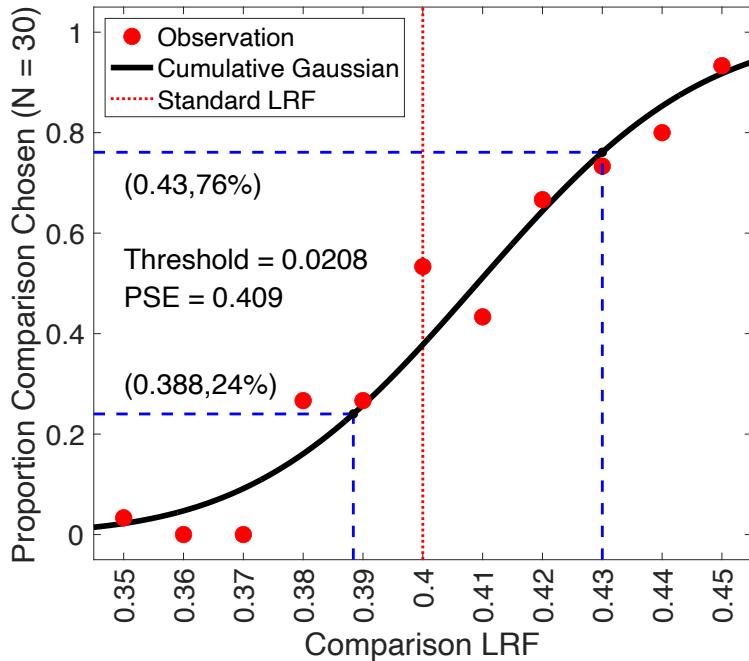
143 (2) *Light source intensity variation*: In this experiment, we measured lightness discrimination thresholds
144 as a function of the amount of variation in the intensity of the light sources while the background was
145 fixed.

146 (3) *Simultaneous variation*: In this experiment, we measured lightness discrimination thresholds as both
147 the background object reflectance spectra and the light source intensity varied simultaneously.

148 To study the effect of variation in surface reflectance of background objects, we generated samples of
149 surface reflectance spectra from a multivariate normal distribution. The distribution was a statistical
150 model of databases of natural surface reflectance measurements. The amount of variation in the surface
151 reflectance of background objects was varied by changing the size of the covariance matrix of the
152 multivariate normal distribution. We measured discrimination thresholds for both chromatic and
153 achromatic variations. In the chromatic variation, the reflectance spectra could take any shape and the
154 objects varied in their luminance and chromaticity. In the achromatic variation, the reflectance spectra
155 were spectrally flat, and the objects were gray.

156 The shape of the spectral power distribution function of the light source was chosen as the CIE D65
157 reference illuminant. The intensity was varied by multiplying the spectral power distribution function by a
158 scalar sampled from a log uniform distribution. The amount of variation was controlled by changing the
159 range of the log uniform distribution.

160 The subsections below provide additional methodological detail.



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 a 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 LRF of the target object in the standard image was 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 for the no-variation ($\sigma^2 = 0.00, \delta = 0.00$) condition. The
 13 discrimination threshold was measured to be 0.0208. The point of subjective equality (PSE), which
 14 corresponds to a proportion of 0.5 in the comparison task, was found to be 0.409. The lapse rate for this
 15 particular fit was determined to be 0.00.
 16

161 **Ethics Statement**

162 All experiments were approved by North Carolina Agricultural and Technical State University
163 Institutional Review Board and were in accordance with the World Medical Association Declaration of
164 Helsinki.

165 **Preregistration**

166 We preregistered the experiments performed in this work before collecting data. The preregistration
167 documents are available at: <https://osf.io/7tgy8/>.¹

168 The experiments were preregistered as Experiment 6 (referred to as *background reflectance variation*),
169 Experiment 7 (referred to as *light source intensity variation*), and Experiment 8 (referred to as
170 *simultaneous variation*). Experiment 6 was a replication of previous work (preregistered as Experiment 3;
171 Singh, Burge, & Brainard, 2022) with additional conditions in which the background objects were
172 achromatic and varied only in their lightness. While the stimuli were different for the three experiments
173 (preregistered Experiments 6, 7, and 8), the experimental method to measure lightness discrimination
174 thresholds was the same.

175 The preregistration documents mentioned that the experiments aimed at characterizing the dependence of
176 human lightness discrimination thresholds on the amount of variation in the background reflectance and
177 the intensity of the light source in the scene. The method of estimating discrimination thresholds was
178 described in the document. We predicted that the thresholds would increase with an increase in the
179 amount of variation. For background reflectance variation, we predicted that the thresholds of achromatic
180 variation would be lower than chromatic variation. We also predicted that the increase in thresholds could
181 be captured by an equivalent noise model (Singh, Burge, & Brainard, 2022). Additionally, we predicted
182 that the threshold for simultaneous variation would be higher than the threshold for individual variations.

183 **Reflectance and Illumination Spectra**

184 We used a statistical model of natural reflectance datasets to generate reflectance spectra of background
185 objects (Singh, Cottaris, Heasly, Brainard, & Burge, 2018; Singh, Burge, & Brainard, 2022). We
186 combined two datasets of surface reflectance measurements (Vrhel, Gershon, & Iwan, 1994; Kelly,
187 Gibson, & Nickerson, 1943). These datasets contain 632 surface reflectance measurements, 170 spectral
188 measurements for the Vrhel et al. dataset and 462 spectral measurements for the Kelly et al. dataset. We
189 mean-centered the dataset by subtracting out the mean surface reflectance over the 632 measurements.
190 Then we used principal component analysis (PCA) to obtain the projection of the mean-centered dataset
191 along the eigenvectors associated with the six largest eigenvalues. These eigenvalues captured more than
192 99.5% of the variance (Singh, Cottaris, Heasly, Brainard, & Burge, 2018). The empirical distribution of
193 the projection weights thus obtained was approximated with a multivariate normal distribution. To get the
194 projection weights of random samples of reflectance spectra, pseudorandom samples were generated from
195 this multivariate normal distribution. Reflectance spectra were constructed by using these projection
196 weights along with the eigenvectors and adding the mean of the surface reflectance dataset. A physical
197 realizability condition was imposed on these spectra by ensuring that the reflectance at each wavelength
198 was between 0 and 1. If a reflectance spectrum did not meet this criterion, it was discarded.

¹ The preregistration documents relevant to this work are those for Experiments 6, 7 and 8. The site also contains preregistrations for previously reported (Experiment 1, 2 and 3; Singh, Burge, & Brainard, 2022) and unreported (Experiment 4) work.

199 To generate achromatic surface reflectance spectra, after generating a physically realizable reflectance
200 spectrum, its average reflectance over all wavelengths was calculated and it was replaced by a spectrum
201 that had this average reflectance at all wavelengths.

202 To control the amount of variation in the reflectance spectra, the covariance matrix of the multivariate
203 normal distribution was multiplied by a covariance scalar (σ^2). A covariance scalar of 0 indicates that
204 there is no variation in the reflectance spectra of the background objects. On the other hand, a covariance
205 scalar of 1 corresponds to the range of reflectance variation observed in the combined natural reflectance
206 datasets.

207 The light source power spectrum was chosen to be the CIE D65 reference illuminant. The D65 spectrum
208 was divided by its mean power over wavelength to obtain its relative spectral shape. The variation in the
209 light source intensity was introduced by multiplying the normalized D65 spectrum by a random number
210 generated from a log-uniform distribution in the range $[1 - \delta, 1 + \delta]$, where δ determines the range of the
211 distribution. We chose a log-uniform distribution for the multiplication parameter because the spectral
212 power distribution function of natural daylight spectra varies over three orders of magnitude and their
213 mean across wavelength can be roughly approximated by a log-uniform distribution (Singh, Cottaris,
214 Heasly, Brainard, & Burge, 2018). All light sources in a scene were assigned the same power spectrum.

215 The values of the two parameters σ^2 and δ for the three experiments were as follows:

216 *Background reflectance variation*: In this experiment, we generated images for nine conditions. Six of
217 these conditions were for chromatic variation at six logarithmically spaced values of the covariance scalar
218 (σ^2): $[0, 0.01, 0.03, 0.1, 0.3, 1.0]$ (same as Singh, Burge, & Brainard, 2022). Three conditions were for
219 achromatic variation at covariance scalar (σ^2): $[0.03, 0.3, 1.0]$. The power spectrum of the light source
220 was the same for all images. Figure 3 shows five typical images for the nine conditions.

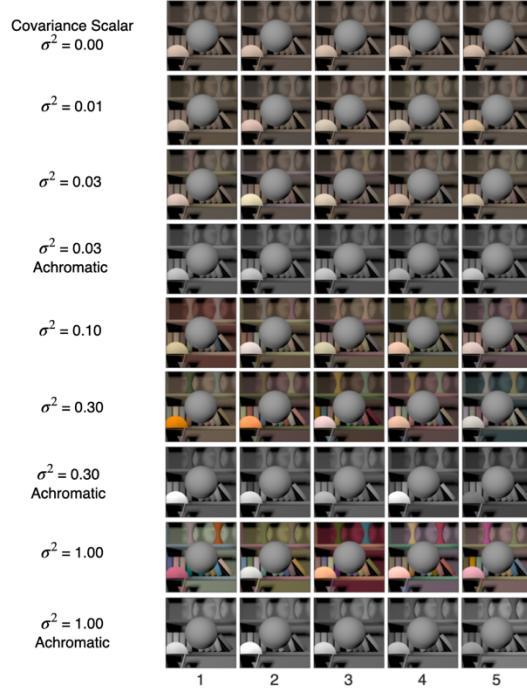
221 *Light source intensity variation*: In this experiment, we generated images for seven linearly spaced values
222 of the range parameter (δ): $[0.00, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30]$. The reflectance spectra of all
223 background objects were the same and were equal to the mean spectrum of the reflectance database. This
224 corresponds to a covariance scalar of 0. Figure 4 shows five typical images for the seven conditions.

225 *Simultaneous variation*: In this experiment, we studied six conditions. These were: no-variation ($\sigma^2 = 0$,
226 $\delta = 0$), chromatic background variation ($\sigma^2 = 1, \delta = 0$), achromatic background variation ($\sigma^2 = 1, \delta =$
227 0), light source intensity variation ($\sigma^2 = 0, \delta = 0.3$), simultaneous variation chromatic background
228 ($\sigma^2 = 1, \delta = 0.3$) and simultaneous variation achromatic background ($\sigma^2 = 1, \delta = 0.3$). Figure 5 shows
229 five typical images for these six conditions.

230 **Image Generation**

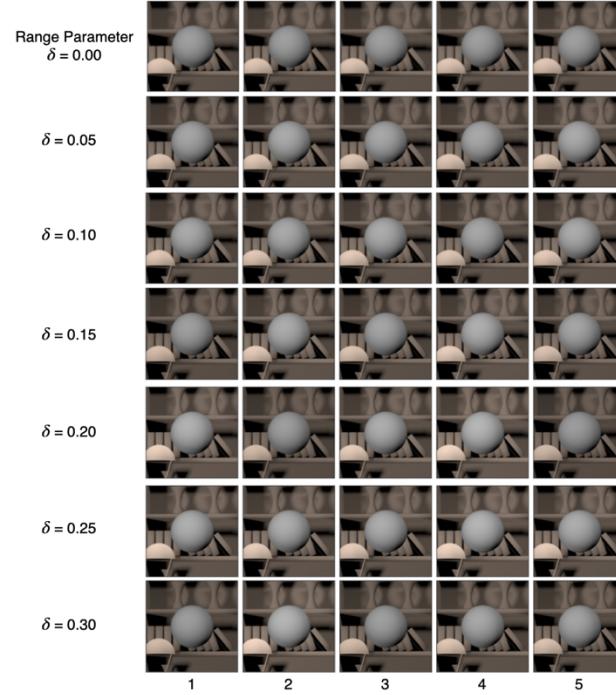
231 The images used in this work were generated using the software Virtual World Color Constancy (VWCC)
232 (github.com/BrainardLab/VirtualWorldColorConstancy) as described in (Singh, Cottaris, Heasly,
233 Brainard, & Burge, 2018). The initial step to generate an image involves constructing a 3D model that
234 serves as the base scene. Then objects and light sources are inserted in the base scene. We chose the
235 *Library* base scene from the base scenes available in VWCC and inserted a spherical object and a
236 spherical light source. The *Library* base scene contained two additional area lights. Then, based on the
237 specific experimental condition, we assign reflectance spectra and spectral power distribution functions to
238 the objects and light sources within the base scene. All light sources within a given scene were assigned
239 identical spectral power distribution functions. All surfaces in the scene were chosen to be matte and had
240 no specularity. Subsequently, we utilize Mitsuba (version 0.5.0), a physically-realistic open-source

1

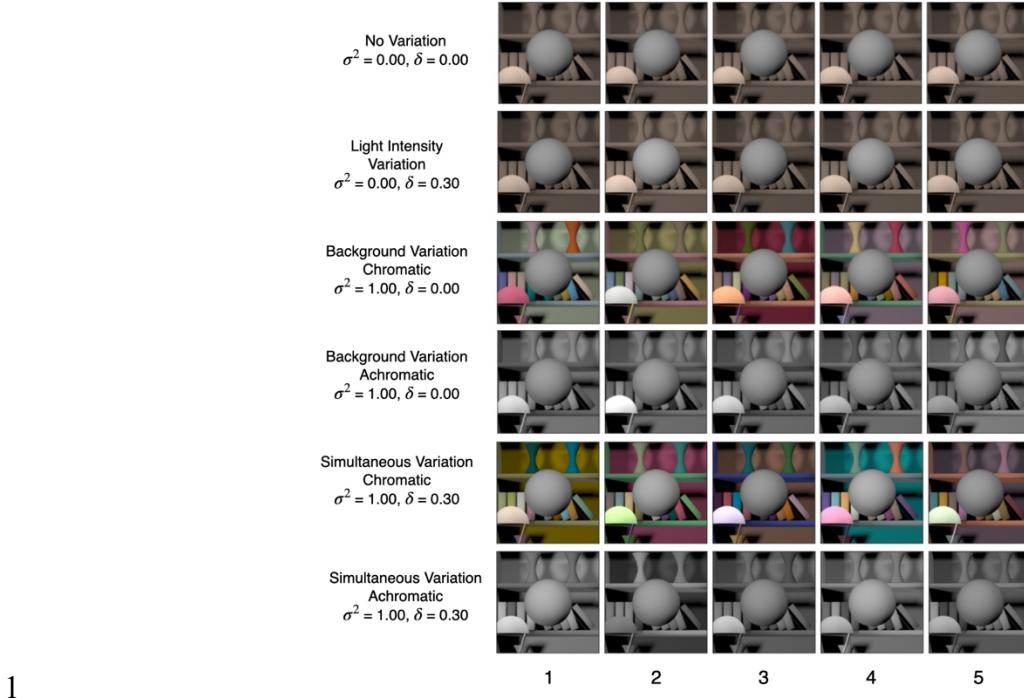


2 **Figure 3: Background reflectance variation:** We studied two types of variations in the reflectance
 3 spectra of background objects in the scene: chromatic variation and achromatic variation. In chromatic
 4 variation, the reflectance spectra could take any shape, and the objects varied in their luminance and
 5 chromaticity. In the achromatic variation, the reflectance spectra were spectrally flat, and the objects
 6 appeared gray and varied only in their luminance. The spectra were chosen from a multivariate normal
 7 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 values of target object LRF in the range [0.35, 0.45]. The target object in each image in the figure
 13 is 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 *simultaneous variation* experiment. We generated 1100 images for each of these conditions, 100
4 images at each value of the target object LRF in the range [0.35, 0.45].

241 rendering system (mitsuba-renderer.org; Jakob, 2010) to produce a 2D multispectral image of the scene.
242 A 201-pixel by 201-pixel part of the image centered at the target object was cropped out to display on the
243 monitor. The cropped multispectral images were converted to gamma corrected RGB images for display
244 on the monitor (Singh, Burge, & Brainard, 2022). For this, the multispectral images were first converted
245 into LMS images using Stockman-Sharpe 2° cone fundamentals. Then monitor calibration data was used
246 to convert the LMS images into RGB images. Then a common scaling was applied to all images used in
247 an experiment to ensure that they were all within the gamut of the monitor. The gamma corrected RGB
248 images were presented on the calibrated monitor during the experiment. When displayed on the
249 experimental monitor at a distance of 75cm from the observers' head position, the image was nearly 2° in
250 visual angle with the target nearly 1° in visual angle.

251 **Stimulus Design**

252 For each condition described above, we generated 1100 images, 100 images at each of the 11 linearly
253 spaced values of the target object LRF in the range [0.35, 0.45]. The standard image target object LRF
254 was 0.4. The comparison image target object LRF varied in the range [0.35, 0.45]. We generated 100
255 images at each comparison level to avoid excessive replication of images in the experiment. For the no-
256 variation ($\sigma^2 = 0.00$, $\delta = 0.00$) condition, we generated one image at each target object LRF level, as the
257 background reflectance and the light source intensity remained fixed in this case. The scene geometry
258 remained fixed during the experiment and the images did not include secondary reflections.

259 The standard image, when presented on the experimental monitor, had an average luminance of 87.1
260 cd/m² for the no-variation condition ($\sigma^2 = 0.00$, $\delta = 0.00$). The average luminance of the target object for
261 the 11 LRF levels were 120.9, 122.3, 123.8, 125.2, 126.5, 127.9, 129.2, 130.5, 131.9, 133.1, 134.4 cd/m².

262 For the simultaneous background and light source intensity variation condition ($\sigma^2 = 1.00$, $\delta = 0.30$), the
263 average luminance of the standard image was 87.8 cd/m². The average luminance of the target object for
264 the 11 LRF levels were 117.7, 119.4, 119.4, 122.3, 123.7, 123.8, 127.8, 126.9, 127.7, 129.1, 129.0 cd/m².

265 **Experimental Structure:**

266 In this study, a trial is defined as the display of a standard and a comparison image on the monitor and the
267 recording of the observer's response. An interval is defined as the presentation of either the standard
268 image or the comparison image within a trial. A block consists of recording 330 trials for one condition,
269 30 trials at each of the 11 comparison image target LRF levels. A permutation consists of recording one
270 block of data for each condition in an experiment. We recorded three permutations for each observer in
271 each experiment. Each permutation had a random order of the blocks.

272 The order of the blocks in a permutation, the LRF levels of the comparison image in trials of a block, and
273 the order of standard and comparison images in a trial were generated pseudorandomly and stored at the
274 beginning of the experiment for each observer. The 30 comparison images at a LRF level of a block were
275 chosen pseudorandomly with replacement from the 100 images available at that LRF level. Before
276 starting a new permutation for an observer, the data for all blocks (conditions) in a permutation was
277 collected.

278 A session consisted of recording three blocks on a single day. An observer performed no more than one
279 session a day. Each block in a session was divided into three sub-blocks of 110 trials. Between these sub-
280 blocks, the observers took a break of a minimum one-minute duration. The observers also took a small
281 break (nearly two to five minutes) between blocks. The observers were informed that they could stop the
282 experiment at any time. If the observer terminated a block, the data was not recorded. No observer
283 terminated a block of the experiment.

284 Each observer first performed a practice session where three blocks of data were recorded for the no-
285 variation ($\sigma^2 = 0$, $\delta = 0$) condition. The observers were excluded from the experiment if their mean
286 threshold for the last two blocks was higher than 0.03. If the observer passed this criterion, then the rest of
287 the data was collected over several days.

288 The experimental procedure was explained to each observer at the beginning of the practice session. The
289 experimenter then obtained the consent for the experiment. Vision tests were performed on the observer
290 to ensure normal visual acuity and normal color vision. After this, the observer went to the experimental
291 room where they were familiarized with the experimental set-up by performing a familiarization block of
292 40 trials. Then the observer was dark adapted by sitting in the dark for about 5 minutes. Then the data for
293 the three blocks of the practice session was recorded. At the end of the practice session, the observer was
294 informed if they could continue the experiment.

295 If the observer was continued, their data was collected over several sessions. The no-variation condition
296 was also included in each permutation of the experiment. So, data for the no-variation condition was
297 collected again. The data for all six observers of an experiment was collected over several weeks. Data for
298 the three experiments was collected in a sequential order starting with *background reflectance variation*,
299 followed by *light source intensity variation*, and finally *simultaneous variation*.
300

301 **Observer Recruitment and Exclusion**

302 The study recruited observers from North Carolina Agricultural and Technical State University and the
303 local Greensboro community. The participants were compensated for their time. The observers underwent
304 a screening process to meet criteria such as a normal visual acuity of 20/40 or better (with corrective
305 eyewear if necessary) and normal color vision, which was assessed using pseudo-isochromatic plates
306 (Ishihara, 1977). The preregistration documents outlined these exclusion criteria (see Methods:
307 Preregistration).

308 We further conducted a practice session to identify observers who could reliably perform the
309 psychophysical task. We first familiarized the observers with the experiment by asking them to respond to
310 40 trials of the experiment for the no-variation condition ($\sigma^2 = 0.00$, $\delta = 0.00$). These trials were
311 divided into 10 easy trials, 10 moderate trials and 20 regular trials. In the easy trials they compared
312 images with target object LRF 0.35 and 0.45. In moderate trials they compared standard image with target
313 object LRF 0.40 to comparison image with target object LRF either 0.35 or 0.45. In the regular trials they
314 compared standard image with target object LRF 0.40 to comparison image with target object LRF
315 chosen randomly in the range [0.35, 0.45]. The data for these 40 trials were not stored. After familiarizing
316 with the experiment, the observer performed three blocks of the experiment for the no-variation condition
317 ($\sigma^2 = 0.00$, $\delta = 0.00$). The threshold was calculated for these three blocks. If the mean threshold of the
318 observer for the last two blocks in the practice session was larger than 0.030, the observer was
319 discontinued. The preregistration document specified this exclusion criterion (See Methods:
320 Preregistration).

321 If the observer met these criteria, they continued with the rest of the experiment.

322 For each observer, the practice session was performed at the beginning of each of the three experiments,
323 irrespective of whether the observer had participated in an earlier experiment. Table 1 summarizes the
324 demographics information of the observers in the three experiments.
325

326

327

Table 1: Observer Information:

Experiment	Background reflectance variation	Light source intensity variation	Simultaneous variation
Total number of observers in practice session	25	15	20
Male/Female	15/10	6/9	11/9
Age Range	19-34 (Mean 23.33)	19-33 (Mean 23.83)	19-28 (Mean 20.8)
Pseudo-names of observers who passed screening. Visual acuities are provided in brackets.	<i>0003</i> (L20/30, R20/20)	<i>0003</i> (L20/30, R20/20)	<i>0003</i> (L20/30, R20/20)
	<i>bagel</i> (L20/20, R20/20)	<i>bagel</i> (L20/20, R20/20)	<i>bagel</i> (L20/20, R20/20)
	<i>committee</i> ^{\$} (L20/25, R20/25)	<i>oven</i> [*] (L20/20, R20/20)	<i>oven</i> (L20/20, R20/20)
	<i>content</i> ^{\$} (L20/30, R20/20)	<i>content</i> ^{\$} (L20/30, R20/20)	<i>content</i> ^{\$} (L20/30, R20/20)
	<i>observer</i> ^{\$} (L20/25, R20/25)	<i>primary</i> ^{\$} (L20/20, R20/20)	<i>manos</i> ^{**} (L20/25, R20/25)
	<i>revival</i> (L20/20, R20/20)	<i>revival</i> (L20/20, R20/20)	<i>revival</i> ^{**} (L20/20, R20/20)

328 \$These observers used personal corrective eyewear during vision test and the experiments.

329 *Observer *oven* reported some difficulties during a few sessions of the *light intensity variation*
330 experiment and their thresholds for two conditions did not fit the expected pattern. We removed their data
331 from the analysis presented in this work. Their data and thresholds are provided in supplementary Table
332 S2.333 **Due to a lack of observers who met the preregistration criteria for the *simultaneous variation*
334 experiment, two observers (*manos*, and *revival*), whose thresholds were close to the preregistration
335 criteria, were also retained for the experiment. Observer *revival* had participated in the previous two
336 experiments and had met the criteria both times. Observer *manos* showed improvement in thresholds with
337 each block, with the threshold for the final block below 0.03. This was a deviation from the
338 preregistration.339 **Apparatus**340 The experiments were performed in a dark room and the stimuli were presented on a color-calibrated
341 LCD monitor (27-in. NEC MultiSync EA271U; NEC Display Solutions). The pixel resolution of the
342 monitor was selected as 1920 x 1080. Its refresh rate was 60Hz and each RGB channel was operated at 8-
343 bit resolution. The experimental computer was an Apple Macintosh with an Intel Core i7 processor. The
344 experimental programs were written in MATLAB (MathWorks; Natick, MA) and utilized Psychophysics
345 Toolbox (<http://psychtoolbox.org>) and mgl (<http://justingardner.net/doku.php/mgl/overview>) libraries. A
346 Logitech F310 gamepad controller was used to collect observers' responses.

347 The distance between the observers' eyes and the monitor was set at 75cm. A forehead rest and chin cup
348 (Headspot, UHCOTech, Houston, TX) were used to stabilize the observers' head position. The observers'
349 eyes were centered both horizontally and vertically in relation to the display.

350 **Monitor Calibration**

351 The monitor was calibrated using a spectroradiometer (PhotoResearch PR655) as described in (Singh,
352 Burge, & Brainard, 2022). For this, we focused the spectroradiometer on a 4.66cm by 4.66cm (3.56° x
353 3.56°) patch on the monitor. The radiometer sampled from a 1° circular spot at the center for the patch
354 due to its optics. We measured the spectral power distribution of the three monitor primaries in the range
355 380nm to 720nm at 4nm intervals. For each primary, the gamma function was determined by measuring
356 the spectral power distribution at 26 equally spaced values in the range [0, 1] where 1 is maximum
357 allowed input and 0 is no input. The linearity of the monitor was also checked by measuring the spectral
358 power distribution at 32 different combinations of the input in the range [0,0,0] and [1,1,1]. The
359 maximum absolute deviation between measured and predicted values were 0.0087 and 0.0081 for x and y
360 chromaticity respectively. For luminance, the maximum absolute deviation was less than 1%.

361 The monitor was calibrated before starting each experiment. Once calibrated, the same settings were used
362 until data for all observers for that experiment was collected. The monitor was then recalibrated for the
363 next experiment. Data was collected in the sequence *background reflectance variation* experiment, *light*
364 *source intensity variation* experiment, and *simultaneous variation* experiment.

365 **Stimulus Presentation**

366 The images were presented on a color calibrated monitor. The size of the image was 2.6cm x 2.6cm
367 corresponding to 2° by 2° visual angle for the observer located at 75cm from the monitor. The size of the
368 spherical target object was ~1°. The rest of the monitor was set at lowest value input [0,0,0]. Except for a
369 4cm x 4cm square part concentric to the image presentation location, the rest of the monitor was also
370 covered with black paper. The images were presented for 250ms each with a 250ms inter-stimulus
371 interval (ISI) between the images. The monitor was dark (lowest value input [0,0,0]) during the ISI. The
372 observer's response was collected after both images were presented and removed from the monitor. The
373 observer could take as long as they wished before responding. The observers were provided auditory
374 feedback about correct or incorrect response after their response. The next trial started 250ms after the
375 feedback. Thus, the actual inter trial interval depended on the response time of the observer.

376 **Code and Data Availability**

377 The data for each experiment and observer is provided as supplementary information (SI). The SI is
378 available at: <https://github.com/vijaysoophie/SimultaneousVariationPaper>. The SI contains the
379 proportion-comparison-chosen data as well as the thresholds for the 3 experimental blocks of each
380 condition, for each experiment and observer. The MATLAB scripts to generate Figures 2, 6 – 12,
381 supplementary Figures S1-S7, and the scripts to obtain thresholds of the linear receptive field formulation
382 of the model are also provided in the SI.

383 **Linear Receptive Field Model**

384 The measured thresholds were fit to a previously developed linear receptive field (LINRF) model (Singh,
385 Burge, & Brainard, 2022) based on signal detection theory (Green & Swets, 1995). In the standard
386 formulation of signal detection theory, the difference between the mean of two Gaussian random
387 variables is directly proportional to the standard deviation of the distribution. The constant of

388 proportionality, called d-prime (d'), measures the distance between the two distributions in the units of
 389 the standard deviation. $d' = 0$ corresponds to the distributions being identical. Larger values of d'
 390 indicate higher discriminability between the two distributions.

391 The linear receptive field (LINRF) model first calculates the responses of a center-surround receptive
 392 field (R) to the standard and comparison images. The receptive field is square in shape to match the
 393 images in the psychophysics experiment. Its center is a circle of radius equal to the size and the location
 394 of the target object. The central region has a spatially uniform positive sensitivity of 1. The surround has a
 395 spatially uniform negative sensitivity of v_s . The receptive field response is calculated as the dot product
 396 of the receptive field (R) with the image (I). A mean-zero Gaussian noise is added to the response. The
 397 noise-added response of the standard and the comparison images are compared using the signal detection
 398 theory framework.

399 Assuming R and I to be column vectors, the response of the receptive field to the standard and the
 400 comparison images can be written as:

$$401 \quad r_{ic} = R^T(I_{c0} + \eta_e) + \eta_i = R^T I_{c0} + \eta \quad (1)$$

$$402 \quad r_{is} = R^T(I_{s0} + \eta_e) + \eta_i = R^T I_{s0} + \eta. \quad (2)$$

403 Here, I_{c0} and I_{s0} represent the comparison and standard images without noise, η_e represents the extrinsic
 404 noise due to variations in the images, η_i represents the intrinsic noise of the observer, and η the overall
 405 noise. Assuming receptive field and noise models to be linear and Gaussian, the overall variance (σ_η^2) can
 406 be related to the variance of the intrinsic noise (σ_{ri}^2) and the covariance matrix of the extrinsic noise (Σ_e)
 407 as:

$$408 \quad \sigma_\eta^2 = (\sigma_{ri}^2 + R^T \Sigma_e R) \quad (3)$$

409 By signal detection theory, the mean difference between the receptive field response to the standard and
 410 the comparison image $R^T(I_{s0} - I_{c0})$ can be related to the variance of the overall noise.

$$411 \quad R^T(I_{s0} - I_{c0}) = d' \sigma_\eta \quad (4)$$

412 We assume that the mean difference $R^T(I_{s0} - I_{c0})$ is proportional to the difference between the LRF level
 413 of the target object in the standard and the comparison images (Δ_{LRF}), i.e., $R^T(I_{s0} - I_{c0}) = C' \Delta_{LRF}$,
 414 where C' is a proportionality constant. Thus,

$$415 \quad \Delta_{LRF} = \frac{d'}{C'} \sigma_\eta. \quad (5)$$

416 In our experiment, the threshold T is defined as the difference between comparison LRF at which the
 417 proportion comparison chosen is equal to 0.76 and 0.50. Assuming that the proportion comparison chosen
 418 is equal to 0.50 when standard and comparison LRF are the same, this choice of threshold ($\Delta_{LRF} = T$),
 419 corresponds to $d' = 1$. Further, we choose $C' = 1$. So, the thresholds (T) can be related to the extrinsic and
 420 intrinsic noise as:

$$421 \quad T = \sqrt{\sigma_{ri}^2 + R^T \Sigma_e R} \quad (6)$$

422 In the case of *background reflectance variation experiment*, the extrinsic noise is varied by multiplying
 423 the covariance matrix of the multinormal distribution from which the sample of reflectance spectra with a
 424 covariance scalar (see Reflectance and Illumination Spectra). Thus, we can write $R^T \Sigma_e R = \sigma^2 \times R^T \Sigma_{e0} R$,
 425 where Σ_{e0} is the covariance matrix of the external noise corresponding to the variation in the dataset of
 426 natural images and σ^2 is the covariance scalar. Thus, the model relates the thresholds (T) in the
 427 experiments to the variance in the intrinsic noise of the observer (σ_{ri}^2) and the variance in the extrinsic
 428 noise ($\sigma_{e0}^2 = R^T \Sigma_{e0} R$) through the relation:

$$429 \quad T^2 = \sigma_{ri}^2 + \sigma^2 \times \sigma_{e0}^2 \quad (7)$$

430 In our experiments, we chose σ^2 on a logarithmic scale. So, we have represented the data using the form:

$$431 \quad \log T^2 = \log(\sigma_{ri}^2 + \sigma^2 \times \sigma_{e0}^2) \quad (8)$$

432 The values of the intrinsic and extrinsic noise standard derivations were estimated by choosing the two
 433 parameters of the model, noise variance (σ_{ri}^2) and surround sensitivity (v_s) that best fits the experimental
 434 data.

435 *Background reflectance variation*: In this case, the Gaussian noise variance ($\sigma_{ri,B}^2$) and the surround
 436 sensitivity ($v_{s,B}$) was chosen to minimize the mean squared difference between the thresholds of the
 437 LINRF model and the experimental thresholds measured at the six values of the covariance scalar. (The
 438 subscript B has been introduced to distinguish the *background reflectance variation* experiment from the
 439 *light source intensity variation* experiment.) For this, we calculated the threshold of the LINRF model
 440 using the two-interval forced choice paradigm for different values of the parameters $\sigma_{ri,B}^2$ and $v_{s,B}$. In each
 441 trial, we randomly sampled a standard and comparison image from our dataset, estimated the response of
 442 the receptive field to the images, and added random samples of intrinsic noise to the responses. The noise
 443 added receptive field response was compared to determine the choice of the trial. We performed 10000
 444 trials for each of the 11 comparison level. The proportion comparison chosen data was used to get the
 445 psychometric function and the threshold, similar to the human experiment. We estimated the model
 446 thresholds for a range of parameter ($\sigma_{ri,B}^2$ and $v_{s,B}$). The mean squared difference between the LINRF
 447 model threshold and the experimental threshold was calculated as a function of the parameters $\sigma_{ri,B}^2$ and
 448 $v_{s,B}$ and was fit to a degree two polynomial of two variables. The resulting polynomial was minimized to
 449 get the best-fit values of the parameters $\sigma_{ri,B}^2$ and $v_{s,B}$. $\sigma_{ri,B}^2$ provided the estimate of the intrinsic noise.
 450 The extrinsic noise ($\sigma_{e0,B}^2$) was estimated by using the best-fit surround sensitivity ($v_{s,B}$) and the sample
 451 covariance matrix of the images (Σ_{e0}) at $\sigma^2=1$ in the relation $R^T \Sigma_{e0} R$.

452 *Light source intensity variation*: To estimate the intrinsic and extrinsic noise variance in the *light source*
 453 *intensity variation* experiment, we fit the thresholds to a functional form similar to Eq. (7) where we
 454 replace the covariance scalar σ^2 by the range parameter δ . We chose the value of the Gaussian noise
 455 variance ($\sigma_{ri,L}^2$) and receptive field surround sensitivity ($v_{s,L}$) to minimize the mean squared difference
 456 between the experimental and LINRF model thresholds measured at the seven values of the range
 457 parameter as explained above. (The subscript L has been introduced to distinguish the *light source*
 458 *intensity variation* from the *background reflectance variation* experiment.) $\sigma_{ri,L}^2$ provided the estimate of
 459 the observer's intrinsic noise.

460 In natural viewing conditions, the intensity of the light source varies over several orders of magnitude
 461 (Singh, Cottaris, Heasly, Brainard, & Burge, 2018). This range of variation cannot be captured in the
 462 experiment due to the limitations of the monitor. To estimate extrinsic noise, we used the best-fit

463 surround sensitivity $v_{s,L}$ and the sample covariance matrices to calculate the quantity $R^T \Sigma_{e0} R$ as a
464 function of the range parameter δ . We fit the resulting values with an exponential function (see Figure
465 S1). The extrinsic noise ($\sigma_{e0,L}^2$) was chosen as the value of the exponential fit at $\delta = 1$. The parameter
466 ($\sigma_{e0,L}^2$) could be used to estimate the extrinsic noise for more naturalistic variations using the exponential
467 fit.

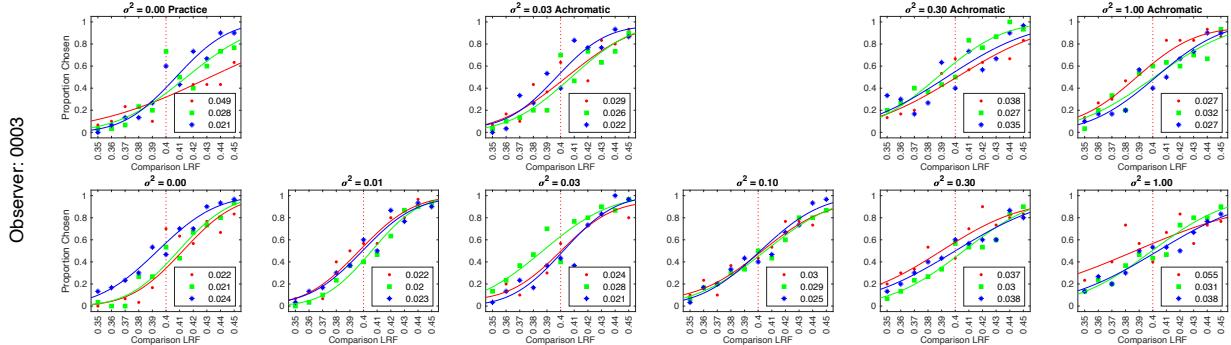
468 3 RESULTS

469 Human Lightness Discrimination Thresholds Increase with Background Reflectance Variation 470

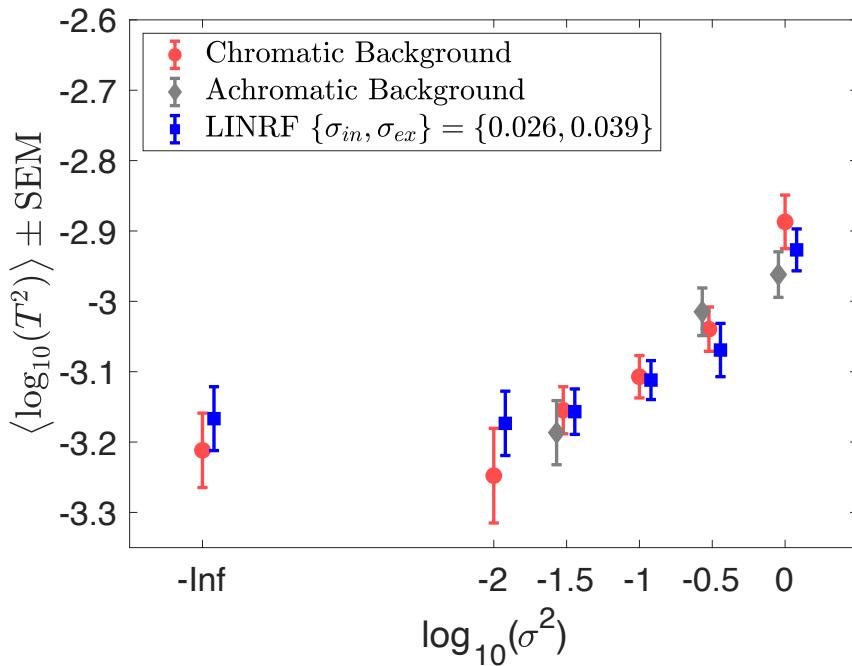
471 We measured the lightness discrimination thresholds of human observers for two types of variation in the
472 reflectance spectra of background objects in the scene: chromatic variation and achromatic variation. In
473 chromatic variation, the reflectance spectra could take any shape and thus the background objects varied
474 in their chromaticity and luminance. In the achromatic variation, each spectrum had the same reflectance
475 at all wavelengths, and thus the spectra varied only in their overall luminance and the objects were gray.
476 The reflectance spectra were sampled from a statistical model of natural surface reflectance spectra
477 databases (see Methods: Reflectance and Illumination Spectra). The model approximated the database
478 with a multivariate normal distribution whose covariance matrix was multiplied by a covariance scalar
479 (σ^2) to control the amount of variation in the sampled spectra. We measured the discrimination thresholds
480 of six human observers at six values of the covariance scalar for chromatic variation and three values of
481 covariance scalar for achromatic variation. For each of the nine conditions and each observer,
482 discrimination thresholds were measured three times in three separate blocks. The psychometric functions
483 for these nine conditions are shown for one observer in Figure 6. The psychometric functions of all six
484 observers are shown in Figure S2. We notice that as the covariance scalar increases, the slope of the
485 psychometric functions decreases, corresponding to an increase in discrimination thresholds. The
486 thresholds for chromatic and achromatic variation are comparable.

487 Figure 7 shows the change in discrimination thresholds as a function of the variance in the reflectance of
488 the background objects. Here we plot the mean log threshold squared (averaged across observers, $N = 6$)
489 as a function of the log of the covariance scalar. The thresholds and standard error of the mean (SEM)
490 from Figure 7 are listed in Table S1. We observe that at small values of the covariance scalar the
491 thresholds are nearly constant. Once the value of the covariance scalar starts increasing, the log threshold
492 squared increases. Further, the thresholds are comparable for chromatic and achromatic variation. The
493 one-way ANOVA p-values of the hypothesis that the mean thresholds for chromatic and achromatic
494 variations are equal are 0.72, 0.57, and 0.16 for covariance scalar 0.03, 0.30, and 1.00 respectively. This
495 indicates that the differences in the mean thresholds are not statistically significant. Figure S3 shows the
496 discrimination thresholds measured in *background reflectance variation* and previously reported
497 thresholds by (Singh, Burge, & Brainard, 2022). The measured thresholds are consistent between the two
498 experiments.

499 We fit the thresholds to the linear receptive field (LINRF) model (Eq. 7) developed by (Singh, Burge, &
500 Brainard, 2022). The LINRF model provides an estimate of the variance of the internal noise of the
501 observer as $\sigma_{ri,B}^2 = 0.026$ and the variance of the extrinsic variability due to the reflectance of
502 background objects as $\sigma_{e0,B}^2 = 0.039$. The equivalent noise level, the ratio of the external variance to
503 intrinsic noise, is ~ 1.5 , indicating that the variability in the representation of object lightness induced by
504 the natural variability in the reflectance of background objects is close to the internal variability of that
505 representation. If the ratio was equal to 1, then we would have concluded that the visual system has
506 discounted the external variability. But the ratio is not significantly large compared to 1, this indicates that
507 the visual system provides a significant level of lightness constancy.



1
2 **Figure 6: Psychometric functions of observer 0003 for background reflectance variation experiment:**
3 We measured the proportion-comparison-chosen data for the nine conditions separately in three blocks
4 for each observer. The figure shows the psychometric function for observer 0003. The psychometric
5 functions for all six observers are shown in Figure S2. A cumulative normal function was fit to the data
6 from each block to determine the discrimination threshold (see Figure 2). The legend provides the
7 estimated lightness discrimination threshold for each block, obtained from the cumulative fit. The first
8 panel in the top row shows the data and thresholds for the practice session. An observer was selected for
9 the experiment only if the average of their last two discrimination threshold measurements in the practice
10 session was less than 0.30. The last three panels in the top row show the data for the three achromatic
11 conditions. The bottom row shows the data for the chromatic variation conditions.



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

508 **Human Lightness Discrimination Thresholds Increase with Light Source Intensity Variation**

509 We measured the lightness discrimination thresholds of human observers as we varied the intensity of
510 light sources in the scene. The shape of the spectrum of the light sources was fixed to be standard daylight
511 spectrum D65. We normalized the spectrum by its mean over wavelengths. The intensity was varied by
512 multiplying the normalized spectrum by a scalar sampled from a log-uniform distribution in the range [1-
513 δ , $1 + \delta$]. The reflectance spectra of the background objects were fixed. We measured lightness
514 discrimination thresholds for seven values of the range parameter δ for six human observers. The
515 psychometric functions of one of the observers for these seven conditions are shown in Figure 8. The
516 psychometric functions of all six observers are shown in Figure S4. Figure 9 shows the change in
517 thresholds as the amount of variation in the light source intensity increases. The data is averaged over five
518 observers (the data for one of the observers has been removed from Figure 9, see Figure S5 for the data
519 for all observers). Table S2 lists the mean thresholds and the SEM measured in this experiment. Similar to
520 the trend for reflectance spectra variation, lightness discrimination thresholds remain nearly constant for
521 small values of the range parameter, and then the log threshold squared increases with an increase in the
522 range parameter. A fit of the mean squared threshold with the LINRF model gives the value of internal
523 noise as $\sigma_{ri,L}^2 = 0.028$. This compares well with the internal noise obtained from the *background*
524 *reflectance variation* experiment ($\sigma_{ri,B}^2 = 0.026$). The variance of the extrinsic variability estimated at the
525 range parameter $\delta = 1.00$ is $\sigma_{eo,L}^2 = 0.052$. The equivalent noise level, the ratio of external variation to
526 intrinsic noise is ~ 1.8 . This indicates that the variation in the lightness representation induced by the
527 variation in light source intensity is close to the internal variation of that representation at these levels. In
528 natural conditions, where light source intensity varies over several orders of magnitude, the extrinsic
529 noise can be estimated by generating images at the level of variation and using the LINRF model.

530 **Thresholds for Simultaneous Variation are Higher Than Individual Variations**

531 We measured lightness discrimination thresholds for simultaneous variation in the reflectance spectra of
532 background objects and the intensity of the light sources in the scene. In this experiment, we studied six
533 conditions: no-variation, variation in the reflectance spectra of background objects with a fixed spectrum
534 of the light sources for achromatic and chromatic backgrounds, variation in the intensity of light sources
535 with a fixed background, and simultaneous variation in the intensity of the light sources and the
536 reflectance spectra of background objects for chromatic and achromatic backgrounds. We measured the
537 lightness discrimination thresholds of six human observers for these six conditions. The psychometric
538 function of one of the observers is shown in Figure 10. The psychometric functions of all six observers
539 are shown in Figure S6. Figure 11 shows the mean log squared threshold of all six observers for these six
540 conditions. Table S3 lists the mean thresholds and SEM from Figure 11. The threshold for simultaneous
541 variation of light source intensity and reflectance spectra of background objects is higher than the
542 condition with individual variations in these properties. As observed earlier, the threshold for achromatic
543 and chromatic conditions are comparable. The one-way ANOVA p-value of the hypothesis that the mean
544 thresholds for chromatic and achromatic variations are equal is 0.19 for the background variation
545 condition and 0.44 for the simultaneous variation condition, indicating that the differences in the mean
546 thresholds are not statistically significant.

547 Figure 11 also shows the squared thresholds of the LINRF model for the six conditions. We used the
548 intrinsic noise and the surround sensitivity (v_s) parameters of the *background reflectance variation*
549 experiment to estimate the threshold of the LINRF model for the no-variation condition, background
550 spectra variation conditions, and simultaneous variation conditions (Experiment 6, Figure 7). For the light
551 intensity variation condition, we used the parameters of the *light source intensity variation* experiment
552 (Experiment 7, Figure 9). See Figure S7 for the predictions of the LINRF model with the same set of
553 parameters for all six conditions.

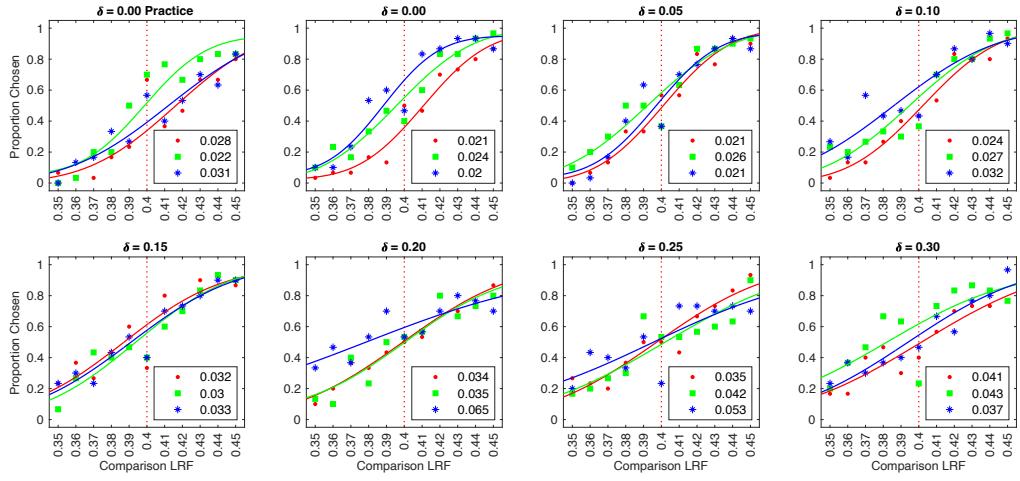
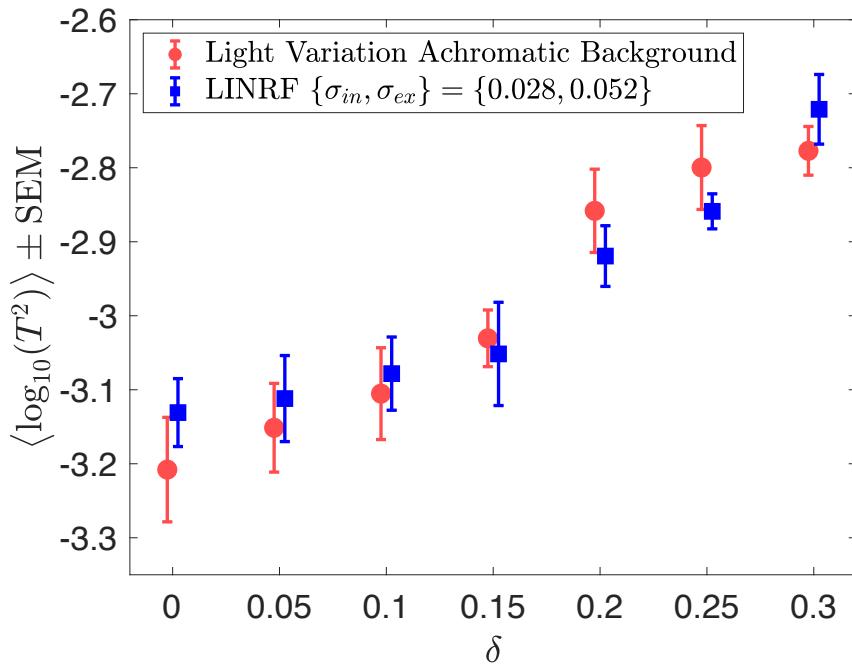
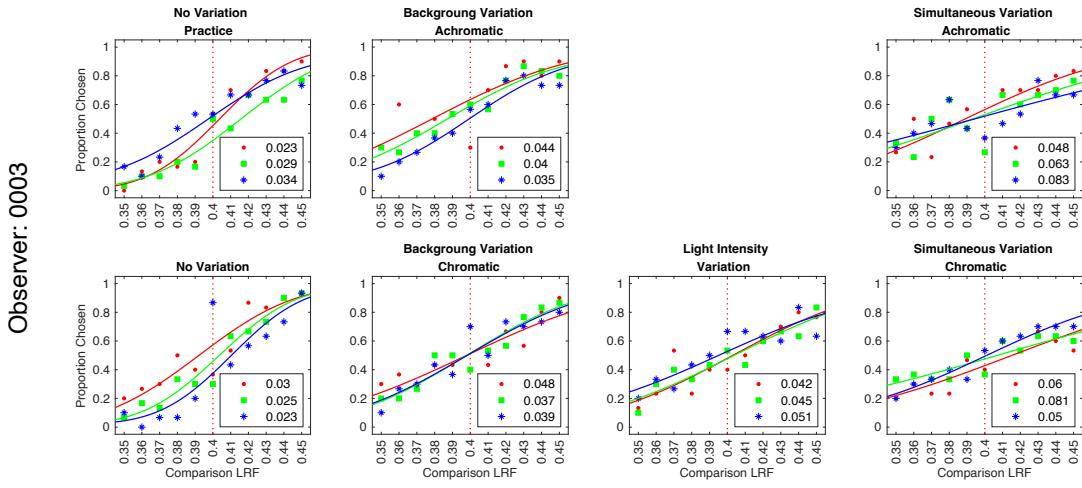
1
2
3
4
5
Observer: 0003

Figure 8: Psychometric functions for observer 0003 for light source intensity variation experiment:
Same as Figure 6, but for the *light source intensity variation* experiment. The figure shows the proportion comparison chosen data for the selection session and the seven conditions for observer 0003. The psychometric 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 estimates of the LINRF model parameters. The legend shows the parameters of the LINRF
8 model fit. The data has been jittered for ease of viewing. The data for all six observers is shown in Figure
9 S5.



1

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

554 We can use the LINRF model to compare the extrinsic variance of the simultaneous variation condition to
555 the variance of the individual variations. According to the linear receptive model, the square of the
556 threshold is proportional to the sum of the variance of observers' intrinsic noise and the extrinsic variation
557 in the stimuli (Eq. 7). The squared threshold at the no-variation condition is equal to the variance of the
558 observers' intrinsic noise. In case of extrinsic variation, the increase in threshold squared compared to the
559 no-variation condition equals the variance of the extrinsic variation. When there is more than one
560 independent source of extrinsic variation, the total variance of the simultaneous variation should be the
561 sum of the variance of the individual variations. This predicts that the increase in threshold squared for
562 simultaneous variation conditions should be equal to the sum of the corresponding increase for the
563 individual variation conditions.

564 Figure 12 shows the increase in the mean squared threshold above the no-variation condition. We
565 compare the mean squared thresholds of the simultaneous variation condition with the sum of the mean
566 squared thresholds of the individual conditions for chromatic and achromatic conditions. The increase in
567 the squared threshold of the simultaneous variation condition is comparable to the sum of the increase in
568 the squared threshold for the individual variations. The one-way ANOVA p-value of the hypothesis that
569 the mean increase in squared thresholds for simultaneous variation is equal to the sum of the mean
570 increase in the squared thresholds of light intensity variation and background object reflectance variation
571 are 0.86 and 0.80 for chromatic and achromatic conditions respectively. The variance of the extrinsic
572 noise calculated for the background variation condition ($\sigma^2 = 1.00, \delta = 0.00$) is 0.0015 and the light
573 intensity variation condition ($\sigma^2 = 0.00, \delta = 0.30$) is 0.0017. As expected, the variance of the
574 simultaneous variation condition ($\sigma^2 = 1.00, \delta = 0.30$), which is 0.0033, is comparable to the sum of
575 individual variances (0.0032).

576 4 DISCUSSION

577 The visual system is known to maintain a stable representation of object lightness despite variations in the
578 proximal signal due to the scene. In this work, we characterize the extent of such stability for variability
579 in the spectra of background objects and light sources in a scene. We measured human observers'
580 threshold of discriminating two objects based on their lightness as a function of the amount of variation in
581 these spectral properties. We observe that for low levels of variability, the thresholds first remained
582 constant, showing that in this regime the performance was determined by observers' intrinsic noise. As
583 the variability increased, the effect of extrinsic variation started dominating the performance and the
584 discrimination thresholds increased. Using a model based on signal detection theory, we related the
585 thresholds in the low variability regime to the internal noise of the observer. The model also related the
586 increase in threshold to the amount of variability in the extrinsic property, thus providing a comparison of
587 the variance in the extrinsic property to the intrinsic noise. The effects of both types of extrinsic variation,
588 the spectra of background objects and the intensity of light sources, were comparable to the effect of
589 intrinsic noise, showing that the visual system provides a good degree of constancy to these variations.
590 Further, for simultaneous variation of these properties, the effects add linearly, resulting in the variance of
591 the simultaneous variation to be equal to the sum of the variance of the individual variations.

592 **Chromatic v/s Achromatic Variations:** Lightness discrimination thresholds of chromatic and
593 achromatic variation in the reflectance spectra of background objects were statistically similar. The
594 chromatic aspect of the variation does not seem to influence lightness discrimination, indicating that
595 lightness and chromaticity are encoded independently. This hypothesis could be tested by measuring
596 chromaticity discrimination thresholds under chromatic and achromatic variation of background objects.

597 **Visual system at threshold level:** For the spectral variabilities studied in this work, the variances in
598 observers' representation of lightness due to extrinsic variations are within a factor of two compared to
599 the variances in observers' intrinsic representation of lightness. If these variances were equal, one could

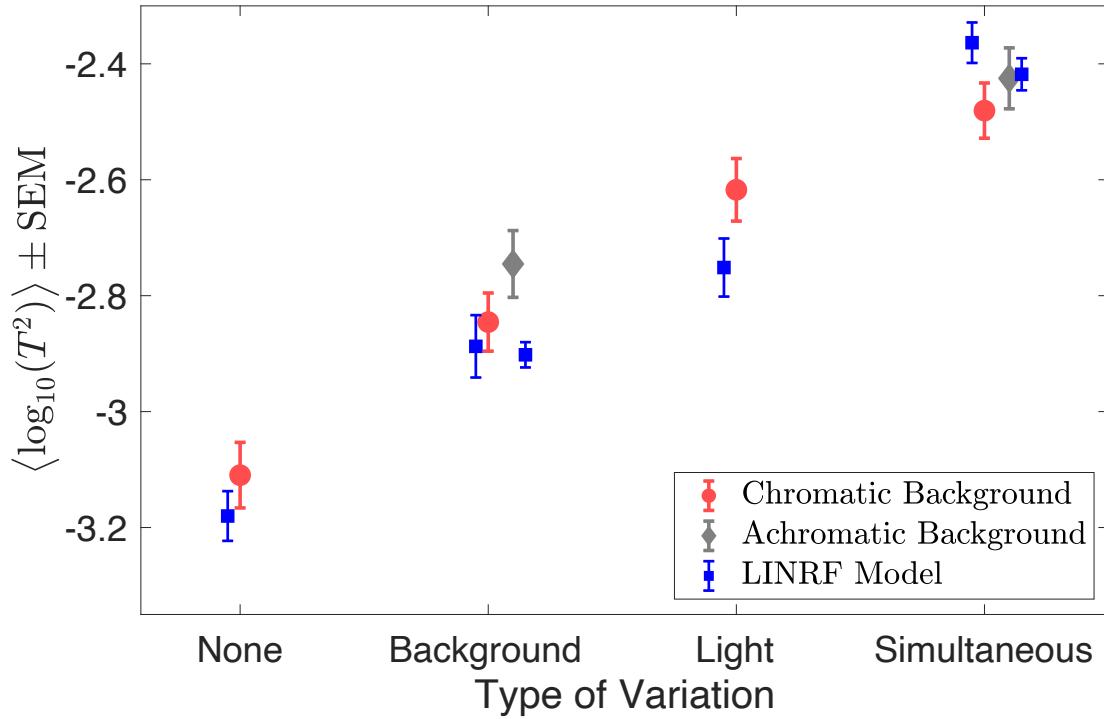
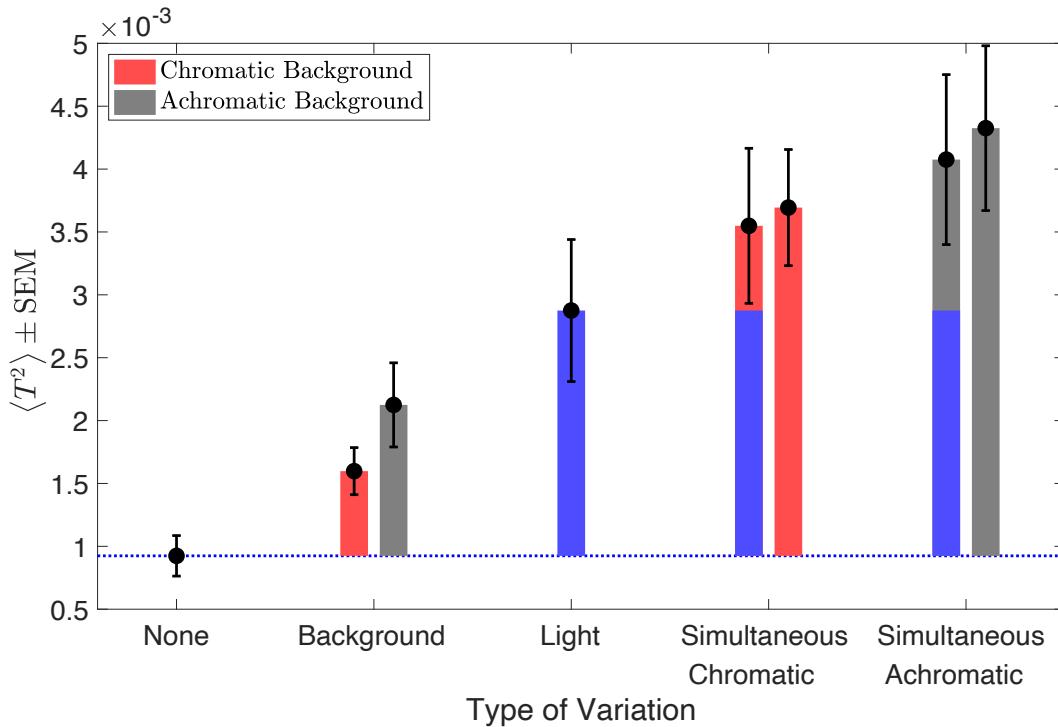


Figure 11: Discrimination thresholds for simultaneous variation of two sources are higher than individual discrimination thresholds. Mean ($N = 6$) log squared threshold for the six conditions in simultaneous variation experiment. The error bars represent ± 1 SEM taken between observers. The data for 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* experiment (Figure 9) for the Light condition. The blue error bars show ± 1 standard deviation estimated over 10 independent estimates of the LINRF model parameters. 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 adds 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 the measured squared threshold and the bars on the left (stacked bars of two different colors)
 8 represent the increase in the sum of the squared threshold of the light intensity variation (blue bar) and the
 9 corresponding background variation conditions (red or gray).

600 conclude that the visual system has fully compensated for the extrinsic variation. As the extrinsic
601 variances are larger than the variance of the intrinsic noise, the visual system has not fully compensated
602 for the external variabilities. But since these variances are within a factor or two, it shows that the visual
603 system provides a large degree of stability in the perceptual representation of lightness and seems to work
604 at near-threshold levels.

605 **Rules of Combination:** The increase in the squared threshold of simultaneous variation of reflectance
606 spectra of background object and intensity of light sources from the no-variation condition was equal to
607 the sum of the increase in the squared threshold of the individual variations. This could be accounted for
608 by assuming that the sources of noise are independent, and their effects add linearly.

609 **5 ACKNOWLEDGEMENTS:** We thank Prof. David Brainard for his comments on the manuscript. This
610 work was supported by the National Science Foundation award BCS 2054900.

611
612 **Table S1: Thresholds for Background Reflectance Variation Experiment (Preregistered**
613 **Experiment 6):**

614 Mean threshold (averaged over blocks) \pm SEM of six human observers for nine background variation conditions
studied in Background reflectance variation (Experiment 6).

Condition	Observer					
	0003	Bagel	Committee	Content	Observer	Revival
$\sigma^2 = 0.00$	0.0221 \pm 0.0010	0.0185 \pm 0.0018	0.0344 \pm 0.0027	0.0223 \pm 0.0012	0.0311 \pm 0.0053	0.0251 \pm 0.0023
$\sigma^2 = 0.01$	0.0215 \pm 0.0009	0.0194 \pm 0.0020	0.0386 \pm 0.0103	0.0193 \pm 0.0012	0.0263 \pm 0.0059	0.0262 \pm 0.0048
$\sigma^2 = 0.03$	0.0242 \pm 0.0019	0.0261 \pm 0.0020	0.0285 \pm 0.0029	0.0246 \pm 0.0046	0.0292 \pm 0.0007	0.0282 \pm 0.0016
$\sigma^2 = 0.03$ Achromatic	0.0255 \pm 0.0019					
		0.0213 \pm 0.0024	0.0343 \pm 0.0055	0.0227 \pm 0.0023	0.0267 \pm 0.0040	0.0263 \pm 0.0016
$\sigma^2 = 0.10$	0.0278 \pm 0.0015	0.0238 \pm 0.0010	0.0284 \pm 0.0017	0.0278 \pm 0.0035	0.0335 \pm 0.0024	0.0281 \pm 0.0013
$\sigma^2 = 0.30$	0.0348 \pm 0.0025	0.0277 \pm 0.0024	0.0344 \pm 0.0020	0.0286 \pm 0.0002	0.0277 \pm 0.0019	0.0301 \pm 0.0038
$\sigma^2 = 0.30$ Achromatic	0.0333 \pm 0.0032					
		0.0284 \pm 0.0028	0.0319 \pm 0.0047	0.0308 \pm 0.0015	0.0358 \pm 0.0030	0.0287 \pm 0.0022
$\sigma^2 = 1.00$	0.0416 \pm 0.0072	0.0316 \pm 0.0008	0.0379 \pm 0.0024	0.0323 \pm 0.0022	0.0405 \pm 0.0042	0.0360 \pm 0.0055
$\sigma^2 = 1.00$ Achromatic	0.0289 \pm 0.0017					
		0.0310 \pm 0.0015	0.0391 \pm 0.0029	0.0384 \pm 0.0058	0.0312 \pm 0.0015	0.0322 \pm 0.0009

615 **Table S2. Thresholds for Light Source Intensity Variation Experiment (Preregistered Experiment**
616 **7):**

617 Mean threshold (averaged over blocks) \pm SEM of six human observers measured for seven lightness intensity
618 conditions studied in Light source intensity variation (Experiment 7). The thresholds of observer Oven were not
619 used in Figure 9.

Condition	Observer					
	0003	Bagel	Oven	Content	Primary	Revival
$\delta = 0.00$	0.0217 \pm 0.0012	0.0181 \pm 0.0001	0.0520 \pm 0.0114	0.0208 \pm 0.0014	0.0329 \pm 0.0061	0.0372 \pm 0.0008
$\delta = 0.05$	0.0228 \pm 0.0018	0.0229 \pm 0.0018	0.0580 \pm 0.0064	0.0207 \pm 0.0007	0.0346 \pm 0.0042	0.0364 \pm 0.0013
$\delta = 0.10$	0.0275 \pm 0.0024	0.0217 \pm 0.0009	0.0325 \pm 0.0022	0.0242 \pm 0.0040	0.0343 \pm 0.0013	0.0376 \pm 0.0072
$\delta = 0.15$	0.0316 \pm 0.0009	0.0238 \pm 0.0011	0.0333 \pm 0.0019	0.0323 \pm 0.0032	0.0345 \pm 0.0042	0.0326 \pm 0.0002
$\delta = 0.20$	0.0447 \pm 0.0100	0.0381 \pm 0.0046	0.0493 \pm 0.0120	0.0276 \pm 0.0016	0.0423 \pm 0.0050	0.0392 \pm 0.0034
$\delta = 0.25$	0.0433 \pm 0.0052	0.0393 \pm 0.0062	0.0461 \pm 0.0060	0.0308 \pm 0.0023	0.0532 \pm 0.0083	0.0387 \pm 0.0025
$\delta = 0.30$	0.0404 \pm 0.0018	0.0429 \pm 0.0033	0.0580 \pm 0.0061	0.0347 \pm 0.0014	0.0465 \pm 0.0047	0.0421 \pm 0.0042

620

621
622
623

Table S3. Thresholds for Simultaneous Variation Experiment (Preregistered Experiment 8):
 Mean threshold (averaged over blocks) \pm SEM of six human observers measured for six conditions studied in
 preregistered experiment 8.

Condition	Observer					
	0003	Bagel	Oven	Content	Manos	Revival
No Variation	0.0261 \pm 0.0022	0.0227 \pm 0.0019	0.0383 \pm 0.0066	0.0246 \pm 0.0004	0.0258 \pm 0.0036	0.0366 \pm 0.0085
Background Variation Chromatic	0.0414 \pm 0.0036	0.0340 \pm 0.0058	0.0498 \pm 0.0050	0.0392 \pm 0.0083	0.0306 \pm 0.0013	0.0383 \pm 0.0033
Background Variation Achromatic	0.0394 \pm 0.0027	0.0319 \pm 0.0015	0.0683 \pm 0.0048	0.0427 \pm 0.0074	0.0435 \pm 0.0071	0.0389 \pm 0.0010
Light Intensity Variation	0.0464 \pm 0.0027	0.0656 \pm 0.0208	0.0592 \pm 0.0091	0.0412 \pm 0.0021	0.0464 \pm 0.0046	0.0474 \pm 0.0069
Simultaneous Variation Chromatic	0.0635 \pm 0.0092	0.0536 \pm 0.0014	0.0639 \pm 0.0106	0.0437 \pm 0.0011	0.0768 \pm 0.0085	0.0528 \pm 0.0037
Simultaneous Variation Achromatic	0.0648 \pm 0.0103	0.0540 \pm 0.0017	0.0826 \pm 0.0166	0.0478 \pm 0.0049	0.0749 \pm 0.0082	0.0561 \pm 0.0028

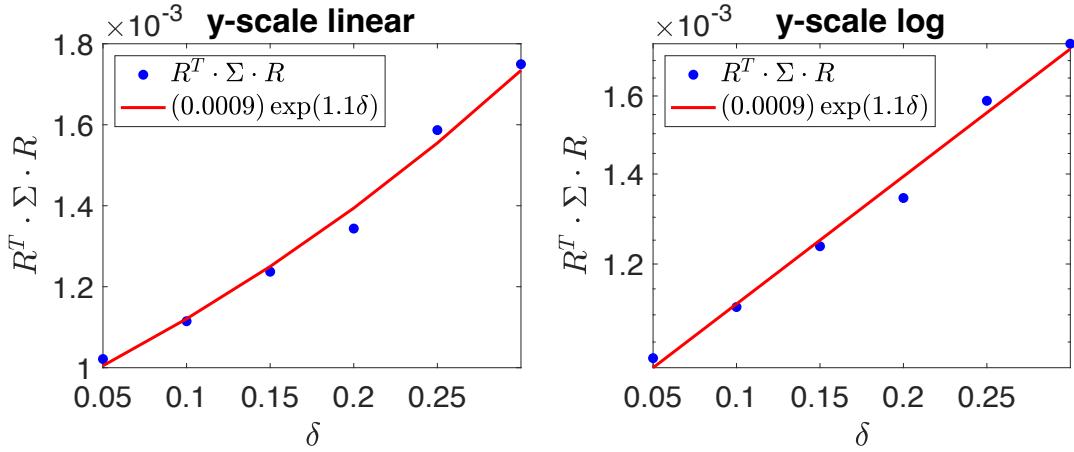
624

760 parameters. The parameters of the *background reflectance variation* condition (blue squares) predict the
761 thresholds of the no-variation condition, the background reflectance variation condition, and the
762 simultaneous variation condition quite well, but fail to predict the threshold of the light source intensity
763 variation condition. Similarly, the parameters of the *light source intensity variation* experiment (black
764 triangles) predict the thresholds of the no-variation condition, the light source intensity variation
765 condition, and the simultaneous variation condition quite well, but fail to predict the threshold of the
766 background variation condition. This could possibly be because the observers in the three experiments
767 were different. Future work would aim at studying these conditions using the same set of observers.

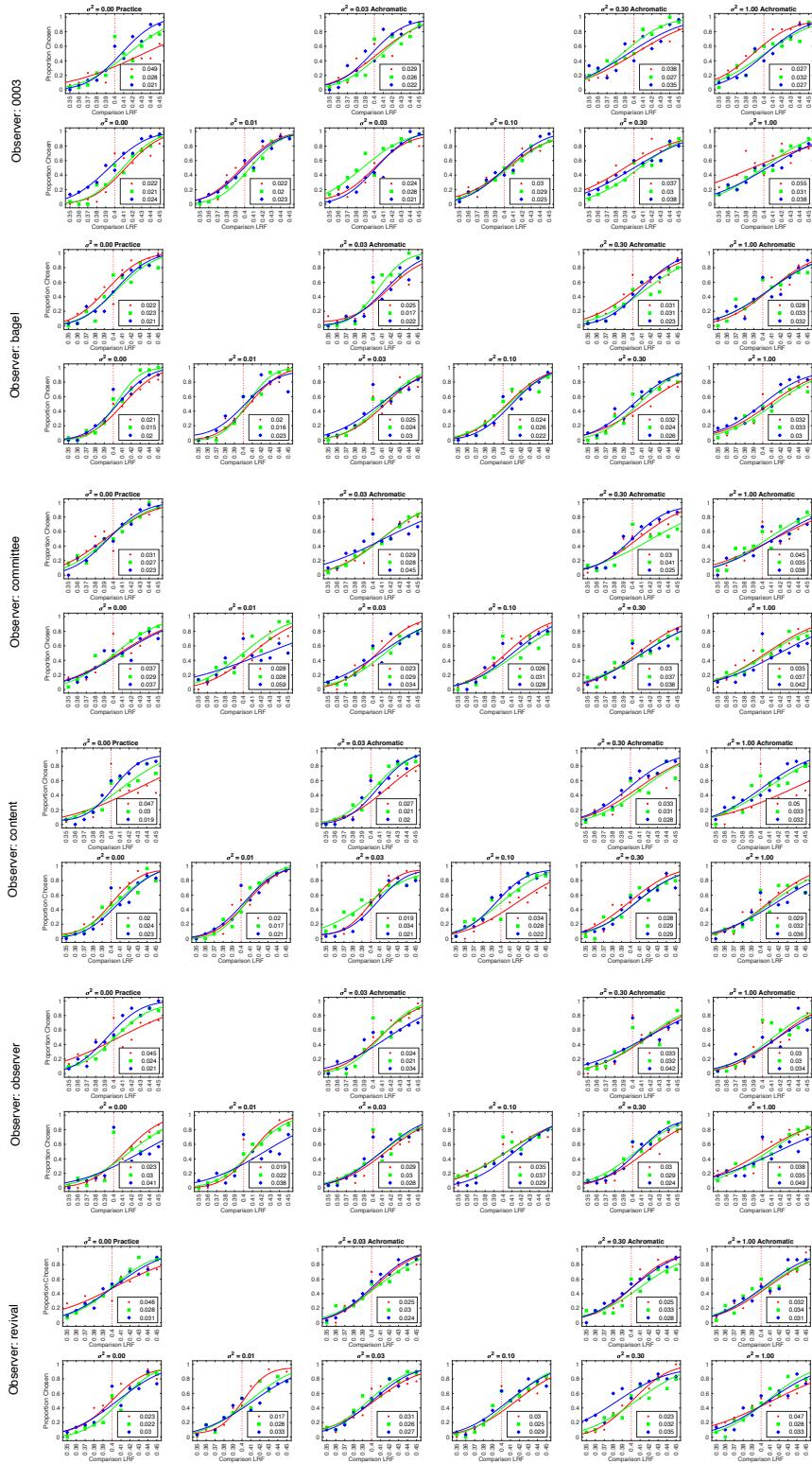
768 REFERENCES

- 769
770 Adelson, E. (2000). Lightness Perception and Lightness Illusions. In M. S. Gazzaniga, *The New*
771 *Cognitive Neurosciences* (pp. 339-351). Cambridge, MA: The MIT Press.
772 American Society for Testing and Materials. (2017). *Standard test method for luminous*
773 *reflectance factor of acoustical materials by use of integrating-sphere reflectometers.*
774 Arend, L. E. (1993). How much does illuminant color affect unattributed colors? *Journal of the*
775 *Optical Society of America*, 10(10), 2134-2147.
776 Arend, L. E., & Goldstein, R. (1987). Simultaneous constancy, lightness, and brightness. *Journal*
777 *of the Optical Society of America A*, 4(12), 2281-2285.
778 Arend, L. E., & Goldstein, R. (1990). Lightness and brightness over spatial illumination
779 gradients. *Journal of the Optical Society of America A*, 7(10), 1929-1936.
780 Arend, L. E., & Spehar, B. (1993). Lightness, brightness, and brightness contrast: 1. Illuminance
781 variation. *Perception & Psychophysics*, 54(4), 446-456.
782 Arend, L. E., & Spehar, B. (1993). Lightness, brightness, and brightness contrast: 2. Reflectance
783 variation. *Perception & Psychophysics*, 54(4), 457-468.
784 Aston, S., Radonjic, A., Brainard, D. H., & Hurlbert, A. C. (2019). Illumination discrimination
785 for chromatically biased illuminations: Implications for color constancy. *Journal of*
786 *Vision*, 19(3), 1-15.
787 Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal*
788 *of the Optical Society of America A*, 15(2), 307-325.
789 Brainard, D. H., & Radonjic, A. (2004). Color constancy. *The visual neurosciences.*, 1, 948-961.
790 Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). Color constancy in the nearly natural
791 image. Asymmetric matches. *Journal of the Optical Society of America A*, 14(9), 2091-
792 2110.
793 Bramwell, D. I., & Hurlbert, A. C. (1996). Measurements of colour constancy by using a forced-
794 choice matching technique. *Perception*, 25(2), 229-241.
795 Craven, B. J., & Foster, D. H. (1992). An operational approach to colour constancy. *Vision*
796 *Research*, 32(7), 1359-1366.
797 Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the
798 statistical regularity of natural daylight? *Journal of Vision*, 4(2), 57-81.
799 Ennis, R., & Doerschner, K. (2019). Disentangling simultaneous changes of surface and
800 illumination. *Vision Research*, 158, 173-188.
801 Foster, D. H. (2003). Does colour constancy exist? *Trends in cognitive sciences*, 7(10), 439-443.
802 Foster, D. H. (2011). Color constancy. *Vision Research*, 51(7), 674-700.
803 Gilchrist, A. (2006). *Seeing black and white*. Oxford University Press.
804 Green, D. M., & Swets, J. A. (1995). *Signal Detection Theory and Psychophysics* (Vol. 1). New
805 York: Wiley.
806 Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal context on
807 color categories and color constancy. *Journal of Vision*, 7(4), 1-15.
808 Ishihara, S. (1977). *Tests for colour-blindness*. Tokyo: Kanehara Shuppon Company, Ltd.
809 Jakob, W. (2010). Mitsuba Renderer.
810 Kelly, K. L., Gibson, K. S., & Nickerson, D. (1943). Tristimulus specification of the Munsell
811 book of color from spectrophotometric measurements. *Journal of the Optical Society of*
812 *America*, 33(7), 355-376.

- 813 Linhares, J. M., Pinto, P. D., & Nascimento, S. M. (2008). The number of discernible colors in
814 natural scenes. *Journal of the Optical Society of America*, 25(12), 2918-2924.
- 815 Luo, M. R., Clarke, A. A., Rhodes, P. A., Schappo, A., Scrivener, S. A., & Tait, C. J. (1991).
816 Quantifying colour appearance. Part I. LUTCHI colour appearance data. *Color Research
817 & Application*, 16(3), 166-180.
- 818 Olkkonen, M., & Ekroll, V. (2016). Color constancy and contextual effects on color appearance.
819 In J. Kremers, R. Baraas, & N. Marshall, *Human color vision, Springer Series in Vision
820 Research Vol. 5* (pp. 159-188). Springer, Cham.
- 821 Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for
822 simulated surfaces. *Journal of Vision*, 9(12), 1-6.
- 823 Olkkonen, M., Witzel, C., Hansen, T., & Gegenfurtner, K. R. (2010). Categorical color
824 constancy for real surfaces. *Journal of Vision*, 10(9), 1-16.
- 825 Patel, K. Y., Munasinghe, A. P., & Murray, R. F. (2018). Lightness matching and perceptual
826 similarity. *Journal of vision*, 18(5), 1-13.
- 827 Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G. D., & Hurlbert, A. (2014). Chromatic
828 illumination discrimination ability reveals that human colour constancy is optimised for
829 blue daylight illuminations. *Plos One*, 9(2), e87989.
- 830 Reeves, A. J., Amano, K., & Foster, D. H. (2008). Color constancy: phenomenal or projective?
831 *Perception & psychophysics.*, 70(2), 219-228.
- 832 Rutherford, M. D., & Brainard, D. H. (2002). Lightness constancy: A direct test of the
833 illumination-estimation hypothesis. *Psychological Science*, 13(2), 142-149.
- 834 Schultz, S., Doerschner, K., & Maloney, L. T. (2006). Color constancy and hue scaling. *Journal
835 of Vision*, 6(10), 1-10.
- 836 Singh, V., Burge, J., & Brainard, D. H. (2022). Equivalent noise characterization of human
837 lightness constancy. *Journal of Vision*, 22(5), 1-26.
- 838 Singh, V., Cottaris, N. P., Heasly, B. S., Brainard, D. H., & Burge, J. (2018). Computational
839 luminance constancy from naturalistic images. *Journal of Vision*, 18(3), 1-19.
- 840 Smithson, H., & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and
841 levels of reference. *Journal of Vision*, 4(9), 1-3.
- 842 Speigle, J. M., & Brainard, D. H. (1996). Is color constancy task independent? *Color and
843 Imaging Conference. 1*, pp. 167-172. Society for Imaging Science and Technology.
- 844 Troost, J. M., & De Weert, C. M. (1991). Naming versus matching in color constancy.
845 *Perception & psychophysics*, 50, 591-602.
- 846 Uchikawa, K., Uchikawa, H., & Boynton, R. M. (1989). Partial color constancy of isolated
847 surface colors examined by a color-naming method. *Perception*, 18(1), 83-91.
- 848 Vrhel, M. J., Gershon, R., & Iwan, L. S. (1994). Measurement and analysis of object reflectance
849 spectra. *Color Research & Application*, 19(1), 4-9.
- 850

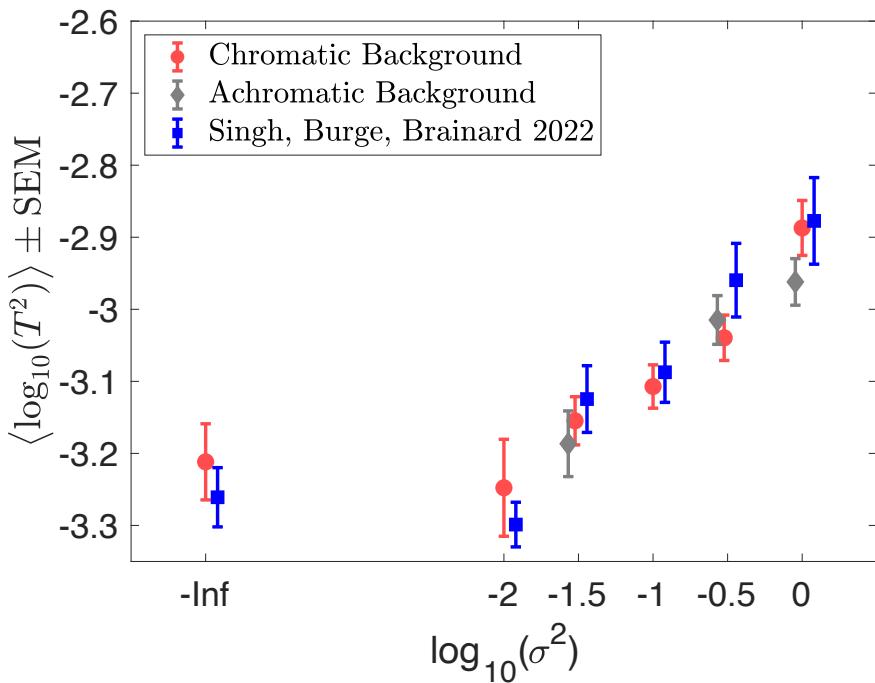


2 **Figure S1: Estimation of extrinsic noise for *light source intensity variation* experiment:** Plot of the
3 variance ($R^T \Sigma R$) as a function of the range parameter δ on a linear (left panel) and logarithmic (right
4 panel) scale. We fit the function with an exponential of the form $A * \exp(B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



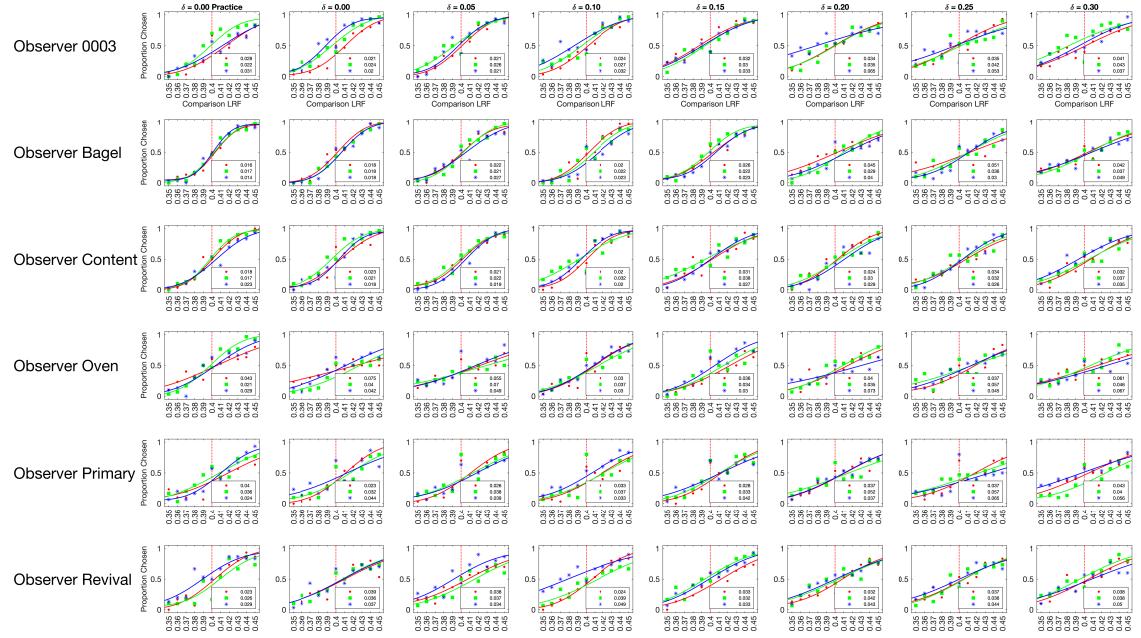
1

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



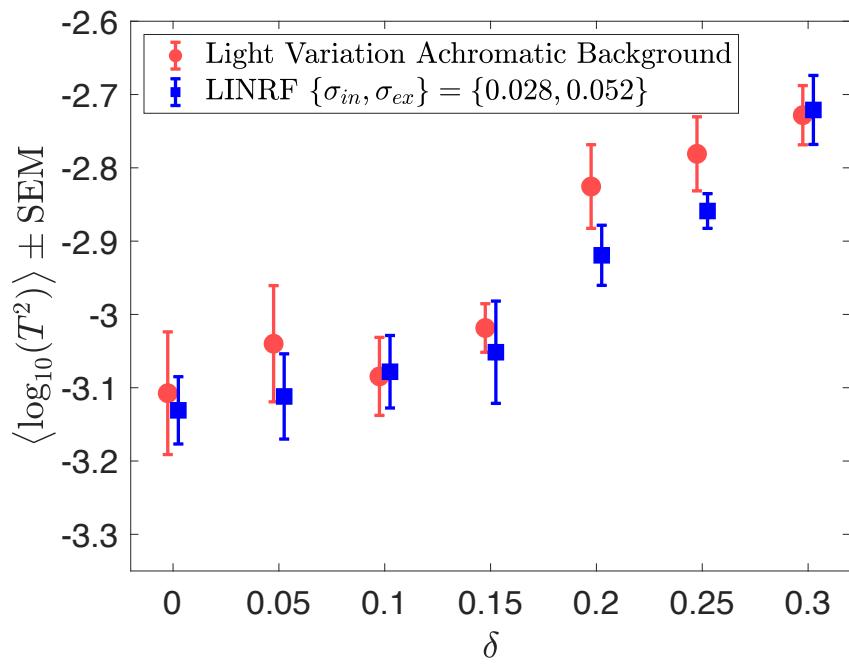
1

2 **Figure S3: Comparison with Singh, Burge, Brainard 2022.** Lightness discrimination thresholds for
3 background variation condition measured in *background reflectance variation* experiment and previously
4 reported data from Singh, Burge, Brainard (2022). The previous experiment only had chromatic
5 conditions and made three threshold measurements for each condition for 4 naïve observers. In this work,
6 *background reflectance variation* experiment had both chromatic and achromatic conditions and
7 measured thresholds for six observers. The experiments were otherwise the same.



1

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

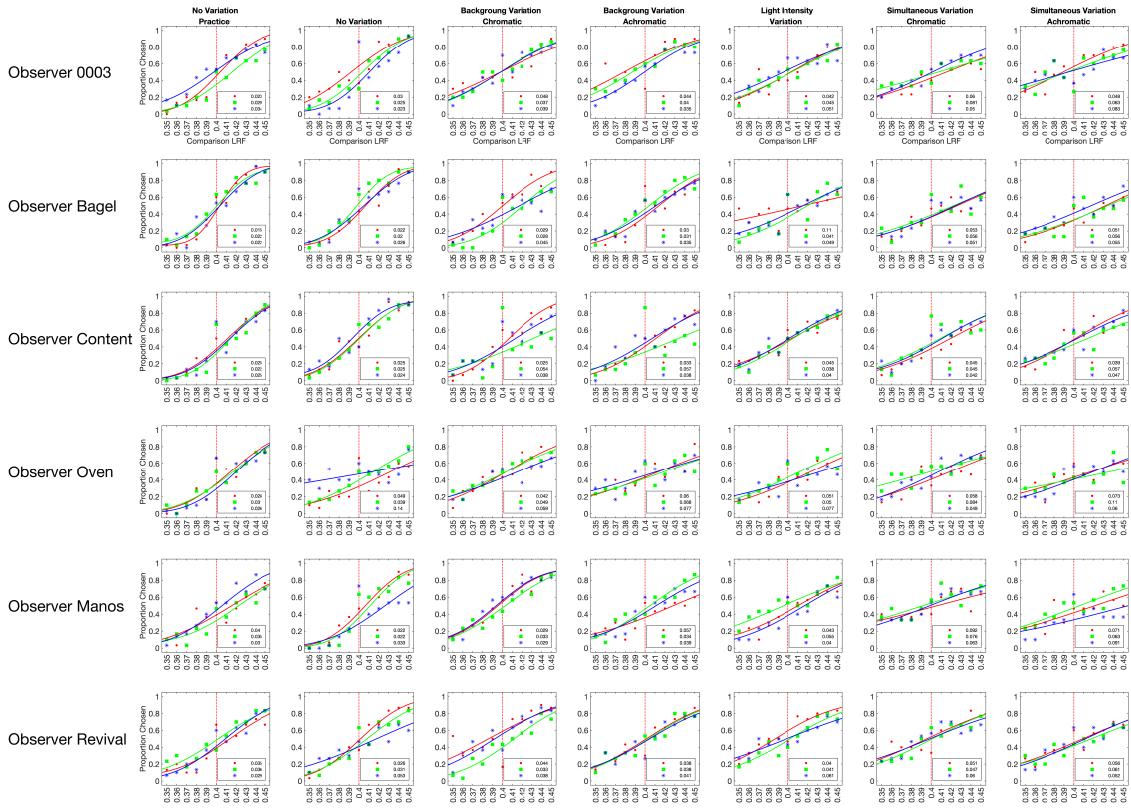


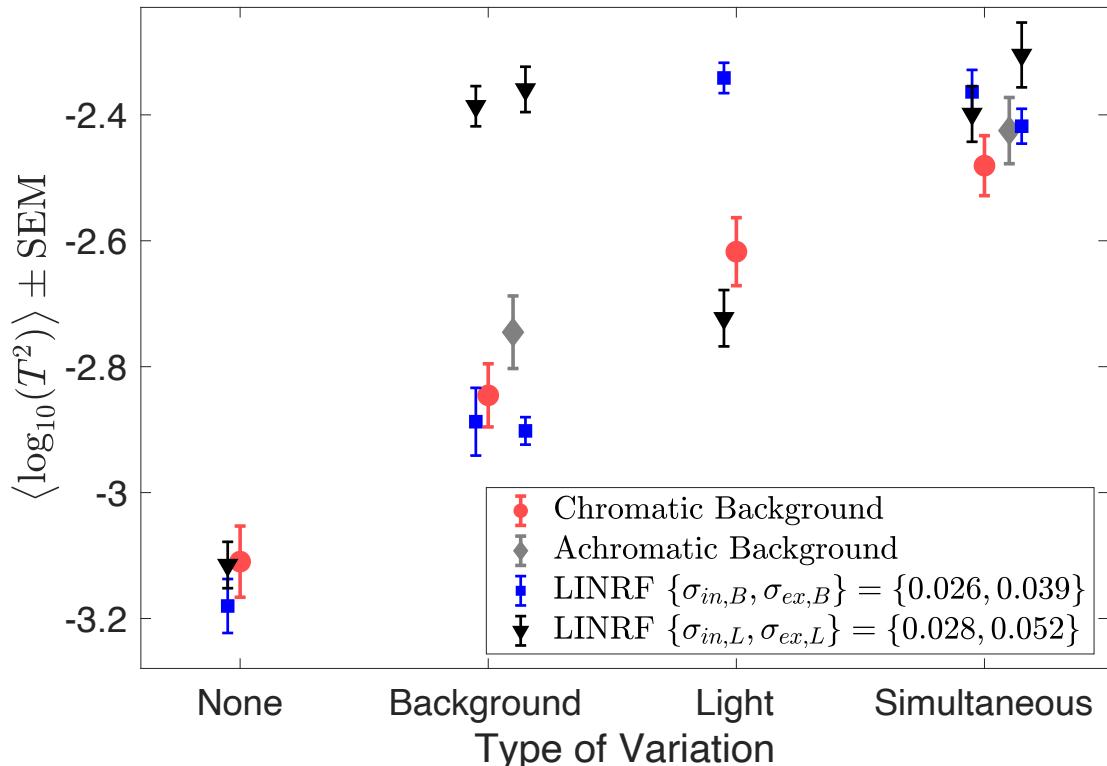
1

2 **Figure S5:** Same as Figure 9, for all six observers retained in the *light source intensity variation*
3 experiment. The 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.** Similar
3 to Figure 10, for all observers retained in the *simultaneous variation* experiment.





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