Experiments with color scrambles reveal Red and Green half-axis mechanisms plus a Gray-tuned mechanism

Christian Herrera[†], Charles Chubb[‡], Charles E. Wright[‡], Peng Sun, [‡], George Sperling[‡]
† VA Loma Linda Healthcare System
‡ Department of Cognitive Sciences, UC Irvine

Abstract

[abstract goes here]

KEYWORDS: Luminance, Equiluminance, Color, Saturation,

1 Introduction

What color-selective mechanisms exist in human vision? This question has been the focus of a great deal of research effort, and yet no clear consensus has emerged. Substantial evidence supports the theory, originally proposed by Krauskopf et al. [1982] that human vision comprises three basic post-retinal mechanisms [Krauskopf et al., 1986, Krauskopf and Gegenfurtner, 1992, Krauskopf et al., 1996a,b, Gegenfurtner and Kiper, 1992, Webster and Mollon, 1991, 1994]. It was proposed that these three mechanisms combine the responses of the short, medium and long-wavelength cone classes, S, M and L additively as follows:

1. The *luminance* mechanism sums the responses of the M and L cones: $M + L$.
2. The *blue-yellow* mechanism takes the difference between the S-cone response and the summed M and L cone responses: $S - (M + L)$.
3. The *red-green* mechanism takes the difference between the L- vs. M-cone responses: $L - M$.

The weights with which the different cone-class responses were combined by these three mechanisms were not well-constrained by the data available to Krauskopf et al. [1982]. The important point was that these mechanisms combine cone-responses linearly.

Eskew [2009] emphasizes, however, that it really does not make sense to think of the red-green mechanism (a similar argument applies to the yellow-blue mechanism) as a single mechanism; rather we should think about it (at least for purposes of psychophysical analysis) as two, complementary, half-wave rectified mechanisms:

1. A *red half-axis* mechanism whose response is given by

$$[L-M]^{+} = \begin{cases} L-M & \text{if } L-M > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

2. And a *green half-axis* mechanism whose response is given by

$$[L-M]^{-} = \begin{cases} |L-M| & \text{if } L-M < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The argument supporting this contention involves two assumptions that are intrinsic to the idea of a mechanism: the *univariance* assumption and the *labeled line* assumption:

1. The univariance assumption holds that the output produced by a mechanism in response to a light is a single number: the activation produced in the mechanism by the light.

2. The labeled line assumption [Graham, 1989, Watson and Robson, 1981] holds that two lights will appear different if and only if they produce significantly different levels of activation in some mechanism.

Suppose, for example, that B is the spectrum of a background light. Let δ_1 and δ_2 be functions mapping the visible wavelengths into $\mathbb{R}$, and consider a stimulus consisting of a dot with spectrum $B+\alpha_1\delta_1$ next to a dot with spectrum $B+\alpha_2\delta_2$ on an otherwise homogeneous background with spectrum B. Suppose all mechanisms except one, that we shall call M, are blind to variations in the amplitude of each of δ_1 and δ_2. In this case, for any values of α_1 and α_2, the response of any mechanism other than M to the lights $B+\alpha_1\delta_1$ and $B+\alpha_2\delta_2$ is equal to its response to B. Suppose, however, that M is sensitive (in different degrees) to both δ_1 and δ_2. For concreteness, suppose M is more sensitive to variations in the amplitude of δ_2. The univariance assumption requires that for any α_1 there exists a value of α_2 such that the spots with spectra $B+\alpha_1\delta_1$ and $B+\alpha_2\delta_2$ appear identical. In particular, this will be true for α_2 adjusted so that the two spots produce equal activation in M. On the other hand, the labeled line assumption implies that if there exists a mechanism, M' distinct from M that is sensitive to variations in the amplitude of δ_1 but not to variations in the amplitude of δ_2, then no matter how one adjusts the amplitude α_2, there will always be a perceptual difference between the lights with spectra $B+\alpha_1\delta_1$ and $B+\alpha_2\delta_2$ because these two lights produce different levels of activation in M'.

Eskew [2009] observes that if the background light with spectrum B is gray, and the dots with spectrums $B+\alpha_1\delta_1$ and $B+\alpha_2\delta_2$ are red and green respectively, then even if the amplitudes α_1 and α_2 are both at detection threshold (under additional conditions designed to insure that performance is mediated by a single mechanism), the two dots appear different [Eskew et al., 2001, Krauskopf et al., 1986, Mullen and Kulikowski, 1990]. This leads Eskew to conclude that human vision should be seen as possessing distinct mechanisms activated by δ_1 and δ_2. He refers to such mechanisms as linear, half-axis mechanisms.

On the other hand, suppose that the $L-M$ mechanism hypothesized by Krauskopf et al. [1982] can produce negative activations. This might be accomplished in several ways. First, the mechanism might use neurons that maintain a non-zero resting firing-rate. A response of 0 would then signal that the light impinging on the retina at a given point in space has $L-M$ activation equal to that of the background to which the subject is adapted. Negative activations (indicating that the light deviates in the green direction from the adapting level) would be signaled by deviations in firing below the resting state firing level, and positive activations (indicating that the light deviates in the red direction from the adapting level) would be signaled by deviations above the resting state firing level. The problem with a neuronal architecture of this sort is that it is costly: substantial resources are required to sustain a non-zero resting-state firing rate in a large population of neurons. A different strategy would make use of separate red and green half-axis channels precisely of the sort suggested by Eskew. Under this strategy, the red half-axis channel would transmit the positive deviations of the $L-M$ mechanism away from the adapted level, and the green half-axis channel would transmit the negative deviations.

Let us suppose that human vision does possess separate red and green half-axis mechanisms. Under one scenario, top-down attention might have entirely separate access to these two mechanisms for purposes of making judgments about various images. It certainly feels,

for example, as though one can selectively attend to the greens in a painting without being influenced by the reds, and vice versa. This experience suggests separate access to mechanisms selective for red and for green. Another possibility, however, is that our hypothetical red and green half-axis mechanisms are bound together in a way that precludes separate access, with the red half-axis mechanism coding the positive activations in a single opponent mechanism and the green half-axis mechanism coding the negative activations.

Several previous studies provide support for the idea that subjects have separate access to red and green half-axis mechanisms. First, adapting to temporal modulations of chromaticity that follow a sawtooth profile (e.g., repeatedly changing gradually to red, then changing abruptly to green and changing gradually again to red, etc.) can increase thresholds differentially for the color receiving the abrupt onset vs. the gradually changing color [Krauskopf et al., 1986, Krauskopf and Zaidi, 1986]. It has also been shown that briefly flashed reddish targets are much more effectively masked by reddish than by greenish noise, and vice versa [Sankeralli and Mullen, 2001]. In addition, it has been shown that targets that deviate from gray in the red direction vs. those that deviate in he green direction are detected by different mechanisms at threshold in the fovea [Gowdy et al., 1999a,b].

The current study uses a different experimental paradigm involving discrimination of chromatic textures to seek evidence of red- vs. green-selective half-axis mechanisms. This paradigm was originally used by Silva and Chubb [2014] to analyze the mechanisms sensitive to textures composed of different grayscales. In the current application, we use textures composed of colors drawn from the red-green cardinal axis of DKL-space [Derrington et al., 1984]. The general strategy, which is sketched in Victor et al. [2017], involves testing subjects in several different tasks, each of which requires the subject to detect the location of a patch of one sort of color texture in a background of a different sort of color texture. The predominant quality of the target patch relative to the background varies across the different tasks. We assume that to perform a given task the subject constructs a task-specific "tool" by taking a linear combination of his-or-her available mechanisms. We further assume that the weights with which the mechanisms are combined are chosen to maximize the expected difference in activation produced by the target patch vs. the background texture. As we shall demonstrate, under these assumptions, the resulting data enables us to determine (1) how many mechanisms exist in human vision that are differentially sensitive to the colors used in the textures, and (2) the sensitivity of each of these mechanisms to the different colors.

2 Methods

2.1 Subjects

There were 5 subjects (2 female) one of which was the first author. Subjects (other than the first author) were naive to the purpose of the experiments. All reported normal or corrected-to-normal vision. None were color-deficient. The UC Irvine Institutional Review Board approved the experimental procedures, and all subjects provided signed consent.

2.2 Equipment

An iMac desktop computer running OS X version 10.6.8 with a 3.06 GHz Intel Core 2 Duo processor and 4 GB memory capacity was used for stimuli presentation and data collection. The computer was equipped with an ATI Radeon HD 4670 graphics chip. The monitor had a resolution of 1920 x 1080 and a viewable diagonal measure of 21.5 inches.

2.3 Color palette acquisition

For each subject, we used a minimum motion task to acquire a palette of colors motion-equiluminant with the background, and no two palettes were exactly the same. The process to acquire an individual equiluminant color palette used the following procedure. We constructed twenty sequences that followed a path around the perimeter of the set of lights achievable with our display system starting at a high-luminance light and ending at a low-luminance light. Different paths passed through different different color regions. For each of the 20 sequences, the subject first used the up and down arrows on the keyboard to move along the path to select the light C_{start} that appeared (according to a minimum boundary contrast criterion) to be equiluminant with a standard gray light G with luminance 52 cd/m^2. Then the subject performed a series experimental of trials in which he-or-she viewed an annular, 8-frame motion stimulus (each frame lasting 33ms) and judged on each trial whether it evoked clockwise vs. counterclockwise motion. This stimulus included a colored light C and was designed to evoke ambiguous motion if C was equiluminant to 52 cd/m^2 gray light. An up-down staircase was used to move along the path of lights starting at C_{start}. A cumulative normal psychometric function was fit to the resulting data to find the color along that path that was equiluminant to G.

We proceeded to find the plane in the space spanned by 3 color fundamentals of our monitor that best fit the 20 lights obtained from the subject. The direction normal to this plane was taken to be the luminance axis for the subject.

For the present experiment, we used a set of motion-equiluminant lights that isolated the L-M cardinal axis of DKL space. Specifically, the lights in this set

1. all produced equal activation in the (Stockman-Sharpe, 2 deg.) S-cone fundamental.
2. were all motion-equiluminant to the achromatic light G with photometric luminance of $52cd/m^2$.

Thus, these lights were chosen to produce differential activation only in whatever mechanisms exist in human vision sensitive to variations in the difference between L-cone vs. M-cone activations. Eight lights were used for each subject. These lights projected onto locations -4α, -3α, -2α, $-\alpha$, α, 2α, 3α, 4α, of the $L-M$ cardinal axis of DKL space, where α was chosen to be as large as possible on our display device given that location 0 corresponded to the achromatic light with photometric luminance of $52cd/m^2$. We will write Ω for the set containing these 8 colors.

It was a mistake on our part to leave a gap in our set of colors at $\alpha = 0$. The lines plotted in Figs. 1a and 1b should be straight, but they are not straight because of this error. Similarly, in Figs. 2a and 2b should be parabolic, but they are too fat because of this error. However, this mistake does not compromise the logic of the study in any important way.

2.4 Ω-scrambles

Stimuli were composed of chromatic textures called Ω-scrambles. Ω is a set, in principle of arbitrary size, whose elements will be used to compose a texture; in our case, the elements of Ω are eight colors. An Ω-scramble is a patch of texture comprising small square elements, each of which is colored with one of the colors in Ω. The "histogram" of the scramble is the probability distribution that gives the proportions with which the different colors in Ω appear in the patch. In practice, to generate a patch of Ω-scramble comprising N spatial elements with histogram as close as possible to p, one loads a virtual urn with N copies of colors drawn from Ω in proportions conforming as closely as possible to p. These colors are then assigned randomly without replacement to the N locations of the patch.

We will write U for the uniform histogram (i.e., $U(\omega) = \frac{1}{n}$ for all $\omega \in \Omega$, where n is the size of Ω. In addition, we call any function $\rho : \Omega \to \mathbb{R}$ a perturbation if

1. $\sum_{\omega\in\Omega} \rho(\omega) = 0$, and
2. $|\rho(\omega)| < \frac{1}{8}$ for all $\omega \in \Omega$.

These two conditions insure that $U + \rho$ and $U - \rho$ will both be probability distributions on Ω. If in fact $|\rho(\omega)| = \frac{1}{8}$ for some $\omega \in \Omega$, we call ρ a maximal perturbation.

We assume that human vision has some number M of mechanisms that are differentially sensitive to Ω-scrambles, and for $m = 1, 2, \cdots, M$, the sensitivity of mechanism m can be characterized by a function $F_m(\omega)$ that reflects the activation produced in mechanism m by a texture element of color ω. The space-average activation produced in mechanism m by an Ω-scramble with histogram p is equal to the dot product of f and p:

$$f \bullet p = \sum_{\omega\in\Omega} F_m(\omega)p(\omega), \tag{3}$$

The difference in activation produced in mechanism m by scrambles with histograms p and q is $F_m \bullet (p - q)$. When $p = U + \rho$ and $q = U - \rho$, the difference in activation is $2F_m \bullet \rho$.

The goal of the current experiment is to determine the sensitivity functions of all of the mechanisms in human vision that are differentially sensitive to Ω-scrambles.

2.5 The basic task

A given stimulus comprised an annulus of Ω-scramble with some histogram $U - \rho$ which contained a disk of Ω-scramble with histogram $U + \rho$ at one of eight possible locations (corresponding to map directions N, NE, E, SE, S, SW W, NW), and the task was to guess the target location. Examples of stimuli are shown in Figs. 1 and 2. Subjects viewed the display from an unrestrained distance of approximately 95 cm. At this distance the outer diameter of the stimulus annulus subtended 6.8° of visual angle, and each texture element subtended 0.11° of visual angle.

A central cue spot that was slightly brighter than the mean-gray background remained visible continuously throughout the experiment. Each trial was initiated with a button press. There followed a 500 ms. gray screen. The stimulus then appeared for 300 ms and was followed by the gray screen. After the display, the subject used the number pad keys

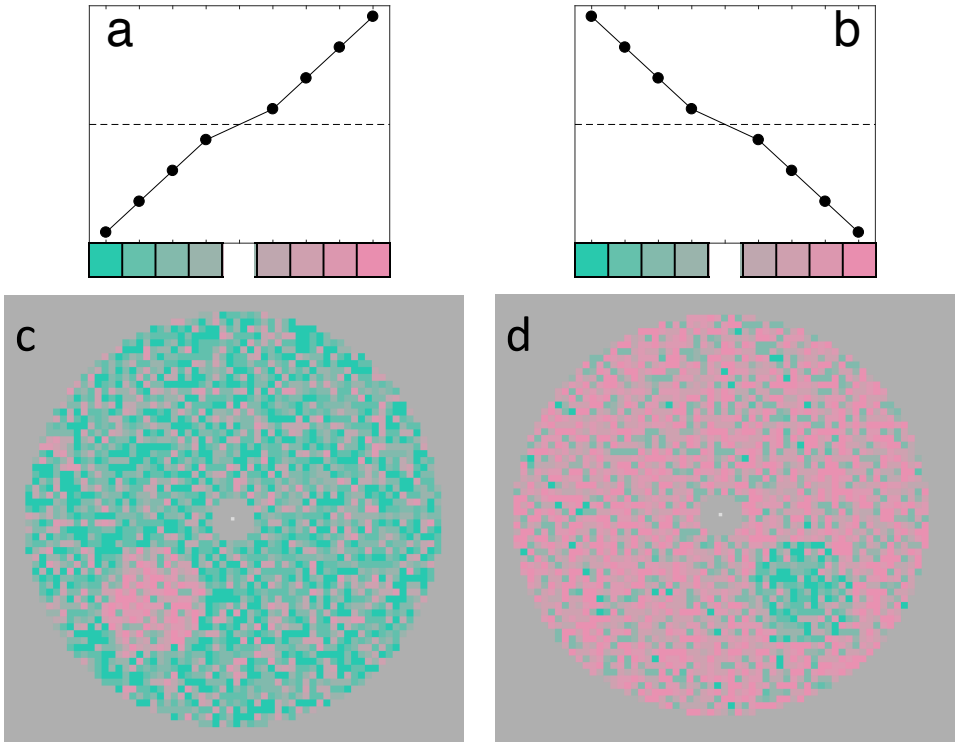

Figure 1: *Examples of stimuli from complementary conditions 1 and 2.* The perturbations (a) λ_1, and (b) $-\lambda_1$. The example stimulus in (c) has a target disk that is composed of Ω-scramble with histogram $U + A\lambda_1$, and a background annulus with histogram $U - A\lambda_1$, where the histogram amplitude A is chosen to make the perturbation $A\lambda_1$ maximal. The task is to indicate the location (using the keys in the number pad of the target disk). The roles of two different types of scramble have been reversed in the example stimulus in (d). In condition 1 (condition 2), each stimulus has a target disk with histogram $U + \rho$ and background annulus with histogram $U - \rho$ for some perturbation ρ that correlates strongly with λ_1 ($-\lambda_1$). In order to clearly indicate the nature of the search task in which the subject was engaged, the differences between the targets and backgrounds shown here are much stronger than they were in the actual experiment.

to indicate the location of the target disk. The mapping was: "7" for NW, "8" for N, "9" for NE, "6" for E, "3" for SE, "2" for S, "1" for SW, "4" for W. A beep sounded after any incorrect response.

2.6 Experimental conditions

Each subject performed 2600 trials in each of four, separately blocked tasks. These four tasks all required judgments of target location as described in Section 2.5. However, the predominant quality that differentiated the target from the background differed in different tasks. In task 1 (task 2), the target was pinker (greener) than the background (as can be seen in Fig. 1, c and d). In task 3 (task 4), the target was higher (lower) in red-green contrast than the background (as can be seen in Fig. 2, c and d). Each of these tasks was

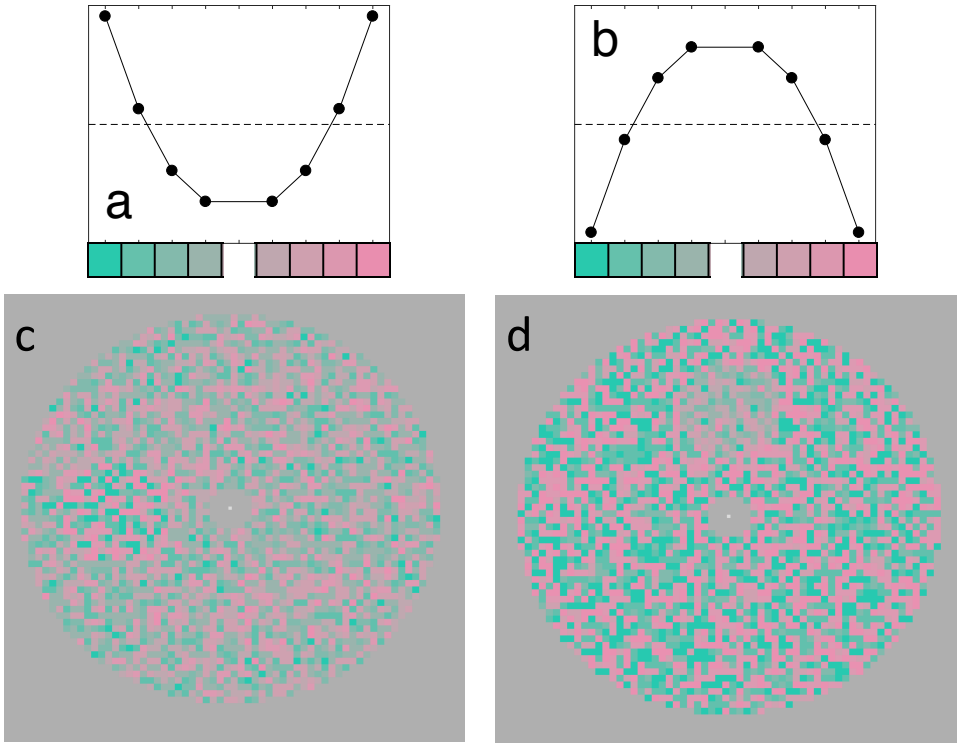

Figure 2: *Examples of stimuli from complementary conditions 3 and 4.* The perturbations (a) λ_2, and (b) $-\lambda_2$. The example stimulus in (c) has a target disk that is composed of Ω-scramble with histogram $U + A\lambda_2$, and a background annulus with histogram $U - A\lambda_2$, where the histogram amplitude A is chosen to make the perturbation $A\lambda_2$ maximal. The task is to indicate the location (using the keys in the number pad of the target disk). The roles of two different types of scramble have been reversed in the example stimulus in (d). In condition 3 (condition 4), each stimulus has a target disk with histogram $U + \rho$ and background annulus with histogram $U - \rho$ for some perturbation ρ that correlates strongly with λ_2 ($-\lambda_2$). In order to clearly indicate the nature of the search task in which the subject was engaged, the differences between the targets and backgrounds shown here are stronger than they were in the actual experiment.

an individual application of the "seed expansion" method [Chubb et al., 2012]. The next section gives a brief overview of the method as it was used in one of the four tasks in the current experiment.

2.6.1 The use of the seed expansion method in the current experiment

For task $t = 1, 2, 3, 4$, a single dominant perturbation ϕ_t is used to produce the difference between the target disk vs the background annulus on each trial. We call ϕ_t the *seed* perturbation for task t. On any given trial in the condition with seed ϕ_t, the Ω-scramble filling the target disk will have a histogram $U + \rho$ for some perturbation ρ correlated strongly and positively with ϕ_t (all correlations are at least 0.894), and the annular background will have histogram $U - \rho$. Thus, the qualitative difference between the target-disk vs the background

will be similar from trial to trial. The point of separately blocking these different tasks is to encourage the subject to optimize his/her strategy for the particular target-vs-background texture difference in each task. In particular, our model assumes that in a given task t, a given subject s combines information from his/her mechanisms to produce a task-specific tool $T_{s,t}$ that gives high (low) values to colors $\omega \in \Omega$ that occur with high (low) density in the target disk (background annulus) in task t. We assume that the subject applies the tool $T_{s,t}$ to the stimulus on a given trial to produce a neural map in which the target location is highly activated in comparison to the background. In Chubb et al. [2012], the function $T_{s,t}$ is called the "expansion" of the seed perturbation ϕ_t. Here we will call $T_{s,t}$ the "tool achieved by subject s in task t".

Without making any assumptions about the nature of the mechanisms underlying performance, we can use a simple model to describe the tool $T_{s,t}$ achieved by subject s in task t. This model assumes:

1. The salience of the target in task t on a trial in which the target has histogram $U + \rho$ (and the background has histogram $U - \rho$) is

$$\mathrm{Sal}_{s,t}(\rho) = T_{s,t} \bullet \rho. \tag{4}$$

2. The probability of a correct response on such a trial is $\Psi_s(\mathrm{Sal}(\rho))$, for the Ψ_s the Weibull function defined by:

$$\Psi_s(x) = 0.125 + 0.855\left(1 - \exp\left(-x^{B_s}\right)\right). \tag{5}$$

   Note that for any given subject s, we assume that the steepness parameter B_s is constant across different tasks. The estimated tools do not change appreciably if we allow the Weibull steepness parameters to vary across tasks for a given subject.

About Ψ_s, notice that

1. Chance performance is $\Psi_s(0) = 0.125$ because the task requires an 8-option forced choice.

2. Ψ_s asymptotes at 0.98 instead of 1.0 to cover the possibility of "finger errors," i.e., errors subjects make make even though they clearly see correct response.

3. Typically, one expects a Weibull function to have two free parameters. However, there is only one (B_s) in Eq. 5. This is because the other typical parameter (a scalar used to divide x) can be absorbed into the tool $T_{s,t}$ in Eq. 4.

2.6.2 The four tasks

To describe the perturbations used in these experiments, we identify the 8 colors $\omega \in \Omega$ ranging from red to green with the values $v_1, v_2, \cdots, v_8$ equal to $-1, -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$. The Legendre polynomials are derived by applying Gram-Schmidt orthonormalization to the sequence of monomials $h_j(v) = v^j$, $j = 0, 1, \cdots, 7$. The Legendre polynomials of order $1, 2, \cdots, 7$ are listed in table 1.

k	$\lambda_k(v_1)$	$\lambda_k(v_2)$	$\lambda_k(v_3)$	$\lambda_k(v_4)$	$\lambda_k(v_5)$	$\lambda_k(v_6)$	$\lambda_k(v_7)$	$\lambda_k(v_8)$
1	-0.5401	-0.3858	-0.2315	-0.0772	0.0772	0.2315	0.3858	0.5401
2	0.5401	0.0772	-0.2315	-0.3858	-0.3858	-0.2315	0.0772	0.5401
3	-0.4308	0.3077	0.4308	0.1846	-0.1846	-0.4308	-0.3077	0.4308
4	0.2820	-0.5238	-0.1209	0.3626	0.3626	-0.1209	-0.5238	0.2820
5	-0.1498	0.4922	-0.3638	-0.3210	0.3210	0.3638	-0.4922	0.1498
6	0.0615	-0.3077	0.5539	-0.3077	-0.3077	0.5539	-0.3077	0.0615
7	-0.0171	0.1195	-0.3585	0.5974	-0.5974	0.3585	-0.1195	0.0171

Table 1: The Legendre polynomials of order 1 to 7.

The seed perturbations for tasks 1, 2, 3 and 4 were $\phi_1 = \lambda_1$, $\phi_2 = -\lambda_1$, $\phi_3 = \lambda_2$, $\phi_4 = -\lambda_2$. (We had originally planned to also include conditions with seed perturbations λ_3, and $-\lambda_3$. However, these conditions proved too difficult.) Examples of stimuli from tasks 1 and 2 are shown in Fig. 1. Fig. 2 gives examples of stimuli from tasks 3 and 4. (These example stimuli have the maximum possible histogram difference to make the texture differences characterizing the different conditions as obvious as possible.)

2.6.3 Trial-by-trial perturbations within a given task

In each of the four tasks, the subject performed 2600 trials, 200 in each of 13 interleaved staircases. This section describes these staircases. For a given task $t = 1, 2, 3, 4$, let $b_1 = \phi_t$, and let

$$b_2 = \begin{cases} \lambda_2 & \text{if } t \text{ is 1 or 2} \\ \lambda_1 & \text{if } t \text{ is 3 or 4,} \end{cases} \tag{6}$$

and for $k = 3, 4, \cdots, 7$, let $b_k = \lambda_k$. Then we construct the perturbations

$$\eta_k^+ = \frac{b_1 + \frac{1}{2}b_k}{\|b_1 + \frac{1}{2}b_k\|} \qquad \text{and} \qquad \eta_k^- = \frac{b_1 - \frac{1}{2}b_k}{\|b_1 - \frac{1}{2}b_k\|} \tag{7}$$

for $k = 2, 3, \cdots, 7$. Note that all of the perturbations b_1, and η_k^+ and η_k^- for $k = 2, 3, \cdots, 8$ are normalized. Note also that the correlation between ϕ_t and each of η_k^+ and η_k^- is 0.8944.

For each of the 13 perturbations $\rho = b_1, \eta_k^+, \eta_k^-$, $k = 2, 3, \cdots, 7$, a staircase was used to test performance using perturbations $A\rho$ for various amplitudes A. Specifically, the staircase for a given perturbation ρ could visit the 30 histogram amplitudes $A = \frac{A_{max}}{30}, \frac{2A_{max}}{30}, \cdots, A_{max}$, for A_{max} the scalar for which the maximum absolute value of $A_{max}\rho$ is equal to $\frac{1}{8}$. Each staircase started at amplitude $A = \frac{A_{max}}{2}$ and ran for 200 trials. In each staircase, A was decremented whenever the previous two trials both yielded correct responses; otherwise A was incremented. These 13 staircases were randomly interleaved to collect the 2600 trials of data in each of tasks $t = 1, 2, 3, 4$.

3 Model Description and Results

3.1 The descriptive model

We first use a simple descriptive model to reveal the structure in our data under very weak assumptions. This model (captured by Eqs. 4 and 5) allows all of the different tools $T_{s,t}$ achieved by different subjects s in different tasks t to be freely determined under the single constraint that subject s uses the same Weibull steepness parameter B_s in all four tasks. (The fits do not differ perceptibly if we allow the Weibull parameters to vary freely across tasks for a given subject.) In addition to the parameter B_s, the other parameters for subject s are the values $T_{s,t}(\omega)$ for all $\omega \in \Omega$ and $t = 1, 2, 3, 4$. However, any given tool $T_{s,t}$ is determined only up to an unmeasurable additive constant (which is equal to the mean activation produced in the tool by Ω-scramble with histogram U). Thus, for a given subject this model has $4 \times 7 + 1 = 29$ free parameters; hence, including all five subjects, the model has $5 \times 29 = 145$ free parameters. The Bayesian procedure we use to fit the model is described in the Appendix.

3.2 Descriptive model results

The results of the descriptive model for all subjects s in all tasks t are plotted by the white lines in Fig. 3. Notice that subjects do indeed achieve very different tools in the four different tasks. This demonstrates that subjects can flexibly deploy a range of different strategies in sensing differences between Ω-scrambles.

We assume that in each different task t, subject s has synthesized the tool $T_{s,t}$ out of the mechanisms available in his-or-her visual system specifically to optimize performance in task t. In order to gain insight into the nature of these underlying mechanisms, we will need to make specific assumptions about how subjects can combine the responses of different mechanisms to produce these tools. These assumptions are described in Section 3.3.

3.3 The tool-construction model

Our goal is to determine (1) the number M of mechanisms (which will turn out to be 3) and (2) the sensitivity of each mechanism to the eight colors used in our stimuli; to accomplish this, we will need to pool the data from all five subjects across all four task conditions. The following model enables this. We assume:

1. All subjects share the same set of mechanisms; however, the relative sensitivity of these mechanisms can differ between different subjects. Thus, there exist functions $F_m : \Omega \to \mathbb{R}$, $m = 1, 2, \cdots, M$, such that the sensitivity function of mechanism m in subject s is
$$F_{s,m}(\omega) = A_{s,m} F_m(\omega) \qquad \text{for all } \omega \in \Omega \tag{8}$$
for some scalar $A_{s,m}$.

2. The tool $T_{s,t}$ achieved by subject s in the task t is a weighted sum of his-or-her mech-

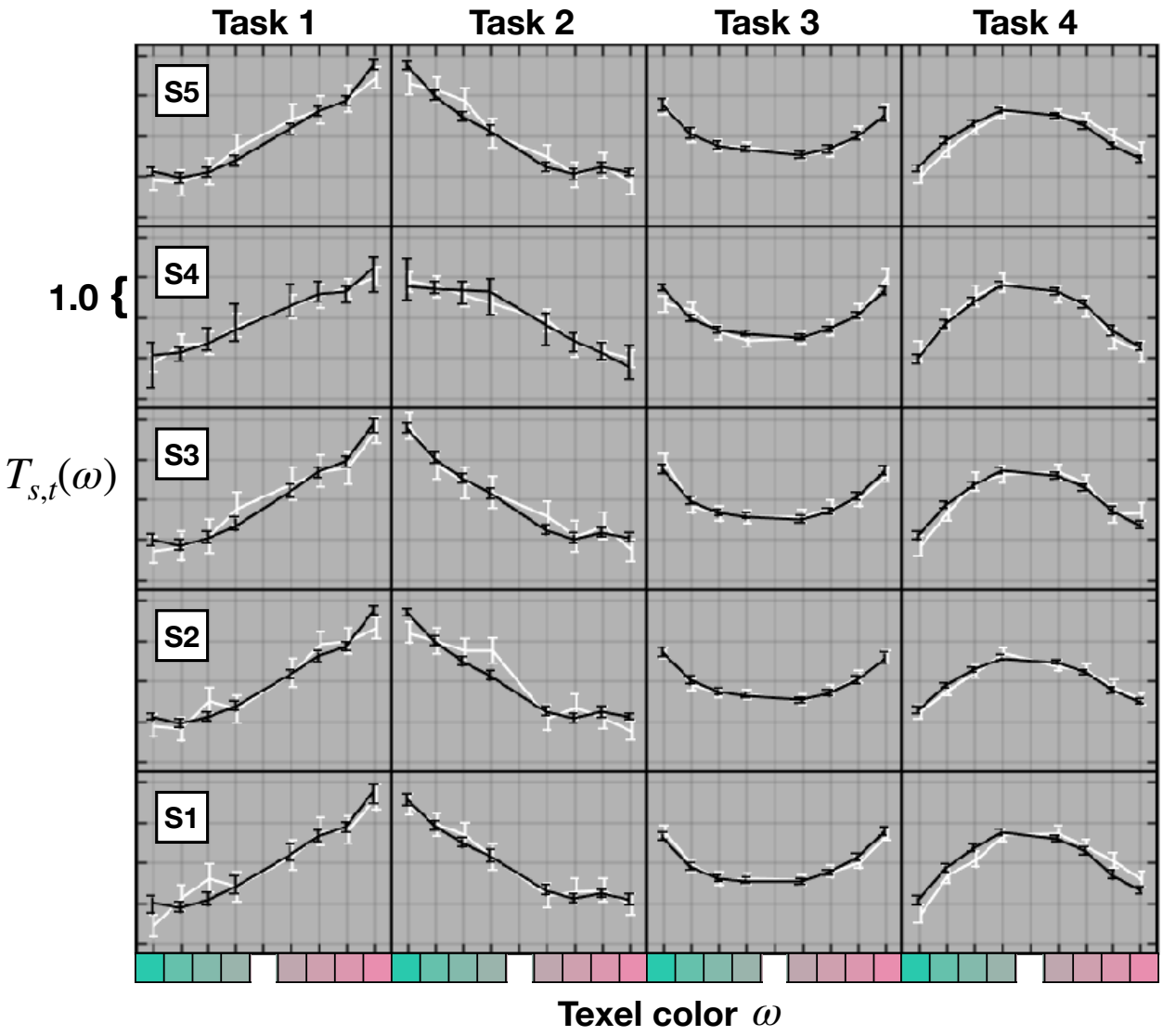

Figure 3: *Tools achieved by all subjects in all tasks.* Each column of panels corresponds to a subject s, and the tools $T_{s,t}$ achieved by that subject for tasks $t = 1, 2, 3, 4$ are plotted in successive rows. The white curve in each panel shows the form of $T_{s,t}$ yielded by the descriptive model. The black curve shows the form predicted by the tool-construction model assuming the three mechanisms shown in Fig. 4.

anism sensitivity functions:

$$T_{s,t}(\omega) = \sum_{m=1}^{M} w_{s,t,m} F_{s,m}(\omega) \qquad \text{for all } \omega \in \Omega. \tag{9}$$

3. The weights $w_{s,t,m}$ in Eq. 9 are chosen to maximize the performance of subject s in task t under the constraints that

$$\sum_{m=1}^{M} w_{s,t,m}^{\beta_s} = 1 \qquad \text{and} \qquad w_{s,t,m} \geq 0 \ \text{ for } m = 1, 2, \cdots, M, \tag{10}$$

where the parameter β_s controls the way in which subject s is able to combine the responses of his-or-her different mechanisms.

Comments on the assumptions:

1. As suggested by Eq. 4, the performance of subject s is optimized in task t by choosing the weights $w_{s,t,m}$ to maximize

$$T_{s,t} \bullet \phi_t. \tag{11}$$

2. The second constraint in Eq. 10 that the weights $w_{s,t,m}$ must be nonnegative reflects two assumptions:

   - Mechanisms produce only nonnegative activations.
   - Targets are localized by maxima in spatial distributions of neural activation.

   These two assumptions jointly imply that only those mechanisms that give larger responses to the target texture than to the background texture are likely to be useful in any given task t.

3. Concerning the combination parameter β_s: If $\beta_s \leq 1$, then the subject's best strategy is always to select the single mechanism whose sensitivity function correlates most strongly with ϕ_t. We therefore assume without loss of generality that $\beta_s \geq 1$. If $\beta_s = \infty$, then the subject's best strategy is always to combine all mechanisms whose sensitivity functions correlate positively with ϕ_t with equal weight. Intermediate values of β_s dictate specific strategies between these extremes.

3.4 Tool-construction model results

We acknowledge that the tool-construction model rests on implausibly strong assumptions. Indeed, in the current study, likelihood ratio tests reject all of the 2-, 3- and 4-mechanism tool-construction models in favor of the descriptive model with very small p-values suggesting that one or more of the assumptions underlying the model are violated. The likelihood ratio test [Wilks, 1944] relies on the fact that for $\hat{\Lambda}_{nested}$ and $\hat{\Lambda}_{fuller}$ the maximum values of the

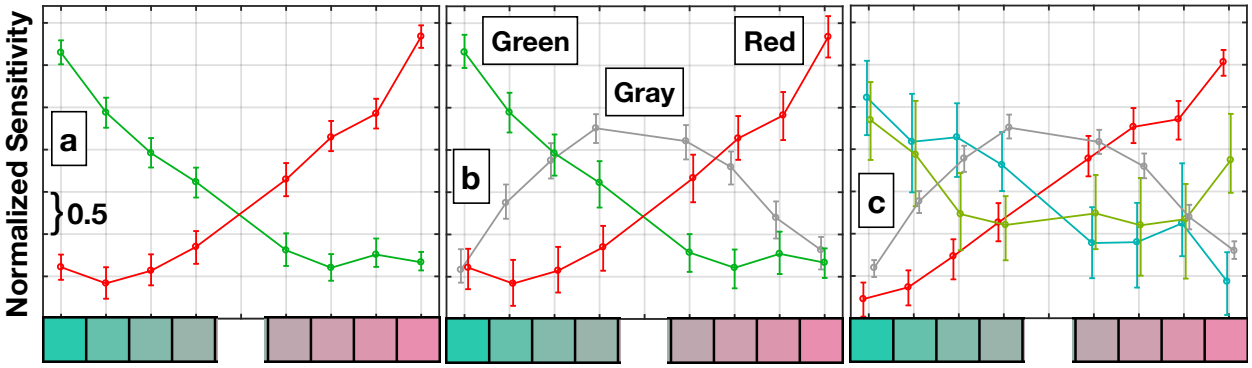

Figure 4: *Normalized mechanism functions.* Three mechanisms are implicated: a green quasi-half-axis mechanism, a red quasi-half-axis mechanism and a gray-selective mechanism. Each of the functions plotted here gives only the relative activation produced by different colors; thus, each has mean 0 and sum-of-squared-values equal to 1.

likelihood functions for the nested and fuller models,

$$X = -2\log\left(\frac{\hat{\Lambda}_{nested}}{\hat{\Lambda}_{fuller}}\right) \tag{12}$$

is asymptotically distributed as chi-square with degrees of freedom equal to the difference between the number of free parameters for the full vs. the nested model. For the comparison of the 4-mechanism model with the descriptive model, the degrees of freedom are 91; however, $X = 140$, yielding $p = 0.0007$. As expected, the comparisons of the 2- and 3-mechanism tool-construction models with the descriptive model yield even smaller p-values.

However, the tool-construction model remains attractive because it offers new leverage into the longstanding problem of estimating the sensitivity functions, F_m, characterizing individual mechanisms in human vision. This problem has proven highly resistant to straightforward empirical methods [Poirson et al., 1990, Poirson and Wandell, 1996]. Even if the detailed assumptions of the tool-construction model are partially violated (as is true in the current case), the estimates the model provides of the senstivity function functions F_m seem likely to approximate reality.

For our purposes, the model parameters β_s and B_s for subjects $s = 1, 2, \cdots, 5$ are of little interest; accordingly, we relegate them to tables 3 and 4 in the Appendix. We note in passing that β_s is between 1.5 and 5 for subjects 1, 2, 3 and 5; however, $\beta_4 \to \infty$ implying that for this subject, the optimal tool in any given task was derived by giving equal weight to all functions $F_{s,m}$ that correlated positively with the target perturbation ϕ_t. B_s was between 1.5 and 2.5 for all five subjects.

Fig. 4 shows the normalized mechanism sensitivity functions predicted under the assumption of 2 mechanisms (panel a), 3 mechanisms (panel b) and 4 mechanisms (panel c). In addition, Fig. 5 shows the tools predicted for all five subjects in all four tasks assuming only two mechanisms.

As Fig. 5 shows, the two-mechanism model accounts well for the results from tasks 1, 2 and 3; however, it fails utterly for task 4. In task 4, all of the predicted tools (shown by

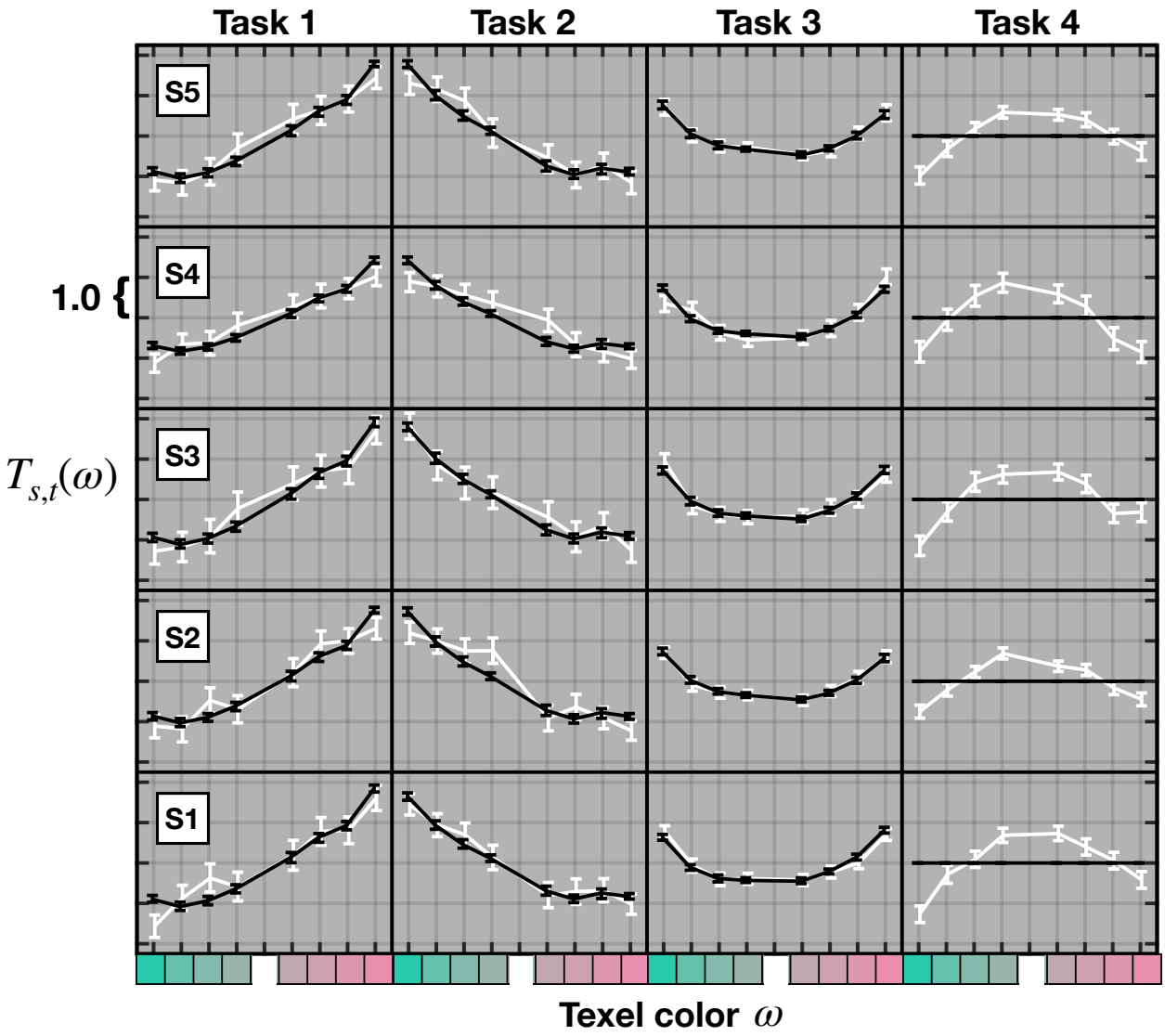

Figure 5: *Tools predicted under the 2-mechanism model.* Each column of panels corresponds to a subject s, and the tools $T_{s,t}$ achieved by that subject for tasks $t = 1, 2, 3, 4$ are plotted in successive rows. The white curve in each panel shows the form of $T_{s,t}$ yielded by the descriptive model. The black curve shows the form predicted by the tool-construction model assuming the two mechanisms shown in Fig. 4a. Note that the two-mechanism model fits the data well for tasks 1,2 and 3 but fails for task 4 (for each subject, the predicted tool is flat).

the black curves) are flat functions implying that subjects should be unable to perform this task. On the contrary, however, as the amplitudes of the white curves indicate, the actual tools achieved in task 4 enable performance that is comparable to performance in the other three tasks. This makes it clear that at least three mechanisms are needed.

The three-mechanism model does a good job of describing the data. The credible intervals on all three functions (Fig. 4b) are relatively tight indicating that model parameters are well-constrained by the data, and the predicted sensitivity functions are smooth in form as one

subject	$\alpha_{s,\text{green}}$	$\alpha_{s,\text{gray}}$	$\alpha_{s,\text{red}}$
1	2.53±0.29	1.61±0.27	2.81±0.34
2	2.62±0.31	1.23±0.21	2.69±0.33
3	2.77±0.33	1.55±0.26	2.93±0.36
4	2.18±0.26	1.73±0.29	2.15±0.28
5	2.68±0.31	1.35±0.23	2.71±0.33

Table 2: *The amplitudes of each of the green, gray and red mechanisms in all five subjects.* Note that the green and red mechanisms have roughly twice the amplitude of the gray mechanism for all subjects.

might expect for neural mechanisms. On the other hand, the predicted tools (shown by the black curves in Fig. 3 provide good matches to the tools estimated by the descriptive model (white curves, Fig. 3), accounting for 93% of the variance in the tools achieved by all five subjects in all four tasks.

On the other hand, the four-mechanism model appears to overfit the data. As expected given the fit provided by the three-mechanism model, the tools predicted by the four-mechanism model closely match the tools estimated with the descriptive model, accounting for 96% of the variance. However (panel c in Fig. 4), although two of the mechanism sensitivity functions (which correspond roughly to the mechanisms labeled "Red" and "Gray" in 4b) seem to be well-constrained by the data, the other two sensitivity functions (which roughly replace the mechanism labeled "Green" in Fig. 4) are closely intertwined, ragged in form, and bracketed by large error bars, suggesting that these sensitivity functions are not well-constrained by the data.

We conclude that the best description of the current results is given by the three-mechanism, tool-construction model. The normalized functions that give the shapes of the sensitivity functions characterizing the three mechanisms are shown in Fig. 4b. The "Green" mechanism is maximally activated by saturated green and shows decreasing activation with decreasing green saturation; this mechanism is invariant across the reds in Ω. The "Gray" mechanism shows maximal activation for gray and decreasing activation for increasingly saturated reds and greens. The "Red" mechanism is maximally activated by saturated red and shows decreasing activation with decreasing red saturation; this mechanism is largely invariant across the greens in Ω.

It should be noted, however, that the Gray mechanism tends to have lower amplitude than either the Green or Red mechanisms. This is shown in Table 2 which gives the sensitivity function amplitudes estimated for our subjects for each of the Green, Gray and Red mechanisms. The mean Green-mechanism sensitivity function amplitude is 2.56; The mean Gray-mechanism sensitivity function amplitude is 1.49; and mean Red-mechanism sensitivity function amplitude is 2.65.

Finally, we note that the curve for each of the red (green) mechanism sensitivity functions appears to give higher weight to achromatic gray than it does to greenish (reddish) colors. However, since we omitted achromatic gray from Ω, we have no actual estimate of the sensitivity of either of these mechanisms to this color.

4 Discussion

The current results suggest that three mechanisms in human vision are differentially sensitive to scrambles composed of lights drawn from the constant-S axis of DKL space. A "Red" mechanism is activated in a graded fashion by lights of increasing red saturation; a complementary "Green" mechanism is activated in a graded fashion by lights of increasing green saturation; and a third "Gray" mechanism is activated in a graded fashion by lights of decreasing saturation.

4.1 Are the Red and Green mechanisms really half-axis mechanisms?

The Red mechanism is approximately invariant in its response to lights varying in green saturation, and conversely, the Green mechanism is approximately invariant in its response to lights varying in red saturation. In this sense, the two mechanisms operate as we might expect half-axis mechanisms to work. If we take it to be a half-axis mechanism, the only peculiar feature of the Red (Green) mechanism is that its response to mean gray seems to be elevated above its response to lights with green (red) saturation > 0.

There are at least two possible explanations for this deviation from the expected form of a half-axis mechanism. First, the visual interpolation that leads us to believe the response of the Red (Green) mechanism to mean gray is elevated relative to its response to green (red) lights may be incorrect. Had we included gray light in our scrambles, we might have discovered that the Red (Green) sensitivity function dropped down farther for mean gray than the curves in Fig. 4b suggest.

We suspect, however, that the slight elevation of the response to mean gray visible in the plots in Fig. 4 does indeed reflect the behavior of the Red and Green mechanisms in response to the stimuli used in the current experiment. Consider a double-opponent neuron N [Shapley and Hawken, 2011] with a central spatial region C and a spatial surround S, and suppose

1. N is excited by
   (a) L-cone activation in C and
   (b) M-cone activation in S;
2. N is inhibited by
   (a) M-cone activation in C and
   (b) L-cone activation in S;
3. N's response is half-wave rectified; i.e., N issues a non-zero response only if the excitatory influences produced by the current input outweigh the inhibitory influences.

Consider the response of the neuron N to a Task-1 stimulus. In Task 1, the stimulus comprises a small, slightly pinkish target disk against a slightly greenish background. Suppose in particular that N receives a portion of the target disk in C and a portion of the

background in S. Note that even if all of the texture elements stimulating C were mean gray, producing equal L- vs. M-cone activation in C, the expected response of N would still be positive because the M-cone activation outweighs the L-cone activation in S. More generally, if C were stimulated with Ω-scramble into which some mean-gray texture elements were mixed, we would expect the neuron N to respond positively to those elements because they produce greater differential L-vs-M-cone activation in C than is produced on average by the greenish scramble stimulating S.

Thus, if the Red mechanism were realized by a retinotopically organized array of double-opponent neurons akin to N, then we would expect to see the slightly elevated response to mean-gray that shows up in the Red mechanism sensitivity function in Fig. 4b. If we reverse the roles of the L- and M-cones in the double-opponent neuron N, we see that a similar argument works to explain the analogous effect in the Green mechanism sensitivity function.

References

C. Chubb, I. Scofield, C-C. Chiao, and G. Sperling. A method for analyzing the dimensions of preattentive visual sensitivity. *Journal of Mathematical Psychology*, 56:427–443, 2012.

A. M. Derrington, J. Krauskopf, and P. Lennie. Chromatic mechanisms in lateral geniculate nucleus of macaque chromatic mechanisms in lateral geniculate nucleus of macaque chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357: 241–265, 1984.

R. T. Eskew. Higher order color mechanisms: A critical review. *Vision Research*, 49:2686–2704, 2009.

R. T. Eskew, J. R. Newton, and F. Giulianini. Chromatic detection and discrimination analyzed by a bayesian classifier. *Vision Research*, 41:893–909, 2001.

K. R. Gegenfurtner and D. C. Kiper. Contrast detection in luminance and chromatic noise. *Journal of the Optical Society of America A*, 9:1880–1888, 1992.

P. D. Gowdy, C. F. Stromeyer, and R. E. Kronauer. Detection of flickering edges: absence of a red-green edge detector. *Vision Research*, 39(25):4186–4191, 1999a.

P. D. Gowdy, C. F. Stromeyer, and R. E. Kronauer. Facilitation between the luminance and red-green detection mechanisms: enhancing contrast differences across edges. *Vision Research*, 39(24):4098–4112, 1999b.

N. Graham. *Visual Pattern Analyzers*. Oxford University Press, New York, 1989.

W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97–109, 1970.

J. Krauskopf and K. R. Gegenfurtner. Color discrimination and adaptation. *Vision Research*, 32:2165–2175, 1992.

J. Krauskopf and Q. Zaidi. Induced desensitization. *Vision Research*, 26(5):759–762, 1986.

J. Krauskopf, D. R. Williams, and D. W. Heeley. Cardinal directions of color space. *Vision Research*, 26:23–32, 1982.

J. Krauskopf, D. R. Williams, D. W. Heeley, and A. M. Brown. Higher order color mechanisms. *Vision Research*, 36:1235–1245, 1986.

J. Krauskopf, H. Wu, and B. Farell. Coherence, cardinal directions and higher-order mechanisms. *Vision Research*, 36:1235–1245, 1996a.

J. Krauskopf, Q. Zaidi, and M. B. Mandler. Mechanisms of simultaneous color induction. *Journal of the Optical Society of America A*, 3:1752–1757, 1996b.

K. T. Mullen and J. J. Kulikowski. Wavelength discrimination at detection threshold. *Journal of the Optical Society of America A*, 7:733–742, 1990.

A. B. Poirson and B. A. Wandell. Pattern-color separable pathways predict sensitivity to simple colored patterns. *Vision Research*, 36:512–526, 1996.

A. B. Poirson, B. A. Wandell, D. Varner, and D. H. Brainard. Surface characterizations of color thresholds. *Journal of the Optical Society of America A*, 7:783–789, 1990.

M. J. Sankeralli and K. T. Mullen. Bipolar or rectified chromatic detection mechanisms? *Visual Neuroscience*, 18(1):127–135, 2001.

R. Shapley and M. Hawken. Color in the cortex–single- and double-opponent cells. *Vision Research*, 51(7):701–717, 2011.

A. E. Silva and C. Chubb. The 3-dimensional, 4-channel model of human visual sensitivity to grayscale scrambles. *Vision Research*, 101:94 – 107, 2014.

J. D. Victor, M. M. Conte, and C. Chubb. Textures as probes of visual processing. *Annual review of vision science*, 3:275–296, 2017.

A. B. Watson and J. G. Robson. Discrimination at threshold: Labelled detectors in human vision. *Vision Research*, 21:1115–1122, 1981.

M. A. Webster and J. D. Mollon. The influence of contrast and adaptation on color appearance. *Vision Research*, 34:1993–2020, 1994.

Michael A Webster and JD Mollon. Changes in colour appearance following post-receptoral adaptation. *Nature*, 349(6306):235–238, 1991.

S. S. Wilks. *Mathematical Statistics*. Princeton University Press, 1944.

5 APPENDIX

Estimates of parameter values β_s and B_s

This section gives the values of the parameters β_s (Eq. 10) and B_s ((Eq. 5) from the three-mechanism tool-construction model.

subject	β_s (maximum likelihood)	lower 95% cred. int.	upper 95% cred. int.
1	3.43	2.64	4.77
2	2.23	1.93	2.99
3	2.75	2.21	3.77
4	310	15.71	935
5	2.09	1.79	2.62

Table 3: Estimates of β_s for all subjects s.

subject	B_s (maximum likelihood)	lower 95% cred. int.	upper 95% cred. int.
1	2.30	2.10	2.42
2	2.26	2.06	2.39
3	2.01	1.84	2.13
4	1.77	1.65	2.01
5	2.40	2.20	2.54

Table 4: Estimates of B_s for all subjects s.

Concerning the Bayesian model-fitting procedure.

Here we describe the details of the Bayesian method used to estimate the joint posterior density characterizing the parameters of each of the two models used in the paper. For each model, we use Markov chain Monte Carlo (McMC) simulation to derive a sample from the posterior density characterizing the model parameters. In order to use this procedure we to specify the likelihood function for each model.

The likelihood function for the descriptive model

The descriptive model allows all of the different tools $T_{s,t}$ achieved by different subjects s in different tasks t to be freely determined under the single constraint that subject s uses the same Weibull steepness parameter B_s in all four tasks. Thus, the descriptive model is fit separately for each subject. For a given subject s, the parameters are $T_{s,t}(\omega)$, for all $\omega \in \Omega$ and $t = 1, 2, 3, 4$. However, any given tool $T_{s,t}$ is constrained to sum to 0. Thus, for a given subject this model has $4 \times 7 + 1 = 29$ free parameters; hence, in all, the model has $5 \times 29 = 145$ free parameters.

For a given subject s and task t, let C_t be the set of all trials in task t on which the subject responded correctly, and let I_t be the set on which he-or-she responded incorrectly, and let $\rho_{t,k}$ be the perturbation used to generate the stimulus on trial k in task t. Then the likelihood function of the descriptive model for a given subject s is:

$$\Lambda_s = \prod_t \prod_{k \in C_t} \Psi_s(T_{s,t} \bullet \rho_{t,k}) \prod_{k \in I_t} (1 - \Psi_s(T_{s,t} \bullet \rho_{t,k})) \,. \tag{13}$$

where Λ_s is implicitly a function of the model parameters for subject s.

The likelihood function for the tool-construction model

The parameters of the tool-construction model

The tool-construction model is a nested version of the descriptive model in which the tool $T_{s,t}$ deployed by subject s in task t is constructed from a fixed set of mechanism sensitivity functions. $M = 3$ mechanisms are required to account for the data. Each of mechanisms $m = 1, 2, 3$ is characterized by a function $F_m : \Omega \to \mathbb{R}$. However, because we assume the sensitivity functions characterizing mechanism m have different amplitudes in different subjects, we can assume without loss of generality that the functions F_m (which determine only the relative activations produced by the different colors in Ω) are normalized: $\sum_{\omega\in\Omega} F_m^2(\omega) = 1$. In addition, our experimental paradigm only allows us to measure the the projection of F_m into the space of perturbations, each of which sums to 0. This means that we can determine F_m only up to the unknown additive constant that reflects the mean activation produced in mechanism m by Ω-scramble with histogram U. Thus, the three functions F_m add $3 \times (8 - 2) = 18$ free parameters. In addition, the amplitudes $A_{s,m}$ of all three mechanisms m in all five subjects s add 15 free parameters. For each subject s we also have a mechanism-combination parameter β_s and Weibull function steepness parameter B_s adding 10 more parameters. Thus, in all the model has $18 + 15 + 10 = 43$ free parameters.

The likelihood function for the tool-construction model is given by Eq. 13 using tools $T_{s,t}$ given by Eq. 9 constructed using mechanism sensitivity functions $F_{s,m}$ given by Eq. 8. The values $A_{s,m}$ in Eq. 8 are model parameters as are the values $F_m(\omega)$ for all $\omega \in \Omega$. Crucially, the weights $w_{s,t,m}$ used by subject s to combine different mechanisms m to produce the tool $T_{s,t}$ he-or-she uses in task t are not free parameters. The assumption that these weights are chosen to optimize performance under the constraints of Eq. 10 determines these weights from the other model parameters. Thus, the values $w_{s,t,m}$ in Eq. 9 need to be derived from the model parameters. As we show in the next section, these weights are given by Eqs. 21 and 22 below.

Deriving the weights $w_{s,t,m}$.

In this section we focus on the case of a single subject s performing a single task t. To simplify notation, we drop the subscripts s and t and refer to the weights $w_{s,t,m}$ as w_m.

Suppose the subject possesses mechanisms $m = 1, 2, \cdots, M$ with sensitivity functions F_m (given by Eq. 8) and uses combination exponent β to construct tools from these mechanisms. We assume that when the subject performs the task with seed perturbation ϕ, he-or-she chooses the combination weights w_m optimally under the constraint that

$$\sum_{m=1}^{M} w_m^\beta = C \tag{14}$$

for some arbitrary constant C (that we will set to 1). In order to compute the likelihood function for the model that rests on these assumptions, we need a formula for the optimal weights w_m given

1. the functions F_m,

2. the combination parameter β,

3. and the target perturbation ϕ.

Specifically, we need a formula for the the weights w_m maximizing $\text{Tool}^T\phi$, where

$$\text{Tool} = \sum_{m=1}^{M} w_m F_m \tag{15}$$

under the constraint of Eq. 14.

We can always start by setting to 0 all w_m for which $F_m^T\phi \leq 0$. This lets us ignore the terms m in the sum for which $F_m^T\phi \leq 0$; we thus assume without loss of generality that $F_m^T\phi > 0$ for all m in the sum. Let F be the matrix whose column vectors are the sensitivity functions F_m, $m = 1, 2\cdots, M$, and let w be the column vector with values $w_1, w_2, \cdots, w_M$. Then

$$\text{Tool} = Fw \tag{16}$$

and our aim is to maximize

$$[Fw]^T\,\phi \quad = \quad w^T V \quad \text{for} \quad V = F^T\phi. \tag{17}$$

We approach the problem by introducing a Lagrange multiplier λ as follows: Define the function

$$\Psi(w,\lambda) = w^T V + \lambda\left(\sum_{m=1}^{M} w_m^\beta - C\right), \tag{18}$$

and note that the values of w_m obtained at a local maximum of Ψ will satisfy the target conditions. The partial derivatives of Ψ with respect to the w_m's and λ are

$$\frac{\partial\Psi}{\partial w_m} = V_m + \lambda\beta w_m^{\beta-1} \qquad m = 1, \cdots, M \tag{19}$$

and

$$\frac{\partial\Psi}{\partial\lambda} = \sum_{m=1}^{M} w_m^\beta - C. \tag{20}$$

By design, setting Eq. 20 to 0 yields Eq. 14. More interestingly, setting Eq. 19 to 0 yields

$$w_m = \gamma\tilde{w}_m \quad \text{for} \quad \tilde{w}_m = V_m^{\frac{1}{\beta-1}} \tag{21}$$

where $\gamma = (\lambda\beta)^{-\frac{1}{\beta-1}}$. Equation 21 determines the relative values of the weights w_m. These are the values $\tilde{w}_m$. Variations in the amplitude γ (which can be produced by varying λ) influence only the value of C in Eq. 14.

We make the w_m's satisfy Eq. 14 with $C = 1$ by setting

$$\gamma = \left(\sum_{k=1}^{n} \tilde{w}_k^\beta\right)^{-\frac{1}{\beta}}. \tag{22}$$

5.0.1 Markov chain Monte Carlo simulation.

The estimation method uses Markov chain Monte Carlo (McMC) simulation. For simplicity, we use uniform prior distributions on all parameters. In any McMC process, one starts with some arbitrary guess at the parameter vector Q (which will ultimately be thrown away) and sets ${}_1S = Q$. (Note: (1) We use pre-subscripts to indicate parameter vector sample number in the McMC process; (2) In the current applications of this method, Q comprises guesses at σ as well as the parameters in V for whichever model is being fit.) Then one iterates the following steps some large number N_{iter} of times:

1. Pick a candidate parameter vector C in the neighborhood of the last sample ${}_{n-1}S$ added to the list. Then

2. for

$$R \quad = \quad \frac{\Lambda(C) f_{prior}(C)}{\Lambda({}_{n-1}S) f_{prior}({}_{n-1}S)} \quad = \quad \frac{\Lambda(C)}{\Lambda({}_{n-1}S)} \tag{23}$$

   (where the righthand equality follows from the fact that we use uniform priors on all parameters)

   - if $R \geq 1$, set ${}_nS = C$;
   - otherwise set

$${}_nS = \begin{cases} C & \text{with probability } R \\ {}_{n-1}S & \text{with probability } 1 - R \end{cases} . \tag{24}$$

The classical result [Hastings, 1970] is that (provided the procedure for selecting candidates C satisfies certain conditions) in the limit as $N_{iter} \to \infty$ this algorithm yields a sample from the posterior density. In practice, one typically throws away the first several thousand samples from the list which are usually not representative of the samples accumulated after the MCMC process has stabilized.

5.0.2 Priors.

The bounds of the uniform priors matter very little provided they are wide enough to include the posterior density. In the McMC process used to fit the descriptive model, for a given subject s, the prior density of each parameter other than the Weibull steepness parameter B_s was taken to be uniform between -1000 and 1000; the prior on B_s was uniform between 0 and 1000. In the McMC process used to fit the tool-construction model, the priors on all parameters other than the mechanism combination parameters were taken to be uniform between -1000 and 1000. For all five subjects s, the β_s for different subjects s

5.0.3 Adaptive candidate selection.

The sampling window used to select the candidate parameter vector C on each iteration of the McMC process dramatically influences the efficiency with which one can estimate the posterior joint density of the parameters. We use an adaptive procedure to adjust the sampling window after every 1000 iterations of the process.

On any given iteration, we generate the candidate C from the last sample ($_{n-1}S$) by first setting

$$\widetilde{C} =_{n-1} S + X \tag{25}$$

where X is a vector of jointly independent Gaussian random variables whose k^{th} entry (corresponding to the k^{th} model parameter) has standard deviation γ_k. (It is the γ_k's that are adjusted after each block of 1000 iterations.) In the case of the model of Eq. **??**, $\widetilde{C}$ includes values $\widetilde{W}(0), \widetilde{W}\left(\frac{1}{6}\right), \cdots, \widetilde{W}(1)$; we generate the candidate vector C from $\widetilde{C}$ by dividing the values $\widetilde{W}(a)$ values by their sum (to impose the constraint that the values $W(a)$ must sum to 1). In the case of the model of Eq. **??** (writing N for the number of different block-types, and τ_k the k^{th} block-type), $\widetilde{C}$ includes values $\widetilde{F}(\tau_1), \widetilde{F}(\tau_2, \cdots, \widetilde{F}(\tau_N)$; we generate the candidate vector C from $\widetilde{C}$ by dividing the values $\widetilde{F}(\tau_k)$ by the square root of the sum of their squared values (to impose the constraint of Eq. **??**).

After a given block of 1000 iterations, we compute the standard deviation ρ_k of the k^{th} model parameter across the previous 1000 iterations. If all of the ρ_k's are 0 (suggesting that no candidate C was ever accepted over the course of the previous 1000 iterations and hence that we have been drawing candidates sufficiently far from $_{n-1}S$ that their likelihoods have all been much smaller than that of $_{n-1}S$), then we decrease all σ_k's by 10%. Otherwise, we adjust the γ_k's by setting $\gamma_k = A\rho_k$ where A is a scalar that was initialized to 0.1 and is itself adjusted after each 1000 trials according to the following rule: If the median value (across the previous 1000 trials) of the likelihood ratio R (Eq. 23) is less than 0.5, A is decreased by 10%; otherwise A is increased by 10%. By adjusting A in this way, we insure that the median likelihood ratio R of Eq. 23 will be near 0.5 which guarantees that the process moves efficiently to scribble in the joint posterior density.

5.0.4 Starting values, burn-in, and number of iterations.

In fitting the model of Eq. **??**, $W(a)$ was initialized to $\frac{1}{7}$ for $a = 0, 1, \cdots, 6$; β_x and β_y were initialized to 0; and σ was initialized to 15. In fitting the model of Eq. **??**, for all block-types τ, $F(\tau)$ was initialized to $\frac{1}{\sqrt{N}}$ for N equal to the number of different block-types, β_x and β_y were initialized to 0, and σ was initialized to 15. 120,000 iterations of the McMC process were performed, the first 20,000 of these were used to allow the McMC process to "burn in," and the last 100,000 were taken as a representative sample of the posterior joint density characterizing the parameters. In each case, the last 100,000 samples were plotted and inspected by eye to insure stability of parameter estimates. Stability of the obtained estimates was confirmed to hold across several other starting locations.