Do common mechanisms of adaptation mediate color discrimination and appearance? Contrast adaptation

James M. Hillis[1,2,*] and David H. Brainard[1]

[1]*Department of Psychology, University of Pennsylvania, 3401 Walnut Street, 302C, Philadelphia, Pennsylvania 19104, USA*
[2]*Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow, G12 8QB, UK*
[*]*Corresponding author: j.mhillis@psy.gla.ac.uk*

Received March 13, 2007; accepted April 2, 2007;
posted April 10, 2007 (Doc. ID 80703); published July 11, 2007

Are effects of background contrast on color appearance and sensitivity controlled by the same mechanism of adaptation? We examined the effects of background color contrast on color appearance and on color-difference sensitivity under well-matched conditions. We linked the data using Fechner's hypothesis that the rate of apparent stimulus change is proportional to sensitivity and examined a family of parametric models of adaptation. Our results show that both appearance and discrimination are consistent with the same mechanism of adaptation. © 2007 Optical Society of America

OCIS codes: 330.5020, 330.1720, 330.5510.

1. INTRODUCTION

Both the appearance of colored stimuli and sensitivity to them depend on the context in which they are presented [1–3]. Thus, both appearance and sensitivity reveal adaptation: a context-dependent change in the visual response to a focal stimulus. Is the adaptation revealed by appearance changes the same as the adaptation revealed by sensitivity changes?

There are two broad reasons why answering this question is of interest. First, as we build models of adaptation and try to connect behavioral performance with its underlying neural mechanisms, it is useful to know the range of phenomena that are amenable to a common explanation. In previous work, we showed that a single parsimonious model of adaptation can simultaneously account for changes in appearance and sensitivity in response to changes in the mean luminance and chromaticity of a spatially uniform background [4]. These results allow data from both matching and threshold experiments to be leveraged jointly in determining the properties of the common adaptive mechanism.

A second motivation is to gain insight into the tradeoffs made by the visual system as it supports distinct functions. Vision is useful both because it allows identification, across scenes, of objects based on their surface reflectance and because it allows discrimination, within scenes, of distinct objects [5,6]. Adaptation can optimize performance for both functions across changes in viewing conditions. In the case of identification, often studied through appearance measurements, adaptation can stabilize the visual response to objects across changes in scene illumination [7–9]. In the case of discrimination, often studied through sensitivity measurements, adaptation can optimize use of the limited neural response range across changes in scene content [10–12]. When a common mechanism of adaptation mediates appearance and sensitivity, it suggests either that the two functions are compatible in their demands on adaptation or that vision has compromised performance for at least one of the functions so as to economize the expenditure of neural resources. When separate mechanisms of adaptation are revealed by the two tasks, we learn that optimizing performance on each was worth an extra expenditure. Elsewhere, we present a theoretical analysis on the extent to which a single mechanism of adaptation can simultaneously optimize performance for both identification and discrimination [13].

Here we extend our previous work [4], which studies adaptation with spatially uniform fields, to the case of adaptation to relatively unstructured luminance and chromatic contrast [14–18]. Adaptation to contrast is thought to occur at a site downstream from that which mediates adaptation to spatially uniform fields [19,20]. Thus generalizing our common-mechanism conclusion to the case of contrast adaptation requires additional data. Indeed, separate studies of contrast adaptation using appearance [18] and sensitivity [17] measurements have drawn qualitatively different conclusions about the nature of the underlying mechanisms.

Because appearance and sensitivity are measured using different psychophysical tasks, the question of whether the two are mediated by the same mechanism of adaptation needs to framed by an explicit model. That is, we need a model of how adaptation affects both appearance and sensitivity. Here we employ a well-established parametric model of early visual processing [21–25] and characterize adaptation in terms of context-dependent changes in the model's parameters. This is the same approach we used in our previous paper [4], and we outline it here.

Figure 1 illustrates our general modeling approach. We consider the case of a univariate mechanism. The x axis represents strength of a focal test stimulus (e.g., its intensity or contrast). The y axis represents the magnitude of the mechanism's response to the test stimulus. The two curves show hypothetical stimulus-response functions, one for each of two different contexts in which the test stimulus is presented. Each curve characterizes the stimulus-response function for one context, and the change in the curves characterizes the adaptation that occurs across the change of contexts.

We link the response functions to performance as follows: To predict appearance matches, we postulate that equality of appearance corresponds to equality of mechanism response. Thus test stimuli with strengths indicated by the downward arrows in the figure are predicted to match across the change in context.

To predict discrimination thresholds measured within a single context, we postulate that test difference sensitivity is proportional to the local slope of the response function [22,26]. In the figure, the slope of the gray response curve is greater at P_1 than at P_2. Thus for this context, a fixed change in stimulus strength, ΔI, produces a larger response difference at P_1 than at P_2 and by assumption is easier to detect. This linking assumption means that measurements of discrimination performance as a function of stimulus strength provide information about the shape of the underlying stimulus-response function [26–28]. Indeed, as we described in detail in our previous paper [4], we can choose a parametric form for the stimulus-response function and find the parameter values that best account for measured discrimination performance.

Here we report measurements of color discrimination for test stimuli presented against a *gray* uniform background and in a set of contexts defined by dynamic luminance and chromatic contrast. We also report asymmetric color matches set by subjects for test stimuli presented across changes from the gray background to each of the other contexts. We chose stimulus conditions where prior work indicated that performance was likely to be mediated by a single postreceptoral mechanism. We used the modeling approach illustrated by Fig. 1 to determine whether changes in appearance and sensitivity induced by contrast can be accounted for by the same mechanism of contrast adaptation. That is, for each pair of contexts studied we asked whether we could find a pair of stimulus–response functions that simultaneously accounted for discrimination performance in each context and asymmetric matches set across the contexts.

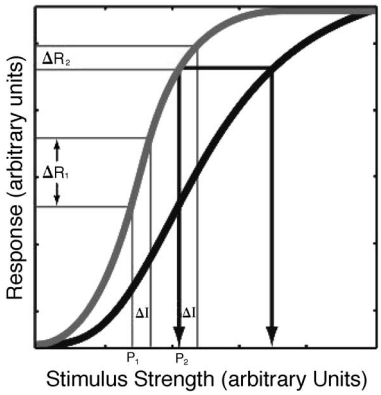

Fig. 1. Linking discrimination and appearance data. The x axis represents stimulus strength (e.g., intensity or contrast), and the y axis represents response magnitude. The two curves are response functions for the same mechanism in two states of adaptation. Because they yield the same responses in the two contexts, stimuli indicated by the downward arrows are predicted to appear the same across the change of context that induced the adaptation. Within a single context, the stimulus difference ΔI yields a larger response difference when added to P_1 as compared with P_2 (compare ΔR_1 with ΔR_2 on the y axis). If additive noise with fixed variance is limiting performance, we expect better discrimination performance at P_1 as compared with P_2.

2. METHODS

The methods were, in essence, identical to those described in detail in our previous paper [4]. Here we provide a summary, along with details about key differences. These concerned primarily the nature of the adapting contexts.

A. Observers

Two observers participated. Observer JMH was the first author, and QRS was a paid volunteer. QRS was unaware of the experimental hypotheses and had little previous experience in psychophysical experiments. Both observers had normal color vision as assessed by an Ishihara test for color blindness. Both observers participated in our previous study on the effects of uniform field changes [4].

B. Apparatus and Test Stimuli

Stimuli were presented on a calibrated RGB monitor controlled with 14-bit intensity resolution for each channel and operating at a 75 Hz refresh rate. Viewing distance was 40 cm.

Test stimuli, spatiotemporal properties: Test stimuli were blurred 1.5 deg diameter spots. They were centered 3 deg, on average, horizontally from a central fixation point. One test spot was located to the left of fixation and the other to the right. To avoid adaptation to the tests, their locations were perturbed slightly from trial to trial. On every trial, test locations were indicated by square frames composed of sparse black points. Intensity of the test spots was ramped on and off gradually.

Test stimuli, chromatic and luminance properties: Test spots were modulations relative to the background. One test color direction modulated the L and M cones together and in equal amounts, with S-cone stimulation held constant. That is, $[L_{test}\ M_{test}\ S_{test}]=I[1\ 1\ 0]$. The other test color direction modulated the S cones while the L- and M-cone stimulation was constant. That is, $[L_{test}\ M_{test}\ S_{test}]=I[0\ 0\ 1]$. We used the Smith–Pokorny estimates of cone spectral sensitivities and standard methods to convert between cone coordinate specifications and video digital-to-analog converter (DAC) settings [29].

C. Adapting Stimuli

Test spots were presented against either a spatially uniform gray field (the gray context) or within dynamic spatially pink noise. Table 1 provides descriptive color names, CIE chromaticity, cone coordinates, and RMS val-

Table 1. Cone Coordinates for Gray and Noise Backgrounds[a]

Condition Name	Color Name	x	y	Y	L	M	S	L_{RMS}	M_{RMS}	S_{RMS}
Gray	Gray	0.31	0.31	22.5	5.7	4.2	0.8	0	0	0
					2724	*1011*	*31*			
L, M noise	Pale green	0.313	0.368	31.1	8.4	6.6	0.8	3.3	2.9	0
	/red	0.306	0.225	13.3	3.1	2.0	0.8			
S noise	Blue-purple/	0.280	0.240	22.5	5.7	4.2	1.3	0	0	0.7
	brownish-yellow	0.306	0.225	22.5	5.7	4.2	0.3			

[a]L, M, S columns correspond to isomerization rate extrema present in the noise fields for each cone type ($\times 10^3$). Italicized numbers for the gray condition correspond to isomerization totals of the background for the duration and area of the test stimuli. These numbers can be used to calculate contrasts for tests and pedestals plotted as isomerization totals in Figs. 2–5. RMS values were computed for isomerization rates ($\times 10^5$) for each cone type.

ues for each context. Cone coordinates are maximum and minimum isomerization rates for light spectra present in the field. See Hillis and Brainard [4] for details on isomerization rate computations. The RMS values were computed for each cone type as $RMS_{c_i} = \sqrt{\Sigma_{x,y}(R^*_{c_i}(x,y) - \bar{R}^*_{c_i})^2}$, where c_i represents either L, M, or S cones, R^* is isomerizations per cone per second for each value present in the field, and $\bar{R}^*$ is the average isomerization rate across the field. This average was constant across all the contexts studied here and is given in the table (gray context entry). The noise pattern covered a 10.5×10.5 deg (256×256 pixel) square. In the appearance experiment, the noise was presented immediately to the left or right of fixation. Thus, the 10.5×10.5 deg noise field appeared to one side of fixation. The rest of the monitor (45 deg $\times$ 37 deg) was uniform gray (i.e., gray surrounded the noise patch). In the discrimination experiment, the dynamic noise patch was presented on both sides of fixation (i.e., the noise covered the central 21×10.5 deg). This was done to improve efficiency of data collection in the discrimination experiment. If we had left the field split, we would have amassed more data for the Gray context than for each noise context. Control experiments showed that the split field with uniform field adaptation did not affect discrimination performance in a manner that affected our analysis [4]. We assumed that the same held for current conditions.

For different adapting contexts, the noise isolated either the L and M cones (LM, for short), or the S cones. For LM-cone noise, the cone coordinates of each pixel were defined as $[L_{noise} \quad M_{noise} \quad S_{noise}] = N(x,y)[1 \quad 1 \quad 0]$, where $N(x,y)$ defined the spatially pink noise patterns that were generated by multiplying the amplitude spectrum of Gaussian noise patterns (256×256 pixels) with a two-dimensional $1/f$ amplitude spectrum. Similarly, for S-cone noise the cone coordinates of each pixel were defined as $[L_{noise} \quad M_{noise} \quad S_{noise}] = N(x,y)[0 \quad 0 \quad 1]$. To make the display dynamic, new noise patterns were shown according to draws from the probability distribution:

$$p(\tau) = \begin{cases} \frac{1}{13} & \text{for } \tau = [156{:}13{:}312] \\ 0 & \text{otherwise} \end{cases},$$

where $p(\tau)$ represents the probability of redrawing the noise at time τ since the previous redraw. The noise adapting stimulus looked like colored drifting fog.

D. Procedure: Discrimination

To assess color difference sensitivity, we used a two-interval forced (2-IFC) choice task. A fixed chromatic spot, the *pedestal*, was presented in one interval and the pedestal plus *test* in the other. The pedestal and test were defined as excursions in a particular color direction from the mean background luminance and chromaticity. That is, the pedestal was defined as $[L_{bg} \quad M_{bg} \quad S_{bg}] + I_p[L_{dir} \quad M_{dir} \quad S_{dir}]$, where $[L_{bg} \quad M_{bg} \quad S_{bg}]$ are background cone coordinates, $[L_{dir} \quad M_{dir} \quad S_{dir}]$ defines the color direction of the pedestal, and I_p is the magnitude of the pedestal excursion from the background. The pedestal+test was defined as $[L_{bg} \quad M_{bg} \quad S_{bg}] + (I_p + \Delta I)[L_{dir} \quad M_{dir} \quad S_{dir}]$, where ΔI defines the magnitude of the test. The test and pedestal were always in the same color direction. Observers were instructed to select the interval with the test.

Before experimental trials began, observers were given practice trials with auditory feedback (used throughout the experiment) where the test was clearly visible. Observers were also told that the pedestal+test would appear more intense than the pedestal alone. Observers were instructed to run practice trials until they were certain of the apparent color change that corresponded to the pedestal+test. Once they were certain, they pressed a button on a game pad controller that initiated a 90 s adaptation period. One pedestal direction and intensity were selected for each experimental session.

There was a minimum of 2 s between each trial during which the gray background or noise pattern covered the full 10.5×10.5 deg patch to the left and right of fixation. During the trial, the region within the square frame that indicated the location of the test was free of background noise but the rest of the background noise remained on the screen.

Of particular concern was to match the stimulus conditions to those used in the asymmetric matching task (described below) as closely as possible. This introduces one uncommon aspect of the procedure: There were two spots presented in each stimulus interval. In one interval the pedestal was presented alone at both the left and right target locations. In the other interval, the pedestal was presented alone at one target location and the pedestal +test was presented at the other location. Observers indicated with a key press which interval they believed contained the test, independent of which side it appeared on. A feedback tone indicated when observers selected the incorrect interval. The intensity of the test was controlled

by staircase procedures. Four randomly interleaved staircases were used in each experimental session: two for tests shown on the right and two for tests on the left. We used 2-down, 1-up and 3-down, 1-up staircase rules to ensure comprehensive sampling of psychometric functions for each pedestal. Typically, one session was run for each pedestal intensity. In cases where the data were particularly noisy, an additional session was run.

E. Procedure: Appearance

To assess the effects of context on the color appearance, we used an asymmetric matching task. Observers fixated the center of a bipartite field: one half of the field had the gray background, and the other half had dynamic pink noise. There was an initial 90 s adaptation period prior to the first time the test spots were displayed and a minimum of 2 s of adaptation between trials. Each time the test spots were displayed, the fixed test spot, called the standard, appeared against the adapting background and the adjustable spot, called the match, appeared against the gray background. As in the discrimination experiment, background noise remained on the screen during the trial and the region within the frame indicating the location of the test was free of noise. The spatial and temporal profiles of the test spots were the same as in the discrimination experiment, with the exception that the spots were flashed once per exposure rather than twice.

We gave observers full trichromatic control of the color coordinates of the adjustable spot in the CIELAB $L^*a^*b^*$ coordinate system. Approximately red/green (a^*), blue/yellow (b^*) adjustments and luminance (L^*) adjustments were made independently using the game pad controls. The standard and match were displayed after each adjustment. The standard and match could also be displayed without making an adjustment by pressing a key on the game pad. Observers could, at any time, choose any one of four step sizes for the adjustments. After completing a match, observers rated the quality of the match with a value between 0 (could not make the match) and 3 (perfect match).

Four matches were completed in a single experimental session (typically lasting 20 min). Each match was made to a different standard, but the standards used within session were always selected from the same color direction. For each direction in LMS color space, we measured match settings for eight standard tests. A total of two to four matches were made for each standard color for each of the three contexts. Stimuli were left–right counterbalanced across sessions.

3. RESULTS

A. S-Cone Tests

Figure 2 shows JMH's discrimination and matching data for S-cone tests. Decrement discrimination thresholds (–S tests) are plotted in the left panels. Increment discrimination thresholds (+S tests) are plotted in the middle column. Filled diamonds in the top row are thresholds for test/pedestal presented on the gray context. Open circles in the middle row are thresholds for tests presented in the LM-cone noise context. Filled gray squares in the bottom row are from the S-cone noise context. For comparison,

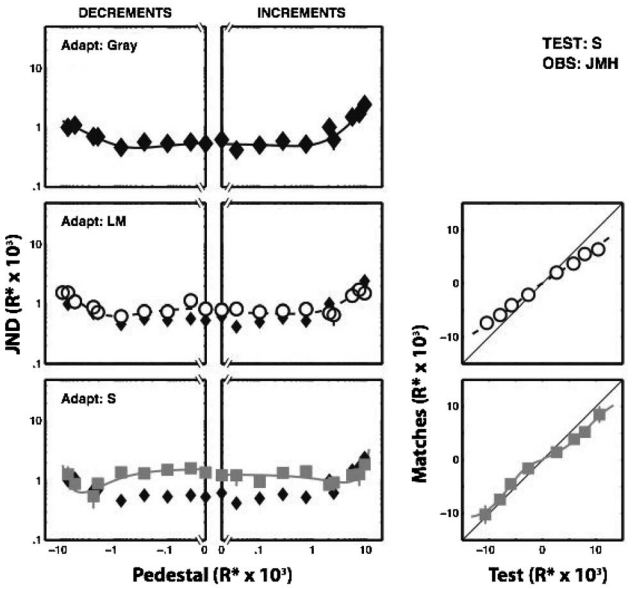

Fig. 2. JMH's discrimination threshold and appearance data for S-cone tests presented in the uniform gray (top row), LM-cone noise (middle row), and S-cone noise (bottom row). Decrement thresholds are plotted in the left column; increment thresholds are in the middle column. Thresholds for the gray condition are replotted (filled diamonds) in the middle and bottom rows for reference. Error bars are 68% confidence intervals. Asymmetric matching data are in the two rightmost panels. The x axis represents the S-cone intensity of the tests that were presented in the adapting noise fields. The y axis represents the intensity of the match set by the observer against the gray half of the display. Error bars are one standard error of the mean. The smooth curves in the plots represent the fit of the CM variant of the $\{g,q,M\}$ model of adaptation.

thresholds from the gray condition are replotted in each panel (filled diamonds). The x axis represents the expected total number of cone isomerizations (R^*) for the pedestal in isolation (i.e., isomerizations expected from the background light were subtracted from the total expected isomerizations at the test location). The y axis represents the expected total number of cone isomerizations for the test in isolation. Thresholds correspond to the test isomerizations required to correctly discriminate the test+pedestal from the pedestal alone 75% of the time, as estimated by fitting the psychometric data with a cumulative normal function. Error bars on the thresholds are 68% confidence intervals determined by bootstrapping data from the maximum-likelihood parameters [30]. Figure 3 shows the same data for QRS.

In the gray condition, S-cone discrimination thresholds do not show the clear drop in thresholds at subthreshold pedestals, called the *dipper* pattern, typically observed in pedestal+test experiments [31,22]. Rather, the S-cone discrimination thresholds against the gray background are roughly constant across the range of subthreshold pedestals used. This finding is inconsistent with the pattern of discrimination results found by Chen *et al.* [32], who used 1 cpd Gabor pedestals and tests and found the typical dipper shape with S-cone tests. The discrepancy between our result and theirs could be due to differences in the spatial bandwidth of the targets or individual subject differences. We did observe threshold increases at

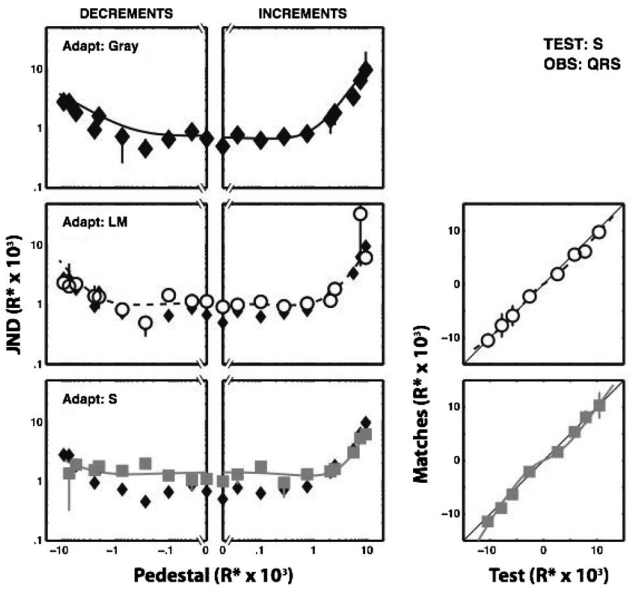

Fig. 3. QRS's threshold and appearance data for S-cone tests. The format is the same as JMH's data shown in Fig. 2.

high pedestal intensities, which is typical of pedestal +test experiments.

In the LM-cone noise condition, S-cone discrimination thresholds are slightly but systematically elevated relative to the gray background at low pedestal intensities. There was no clear effect of LM-cone noise on S tests at high pedestal intensities.

In the S-cone noise condition, discrimination thresholds are clearly elevated at low pedestal intensities. At high pedestal intensities, thresholds appear to be the same or lower in noise condition as compared with the gray context.

The rightmost panels in Fig. 2 show asymmetric matching results for S-cone tests. Fixed-test (presented against adapting noise) S-cone isomerization totals are plotted on the x axis, and the observer's S-cone match settings are plotted on the the y axis (matching stimulus presented against the uniform gray background). Error bars are the standard error of the mean. The diagonal line indicates physical identity of test-match pairs. JMH's settings appear to deviate from physical identity: Matches set against the gray background were smaller excursions from the background intensity than the tests presented in the noise contexts. Surprisingly, the effect appears to be larger against the LM-cone noise as compared with the S-cone noise. For QRS's data, on the other hand, there appears to be little or no effect of either LM-cone or S-cone adapting noise on matches. LM-cone values of JMH's and QRS's match settings for these S-cone tests (not shown) were not systematically different from zero. To test whether S-cone matches were different from the tests, we employed a simple linear regression, which tested whether the regression coefficient differed from one (physical identity). For JMH's data in both adapting conditions it did ($p<0.001$), while for QRS's data for both conditions it did not ($p>0.1$).

B. LM-Cone Tests

Figures 4 and 5 show JMH's and QRS's results for LM-cone tests presented against the uniform gray and LM-cone noise contexts. The format of these figures is the same as that of Figs. 2 and 3 except that the axes now correspond to LM-cone isomerizations. We observed no effect of S-cone noise on LM-cone detection thresholds or asymmetric matches, so we did not measure discrimination functions for this noise context. For LM-cone noise, trends for LM-cone tests were similar to those for S-cone tests. When adapted by LM-cone noise, discrimination thresholds were elevated at low intensity pedestals relative to the gray condition. At high pedestal intensities, thresholds in the noise adapted condition were the same as or lower than in the gray condition. Match intensities against the gray background were lower in magnitude than the test intensities presented in the noise background (rightmost panels) for both observers (test on whether regression coefficient differed from one, $p<0.001$).

We noted that for both S- and LM-cone tests, discrimination thresholds at high pedestal intensities were the same or lower in the noise conditions as compared with the gray context. To get a clearer picture of the pattern of results at high pedestals, Fig. 6 shows the threshold difference between gray and noise. The y axis in each panel is discrimination thresholds in the noise condition minus

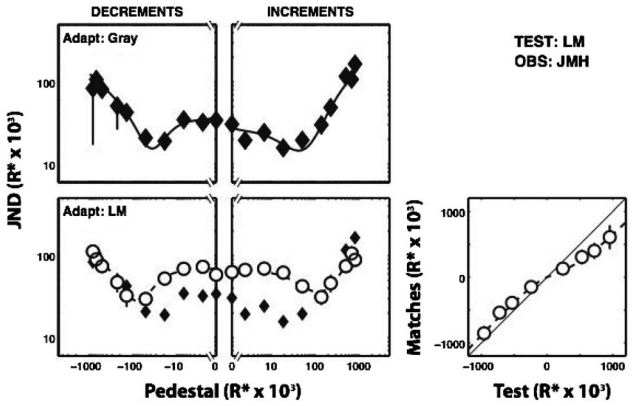

Fig. 4. JMH's threshold and appearance data for LM-cone tests presented on the uniform gray (top row) and LM-cone noise (bottom row). The format is the same as in Figs. 2 and 3 except the axes now correspond to L- and M-cone test isomerization totals.

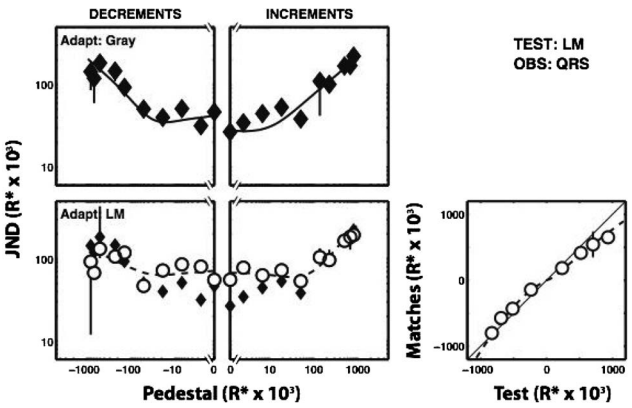

Fig. 5. QRS's threshold and appearance data for LM-cone tests presented on the uniform Gray (top row) and LM-cone noise (bottom row).

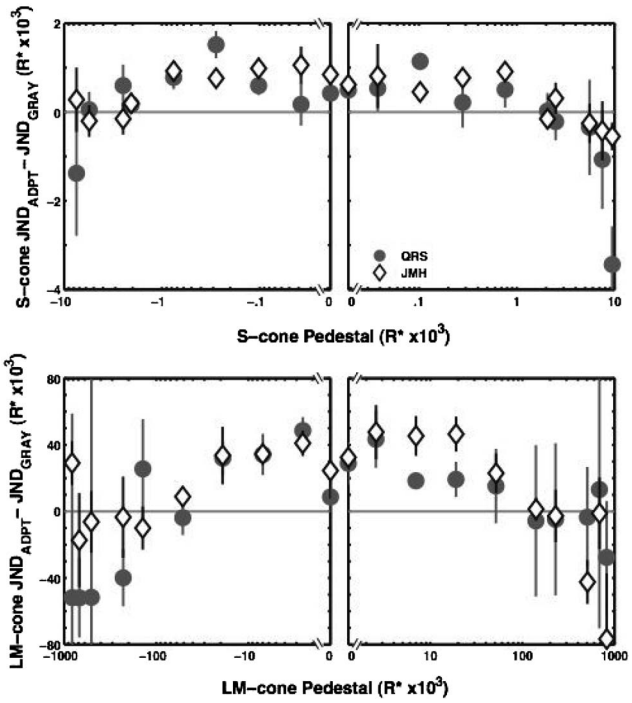

Fig. 6. Threshold difference between gray and noise contexts for S-cone tests in S-cone noise (top panels) and for LM-cone tests in LM-cone noise. The left and right columns of the panels are for decrements and increments, respectively. Diamonds are for JMH, and circles are for QRS. Error bars are 95% confidence intervals.

discrimination thresholds in the gray context. Values greater than zero therefore indicate that thresholds were higher in the noise context. The x axes correspond to pedestal intensity. The top two panels are S-cone decrement (left panel) and increment (right panel) discrimination differences between the gray and the S-cone noise contexts. The bottom two panels are LM-cone decrement and increment discrimination differences between the gray and the LM-cone noise contexts. Different symbols are for different observers. At low pedestals, it is clear that thresholds are higher in the noise contexts as compared with the gray context. This effect declines with pedestal intensity, and there is a trend for the differences to fall below zero at the highest pedestal contrasts. Although this trend is not statistically significant in our data (as determined by t-tests for the top three pedestal levels for all conditions and observers, corrected for multiple comparisons at the $p=0.05$ level), this effect has been observed in other studies of contrast adaptation [33–35]. To the extent that this effect is real, it suggests that contrast adaptation acts to expand the operating range of the adapting mechanism.

C. Does the Common Mechanism Hypothesis Hold?

Modeling

Contextual effects on color appearance and color discrimination were measured using different tasks. To determine whether the contextual effects observed in the two tasks were mediated by the same mechanism, we adopted an explicit model of that mechanism. We hypothesized that an adaptive contrast-encoding process is common to the two tasks. We chose a well-established model of image contrast encoding [22–25] based on the Naka–Rushton [21] function:

$$R = M\frac{(gI+s)^p}{(gI+s)^q+1}, \tag{1}$$

where I is the intensity of the stimulus expressed as a difference from background intensity, g controls the gain or slope, s is the "subtractive" term that determines response to mean background color, the exponents p and q determine the shape/curvature of the response nonlinearity, and M scales the response function. Examples of response functions defined by Eq. (1) are shown in Fig. 1. We, following others [36,37], used Fechner's proposal that discrimination sensitivity is proportional to appearance changes to link the two types of data. (See Fig. 1 and its discussion in the text.) A more formal description of the procedure that links the response function to predictions of the appearance and discrimination data is provided in our previous paper [38].

To model adaptation, we need to specify which parameters of Eq. (1) can change with context. For example, in what we call a $\{g,q,M\}$ model of adaptation, the parameters s and p are held fixed across all viewing contexts while the parameters g, q, and M are allowed to vary with context. In a $\{g,q,M\}$ model, we refer to s and p as *fixed parameters* and to g, q, and M as *adaptation parameters*. Because we can choose different combinations of fixed and adaptation parameters, Eq. (1) leads to a family of 31 parametric adaptation models: 1 model with five adaptation parameters, 5 models with four adaptation parameters, 10 models with three adaptation parameters, 10 models with two adaptation parameters, and 5 models with one adaptation parameter. There is also a no-adaptation model in which all five parameters remain fixed across context.

Once we choose a parametric adaptation model (e.g., the $\{g,q,M\}$ model), we can examine whether appearance and discrimination are mediated by a common mechanism of adaptation by comparing the performance of three variants of the model: (1) the common mechanism (CM) variant, (2) the independent adaptation (IA) variant, and (3) the independent channel (IC) variant. These three variants are distinguished by which parameters are set in common for appearance and discrimination and which are allowed to take on different values for the two judgments.

In the CM variant, both the fixed and adaptation parameters have the same values for appearance and discrimination. This variant describes the case where the signals mediating performance on both judgments pass through a common contrast-encoding mechanism whose properties adapt to the viewing context. For the $\{g,q,M\}$ parametric model and data from N contexts, the CM variant has $2+3N$ free parameters (two fixed parameters common to both contexts and three adaptation parameters for each context.)

In the IA variant, the fixed parameters are held common for the two tasks but the adaptation parameters are allowed to vary separately for appearance and discrimination. This variant describes the case where context-independent processing is common to all visual signals,

but adaptation occurs in a judgment-dependent fashion. For the $\{g,q,M\}$ parametric model and data from N contexts, the IA variant has $2+6N$ free parameters (two fixed parameters common to both contexts and three adaptation parameters for each context for each of two judgments.)

Finally, in the IC variant, both fixed and adaptation parameters vary with judgment. This describes the case of independent parallel processing of the signals relevant to each task. For the $\{g,q,M\}$ parametric model and data from N contexts, the IC variant has $4+6N$ free parameters (two fixed parameters for each of two judgments but common to all contexts within judgment and three adaptation parameters for each context for each of the judgments.)

Given a particular choice of parametric adaptation model, we can evaluate whether appearance and discrimination are mediated by a common mechanism of adaptation by deciding whether the CM, IA, or IC variant provides the best account of the data. These variants differ in the number of free parameters and are nested within one another (CM is nested within IA, which is nested within IC). Variants of nested models with more free parameters always produce higher likelihoods. To decide which variant accounts for the data best, we applied model selection criteria that take the number of free parameters into account. These model selection criteria are also required to choose a particular parametric adaptation model within which to compare the CM, IA, and IC variants.

Statistical methods for model selection are a topic of active investigation [39], and there is no turnkey method that may be applied when the set of models under consideration have different numbers of parameters, are nonlinear (model predictions a nonlinear function of model parameters), and are nonnested. Two model selection criteria are fairly widely used, however, to evaluate the balance between model fit and model complexity. These are Akaike's information criterion [40] (*AIC*) and the Bayesian information criterion [41,42] (*BIC*). These criteria assign an overall score to each of a set of models, based on the likelihood of the data (obtained through a maximum-likelihood fit of the model to the data) and the number of parameters in the model. The lower the value of *AIC* or *BIC*, the more preferred the model. A model's score decreases with the maximized likelihood but is penalized for the number of free model parameters. In the case of the *BIC*, the penalty varies with the size of the data set. The formulas for the two criteria are

$$AIC = -2 \cdot \ln(L) + 2 \cdot k, \tag{2}$$

$$BIC = -2 \cdot \ln(L) + \ln(n) \cdot k, \tag{3}$$

where L is the likelihood of the data from the maximum-likelihood fit, k is the number of parameters in the model, and n is the number of data points. From Eqs. (2) and (3) it is clear that the *BIC* penalizes additional parameters more aggressively than the *AIC* when the number of data points exceeds eight. We took the number of data points as the number of forced choice judgments plus the number of individual match settings.

We fit models to the data using a maximum-likelihood criterion. For each test color direction (LM increments, LM decrements, S increments, and S decrements) and observer, we used a numerical search to find maximum-likelihood model parameters. From here on, we refer to the combined matching and discrimination data from each test direction/observer combination as a *data set*, so that there were eight data sets in all. Within a data set, we fit the data from all contexts simultaneously. For example, in the case of S-cone increment tests, we had matching and discrimination data for the gray condition and two contrast adapting conditions (S-cone noise and LM-cone noise). Fitting the CM variant of a two-parameter adaptation model to this data set thus required a nine-parameter numerical search.

Calculating model likelihoods for the binomial discrimination data was straightforward (e.g., see [38]). For the appearance data, we assumed appearance matches were corrupted by Gaussian noise. To estimate the variance of the noise, we used the variance of the residuals between the individual matches and their corresponding means. To a first approximation, the residuals were normally distributed. By using the Gaussian noise assumption and variance estimated from the residuals, we could calculate the likelihood of observing the individual match settings.

We fit each variant (CM, IA, IC) of the one-, two-, and three-parameter adaptation models ($3 \times 25 = 75$ models) to each data set. We also fit the no-adaptation model. We combined information across all eight data sets to obtain overall *AIC* and *BIC* scores for each model. We did this by summing the log likelihoods obtained from each of the eight data sets to obtain an overall log likelihood and by summing the number of parameters and data points from each of the data sets to compute overall values for k and n. The total value of n was 50012, so that $\ln(n) = 10.82$ in Eq. (3).

Figure 7 shows the overall differences between *AIC* and *BIC* scores and their respective minima [i.e., $\Delta AIC = AIC - min(AIC)$ and $\Delta BIC = BIC - min(BIC)$] for the the three variants of the one-, two-, and three-parameter models of adaptation. The top-left panel shows the ΔAIC scores, and the bottom-left panel shows the ΔBIC scores. Within each panel, the blocks of bars from left to right correspond to results for the CM, IA, and IC variants. Within each block of bars, the black bars, gray bars, and white bars represent one-, two-, and three-parameter models, respectively. The adaptation models each of these bars corresponds to are listed in Table 2.

For our purpose, the important feature of the results shown in Fig. 7 is that the CM variants are preferred by both *BIC* and *AIC* over their IA and IC counterparts. For the adaptation models most preferred by *BIC*, the *BIC* prefers the CM variant; for the adaptation models most preferred by the *AIC*, the *AIC* prefers the CM variant. To elucidate this fact, we replotted the ΔAIC and ΔBIC scores as a function of parametric model rank in the two right panels. The top-right panel shows the ΔAIC value as a function of *AIC* model rank. The bottom-right panel shows the ΔBIC value as a function of *BIC* model rank. CM, IA, and IC scores within each parametric model are represented, respectively, by black symbols with gray out-

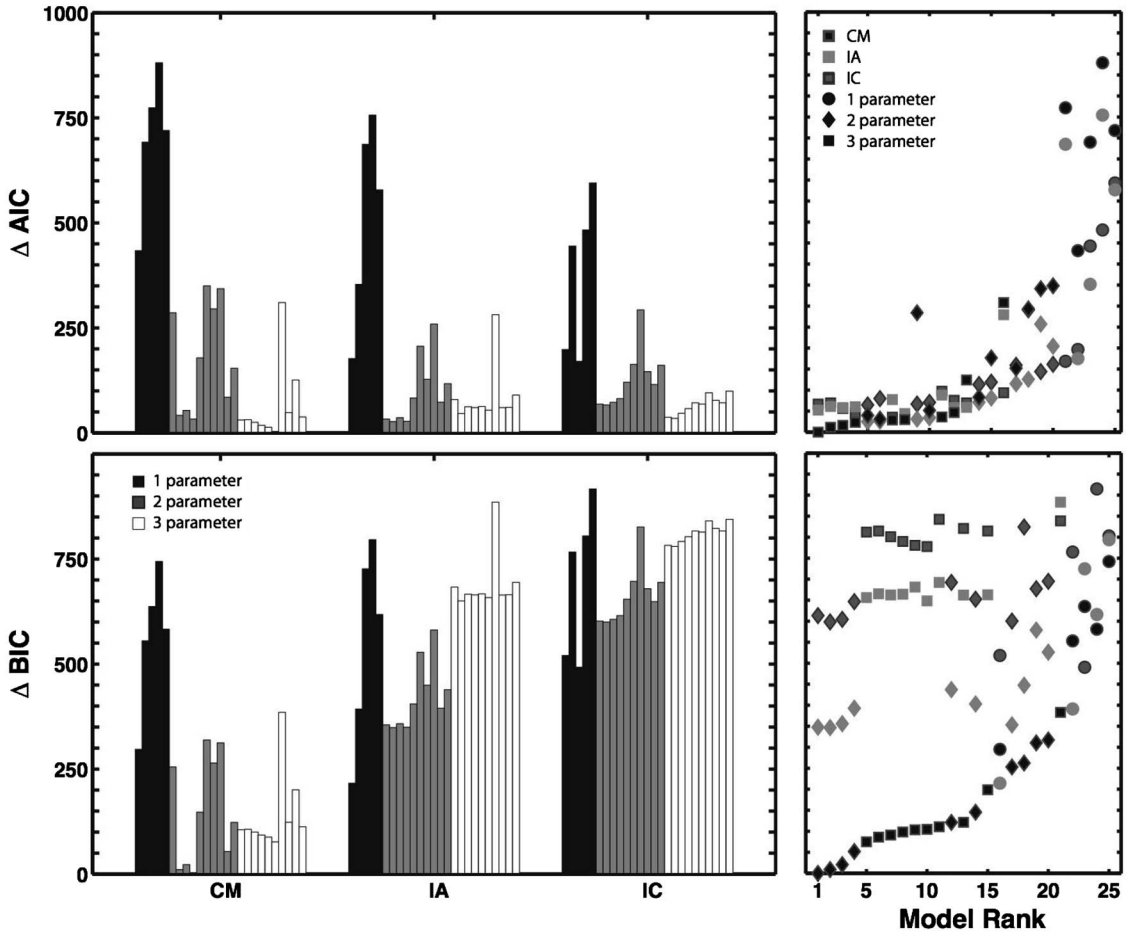

Fig. 7. (Top panel) Δ*AIC* and (bottom panel) Δ*BIC* scores for the three model variants. For the two bar plots on the left, the blocks of bars represent, from left to right, results for the CM, IA, and IC variants. Black, gray, and white bars represent one-, two-, and three-parameter adaptation models, respectively. Table 2 lists the adaptation models included in each of the blocks of bars. The two right panels show the Δ*AIC* and Δ*BIC* scores as a function of their rank. The top-right panel shows the Δ*AIC* value as a function of *AIC* rank. The bottom-right panel shows the Δ*BIC* value as a function of *BIC* rank. For each criterion (*AIC* and *BIC*), the parametric models were ranked according to the score assigned to the most preferred variant of each model, so that the CM, IA, and IC variant scores shown for any rank correspond to the same parametric model. The CM, IA, and IC results are represented, respectively, by black symbols with gray outlines, light gray symbols, and light gray symbols with dark gray outlines. Within this symbol color scheme, one-, two-, and three-parameter adaptation models are represented by circles, diamonds, and squares, respectively.

lines, light gray symbols, and light gray symbols with dark gray outlines. Within this symbol color scheme, one-, two-, and three-parameter adaptation models are represented by circles, diamonds, and squares, respectively. In these plots, it is clear that the CM variant is preferred for the models ranked best by both *AIC* and *BIC*. The smallest *AIC* difference between the best CM variant and the best IA and IC variants are 24.93 and 32.60, respectively. Differences of 10 or greater are generally considered to lend essentially no support for the model considered [43]. We therefore consider this clear evidence against the IA and IC variants.

Figure 7 also makes clear that the conclusion one draws about the CM hypothesis depends on the parametric model of adaptation used to compare the CM, IA, and IC variants. For example, if we had simply analyzed our data using the one-parameter g model of adaptation, we would have concluded from the *BIC* scores that the IA variant was the most preferred.

It is also possible to use the *AIC* and *BIC* to compare model variants for the eight individual data sets rather

Table 2. List of Adaptation Models Corresponding to the *AIC* and *BIC* Scores Shown in Fig. 7[a]

Number of Parameters		
One	Two	Three
g	g,s	g,s,p
s	g,p	g,s,q
p	g,q	g,s,M
q	g,M	g,p,q
M	s,p	g,p,M
	s,q	g,q,M
	s,M	s,p,q
	p,q	s,p,M
	p,M	s,q,M
	q,M	q,p,M

[a]In Fig. 7, black, gray, and white bars correspond to one-, two-, and three-parameter models, respectively. Each column on this table lists the models plotted for each of these three groups. For example, in Fig. 7, the five black bars in each group of bars correspond to the one-parameter adaptation models; the one-parameter models are listed in the first column of this table. The models listed from top to bottom correspond to bars from left to right.

than for the entire data set. Supplemental Tables 1 and 2 provide these comparisons for the $\{g,M\}$ and $\{g,q,M\}$ models, respectively. (Note: All supplemental materials cited in this paper are available at http://color.psych.upenn.edu/supplements/com_contrast/SupMaterial.html.) For the $\{g,M\}$ model, ranked highest by the *BIC*, the *BIC* prefers the CM variant over the IA and IC variants for each individual data set. For the $\{g,q,M\}$ model, most preferred by the *AIC*, the *AIC* prefers the CM variant over the IA and IC variants for each individual data set.

The second broad feature of Fig. 7 is that differences among the top-ranked adaptation models appear relatively small. *AIC* ranks a few of the three-parameter CM models most highly, while the *BIC* favors the two-parameter CM models but does not sharply distinguish among those it ranks in the top half. For the *AIC* the best adaptation models (ranked from highest to lowest) were $\{g,q,M\}$, $\{g,p,M\}$, $\{g,p,q\}$, and $\{g,s,M\}$ with ΔAIC values of 0, 11.41, 16.04, and 23.34, respectively. For the *BIC*, the top adaptation models were $\{g,M\}$, $\{g,p\}$, $\{g,q\}$, and $\{p,M\}$ with ΔBIC values of 0, 8.69, 20.71, and 51.83, respectively). The best one-parameter model was $\{g\}$, and this ranked worse than the better two- and three-parameter models under both *AIC* and *BIC* ($\Delta AIC=432$; $\Delta BIC=295$). For the no-adaptation model, *AIC* and *BIC* scores were higher than for any other model ($\Delta AIC=1274$; $\Delta BIC=1031$).

Because of the inconsistencies between *AIC* and *BIC* rankings and the uncertainty in statistical significance of ΔAIC and ΔBIC values [43], we thought it important to examine the quality of the fits directly against the data. The smooth curves in Figs. 2–5 show the fit of the CM variant of the three-parameter $\{g,q,M\}$ model (ranked first by *AIC*) along with the data. The model fits capture most features of the data. Supplemental Fig. S1 replots these fits and also shows the data together with the CM variant fits of the $\{g,p\}$, $\{g,q\}$, $\{g,M\}$, $\{p,M\}$, $\{g,q,M\}$, and $\{g,p,M\}$ models. Visually, neither of the three-parameter models most favored by *AIC* are obviously superior to the two-parameter models most favored by *BIC*.

Supplemental Tables 1–4 provide results of an additional analysis based on comparing models and variants using the generalized likelihood ratio test (GLRT) [44]. The GLRT can only be applied pairwise to nested models, and we limited our comparisons using the GLRT to the $\{g,M\}$, and $\{g,q,M\}$ models. The $\{g,M\}$ model is nested within the $\{g,q,M\}$ model within any variant. For the CM variant, the GLRT rejects the $\{g,M\}$ model in favor of the the $\{g,q,M\}$ model for the pooled data set (Supplemental Table 3, $p<0.05$). For the IA variant, the GLRT does not reject the $\{g,M\}$ model in favor of the $\{g,q,M\}$ model for the pooled data set at the same p value (Supplemental Table 4). GLRT comparisons of the CM and IA variants do not reject the CM variant for the $\{g,q,M\}$ but do for the $\{g,M\}$ model (Supplemental Tables 1 and 2, pooled data set, $p<0.05$). The pattern of results for individual data sets is also complex (Supplemental Tables 1–4). Overall, it is difficult to draw a clear conclusion using the GLRT analysis, since the method requires pairwise comparisons and in our case these comparisons are not mutually consistent with each other. In addition, some authors have suggested that the GLRT is overly aggressive in rejecting simpler models [43,45]. Whether the GLRT analysis supports the common mechanism hypothesis depends on the particular adaptation model considered.

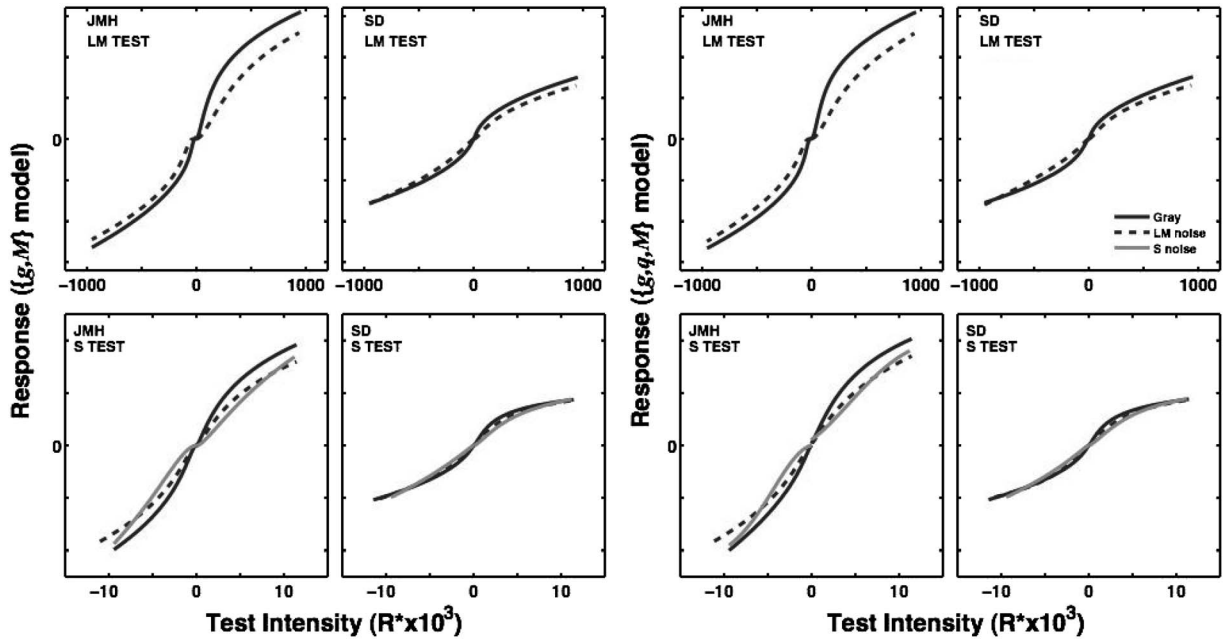

Fig. 8. S-, L-, and M-cone test response functions for the $\{g,M\}$ and $\{g,q,M\}$ models of the CM variant. The four plots grouped on the left show response functions; the $\{g,M\}$ model and the four plots grouped on the right show response functions for the $\{g,q,M\}$ model. Within each group of four panels, JMH's response functions are in the left column and QRS's are in the right column. Response functions for the S-cone tests are in the bottom row and those for the LM-cone tests are in the top row. Test intensity is on the x axes, and response magnitude is on the y axes. The solid black, dashed black, and solid gray curves are response functions inferred from the uniform gray, LM-cone noise, and S-cone noise contexts, respectively.

We can also compare the visual fits with the data of the CM, IA, and IC variants of a particular parametric model. Supplemental Fig. S2 provides this comparison visually for the models ranked first by *AIC* and *BIC* [$\{g,q,M\}$ and $\{g,M\}$, respectively]. The IA or IC variants do not capture the broad trends in the data any better than the CM variant. Differences among the variant fits are most obvious for the appearance data. There are two prominent features that show up in the IA and IC fits to appearance data that do not appear for the CM fits. First, in some cases, there is a clear discontinuity in the transition between increment and decrement matches. Second, the IA and IC fits to the appearance data do not extrapolate well in a few cases. We take the discontinuities and extrapolation anomalies to be consistent with the conclusion we drew from the *AIC* and *BIC* analyses, namely, that the IA and IC variants of the $\{g,M\}$ and $\{g,q,M\}$ models overfit the data.

It is interesting to interpret the data in terms of the underlying stimulus-response functions derived from the model parameters. These are shown for the $\{g,M\}$ and $\{g,q,M\}$ models in Fig. 8. The x axes show stimulus intensity (expressed as a difference from the background), and y axes show response magnitude. The four plots grouped on the left show response functions for the CM variant of the $\{g,M\}$ adaptation model. The four plots grouped on the right show response functions for the CM variant of the $\{g,q,M\}$ model. The response functions from the two models are essentially identical, a point we return to below. The fact that QRS generally had higher thresholds than JMH is reflected in the fact that her response functions are shallower. The solid black, dashed black, and solid light gray curves are, respectively, the response functions inferred for the gray, LM-cone noise, and S-cone noise contexts. Response functions in the noise contexts appear to be flatter at high intensities in many of the noise conditions as compared with the response functions in the gray context. Changes in the curvature of the response functions are consistent with previous findings [33–35] and the idea that contrast adaptation serves to distribute the limited response range evenly across the range of intensities present in the image [33,34,46,47]. Changes in the curvature of the response functions also lead to the sigmoidal shape of the match versus test predictions for the appearance data. This has been observed by others and is sometimes called the *crispening effect* [15,48].

In principle, the effects of the various parameters of Eq. (1) on the underlying mechanism response function are quite different from one another, and one might wonder why the data fail to distinguish sharply among parametric models with the same number of adaptation parameters. The reason becomes apparent if we examine the actual response functions corresponding to the best fit for two of the top-ranked two-parameter models: $\{g,M\}$ and $\{g,p\}$. Over the range of stimuli we employed, these two parametric models yield essentially identical response functions (see Supplemental Fig. 3). Although the two parameters $\{p\}$ and $\{M\}$ would have different effects over a larger response range, they can mimic each other's effect for the stimuli we employed.

Finally, we note that our parametric modeling is based on the assumption of fixed additive noise (see [38]). The data could also be fit under the assumption of multiplicative noise, and this might lead to different conclusions about the parametric variation of the response function with adaptation. On the basis of a few informal analyses, we concluded that our data not have the power to distinguish between models based on different noise assumptions.

4. DISCUSSION

Adaptation can can take place anywhere along the visual pathways. Considerable evidence indicates visual information flows in parallel through channels that serve specialized functions. These include ON and OFF pathways in the retina, parvocellular and magnocellular pathways through the lateral geniculate nucleus, and ventral and dorsal pathways in the visual cortex [49–53]. Given this streaming of information, it is possible that appearance judgments rely on a processing stream different from that of discrimination judgments and therefore could reveal different adaptation. Alternatively, the adaptation revealed by appearance judgments may be a result of changes in elements of the visual pathway that are common to sensitivity measures. We have referred to this latter possibility as the common mechanism (CM) hypothesis [54].

To test the hypothesis that the effects of spatiotemporal contrast on color appearance and sensitivity are mediated by a common mechanism of adaptation, we measured effects of background contrast on color appearance and sensitivity under well-matched conditions. Consistent with other studies, we found that (1) targets presented against patterned backgrounds appear lower in contrast than when presented against uniform backgrounds [14–18,55,56] and (2) color difference sensitivity changes in different ways for high- and low-contrast pedestals [33–35]. Our results and analyses lead us to the conclusion that these effects of relatively unstructured contrast on color appearance and discrimination performance can be explained by a common mechanism of contrast adaptation. Most previous work measuring appearance and detection/discrimination is in agreement with the conclusion that a common mechanism of adaptation controls both appearance and sensitivity [36,48,57–60]. We have not attempted to distinguish between effects of spatial and temporal contrast.

We do note that our formal analysis depends on the efficacy of the model selection criteria we employed. Overall, the *AIC* and *BIC* analyses both favor CM model variants. An analysis using the GLRT is less clear-cut: Whether the CM model variant is rejected by the pooled data using this approach depends on which parametric model of adaptation is adopted, and which parametric model is favored via a pairwise GLRT test depends on whether this test is performed on the CM or IA variants. The difficulty of handling our suite of models in a sensible manner via pairwise tests is one factor that led us to adopt the *AIC*/*BIC* approach. Although it remains possible that there exist other parametric models of the visual response function or other model selection approaches that would reject the CM hypothesis, our visual

examination of the quality of the various model and variant fits suggests that the size of the effect driving any such hypothetical rejection would be small.

One recent study reports a dissociation in the time course of adaptation for appearance and discrimination tasks and interprets this as indicating that separate mechanisms control adaptation for the two tasks [61]. This conclusion strikes us as premature. The stimuli in that study were brief 83 ms pulses presented at various intervals after a change in the adapting context. The dynamics of the response to the pulsed tests and of the adaptation to context change may be quite rich. Without an explicit model of both these dynamics and how the response is integrated over time to determine discrimination and appearance, it is problematic to intuit what patterns in the data are or are not consistent with mechanisms of adaptation common to the two tasks. For example, if the various parameters of the adaptation models we use here respond with different time courses to the change of context, it is easy to imagine that the time course of effects visible in the raw data will differ for discrimination and appearance.

Another apparent dissociation between adaptation for discrimination and appearance tasks is revealed by comparing the conclusions of Krauskopf *et al.* [17], who studied the effect of contrast adaptation on detection thresholds, and those of Webster and Mollon [18], who studied the effect of contrast adaptation on color appearance. The data of Krauskopf *et al.* [17] are consistent with a gain control model where adaptation occurs within separate color opponent channels. The data of Webster and Mollon [18,62] require a more complex model. There are sufficient differences between the experiments reported in the two papers, however, that the comparison does not speak strongly to the CM hypothesis as we have framed it. First, as noted by Webster and Mollon [62], the temporal properties of the test and adapting stimuli used in the two studies differed sufficiently to make comparison tenuous. Second, the data from the two studies are for quite different contrast regimes. Given the nonlinear shape of the contrast-response function needed to account for our sensitivity data, extrapolating from data at detection threshold [17] to predictions about what happens at suprathreshold contrasts [18] is not secure. Third, in our hands, fitting data from both paradigms simultaneously (CM model variant) and comparing this fit with that obtained when data from each paradigm are fit separately (IA and IC variants) is fundamental to drawing conclusions about the CM hypothesis. This approach is difficult, if not impossible, when the data are from different observers studied with different stimulus conditions in different labs. These caveats noted, it is also true that our study focused on conditions designed to tap the behavior of S-cone and LM-cone color mechanisms in isolation. A generalization of our paradigm to intermediate test and adapting color directions might reveal failures of the CM hypothesis.

Our result indicating that the CM hypothesis holds for contrast adaptation is somewhat surprising in light of the fact that, in general, a simple mechanism cannot simultaneously optimize for both discrimination sensitivity and color constancy [13,63]. This may be due to the fact that we have measured effects of relatively unstructured and abstract backgrounds. In this case, observers' phenomenology is one of making judgments about spots of light rather than about illuminated objects. It would be of great interest to know whether the CM hypothesis continues to hold for images of real scenes. Indeed, recent results from our lab [38] indicate a dissociation in adaptation for appearance and discrimination judgments when the context is a structured scene that includes cues to variations in illumination.

REFERENCES AND NOTES

1. G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (Wiley, 1982).
2. P. K. Kaiser and R. M. Boynton, *Human Color Vision*, 2nd ed. (Optical Society of America, 1996).
3. R. T. Eskew, J. S. McLellan, and F. Giulianini, "Chromatic detection and discrimination," in *Color Vision: From Molecular Genetics to Perception*, K. Gegenfurtner and L. T. Sharpe, eds. (Cambridge U. Press, 1999), pp. 345–368.
4. J. Hillis and D. H. Brainard, "Do common mechanisms of adaptation mediate color discrimination and appearance? Uniform backgrounds," J. Opt. Soc. Am. A **22**, 2090–2106 (2005).
5. J. Mollon, "Color vision," Annu. Rev. Psychol. **33**, 41–85 (1982).
6. G. H. Jacobs, *Comparative Color Vision* (Academic, 1981).
7. E. Land and J. McCann, "Lightness and retinex theory," J. Opt. Soc. Am. **61**, 1–11 (1971).
8. L. T. Maloney, "Physics-based approaches to modeling surface color perception," in *Color Vision: From Genes to Perception*, K. T. Gegenfurtner and L. T. Sharpe, eds. (Cambridge U. Press, 1999), pp. 387–416.
9. D. H. Brainard, "Color constancy," in *The Visual Neurosciences*, L. Chalupa and J. Werner, eds. (MIT, 2004), Vol. 1, pp. 948–961.
10. H. Barlow, "Possible principles underlying the transformations of sensory messages," in *Sensory Communication*, W. Rosenblith, ed. (MIT, 1961), pp. 217–234.
11. S. Laughlin, "A simple coding procedure enhances a neuron's information capacity," Z. Naturforsch. **36**, 910–912 (1981).
12. H. B. Barlow and P. Foldiak, "Adaptation and decorrelation in the cortex," in *The Computing Neuron*, C. Miall, R. Durbin, and G. Mitchison, eds. (Addison-Wesley, 1989), pp. 54–72.
13. A. Abrams, J. Hillis, and D. H. Brainard, "The relation between color discrimination and color constancy: when is optimal adaptation task dependent?" Neural Comput. (to be published).
14. C. Chubb, G. Sperling, and J. A. Solomon, "Texture interactions determine perceived contrast." Proc. Natl. Acad. Sci. U.S.A. **86**, 9631–9635 (1989).
15. R. O. Brown and D. I. A. MacLeod, "Color appearance depends on the variance of surround colors," Curr. Biol. **7**, 844–849 (1997).
16. M. D'Zmura and B. Singer, "Contast gain control," in *Color Vision: From Molecular Genetics to Perception*, K. Gegenfurtner and L. T. Sharpe, eds. (Cambridge U. Press, 1999).
17. J. Krauskopf, D. R. Williams, and D. W. Heeley, "Cardinal directions of color space." Vision Res. **22**, 1123–1131 (1982).
18. M. Webster and J. Mollon, "Changes in colour appearance following postreceptoral adaptation," Nature **349**, 235–238 (1991).
19. G. Sclar, P. Lennie, and D. DePriest, "Contrast adaptation in striate cortex of macaque," Vision Res. **29**, 747–755 (1989).
20. D. J. Heeger, "Normalization of cell responses in cat striate cortex," Visual Neurosci. **9**, 181–197 (1992).
21. K. Naka and W. Rushton, "S-potentials from colour units in

the retina of fish (Cyprinidae)," J. Physiol. (London) **185**, 536–555 (1966).
22. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458–1471 (1980).
23. E. H. Adelson, "Saturation and adaptation in the rod system," Vision Res. **22**, 1299–1312 (1982).
24. W. S. Geisler, "Mechanisms of visual sensitivity: backgrounds and early dark adaptation," Vision Res. **23**, 1423–1432 (1983).
25. J. Foley, "Human luminance pattern-vision mechanisms: masking experiments require a new model," J. Opt. Soc. Am. A **11**, 1710–1719 (1994).
26. G. Fechner, *Elements of Psychophysics*, Henry Holt Edition in Psychology (Holt, Rinehart & Winston, 1966).
27. D. Krantz, "Integration of just-noticable differences," J. Math. Psychol. **8**, 591–599 (1971).
28. R. D. Luce and W. Edwards, "The derivation of subjective scales from just noticeable differences," Physiol. Rev. **65**, 222–236 (1958).
29. D. H. Brainard, D. G. Pelli, and T. Robson, "Display characterization," in *Encylopedia of Imaging Science and Technology*, J. P. Hornak ed. (Wiley, 2002), pp. 72–188.
30. F. Wichmann and N. Hill, "The psychometric function: II. Bootstrap-based confidence intervals and sampling," Percept. Psychophys. **63**, 1314–1329 (2001).
31. J. Nachmias and R. Sansbury, "Grating contrast: discrimination may be better than detection," Vision Res. **14**, 1039–1042 (1974).
32. C. C. Chen, J. M. Foley, and D. H. Brainard, "Detection of chromoluminance patterns on chromoluminance pedestals I: threshold measurements," Vision Res. **40**, 773–788 (2000).
33. A. G. Shapiro and Q. Zaidi, "The effects of prolonged temporal modulation on the differential response of color mechanisms," Vision Res. **32**, 2065–2075 (1992).
34. Q. Zaidi, B. Spehar, and J. DeBonet, "Adaptation to textured chromatic fields," J. Opt. Soc. Am. A **15**, 23–32 (1998).
35. A. G. Shapiro, J. L. Beere, and Q. Zaidi, "Time-course of S-cone system adaptation to simple and complex fields," Vision Res. **43**, 1135–1147 (2003).
36. E. Heinemann, "The relation of apparent brightness to the threshold for differences in luminance," J. Exp. Psychol. **61**, 389–399 (1961).
37. A. J. M. Houtsma, N. I. Durlach, and L. D. Braida, "Intensity perception XI. Experimental results on the relation of intensity resolution to loudness matching," J. Acoust. Soc. Am. **68**(3), 807–813 (1980).
38. J. Hillis and D. H. Brainard, "A shadowy dissociation between discriminability and identity," J. Vision **4**, 56a (2004).
39. I. Myung, M. Forster, and M. Browne, "Guest editor's introduction: special issue on model selection," J. Math. Psychol. **44**, 1–2 (2000).
40. H. Akaike, "A new look at the statistical model identification," IEEE Trans. Autom. Control **19**, 716–723 (1974).
41. G. Schwartz, "Estimating the dimension of a model," Ann. Stat. **6**, 461–464 (1978).
42. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer Series in Statistics (Springer, 2001).
43. K. P. Burnham and D. R. Anderson, *Model Selection and Multi-Model Inference* (Springer, 2002).
44. P. Bickel and K. Doksum, *Mathematical Statistics* (Holden-Day, 1977).
45. A. E. Rafferty, "Bayesian model selection in social research," Sociol. Methodol. **25**, 111–163 (1995).
46. H. B. Barlow, D. MacLeod, and A. van Meeteeren, "Adaptation to gratings: no compensatory advantages found," Vision Res. **16**, 1043–1045 (1976).
47. M. J. Wainright, "Visual adaptation as optimal information transmission," Vision Res. **39**, 3960–3974 (1999).
48. P. Whittle, "Brightness, discriminability and the 'crispening effect'," Vision Res. **32**, 1493–1507 (1992).
49. R. W. Rodieck, *The First Steps In Seeing* (Sinauer, 1998).
50. D. M. Dacey, "Parallel pathways for spectral coding in primate retina," Annu. Rev. Neurosci. **23**, 743–775 (2000).
51. D. Hubel and T. Wiesel, "Functional architecture of macaque monkey visual cortex," Proc. R. Soc. London, Ser. B **198**, 1–59 (1977).
52. L. Ungerleider and M. Mishkin, "Two cortical visual systems," in *Analysis of Visual Behavior*, D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield, eds. (MIT, 1982), pp. 549–586.
53. D. Felleman and D. Van Essen, "Distributed hierarchial processing in the primate cerebral cortex," Cereb. Cortex **1**, 1–47 (1991).
54. In our hands, the common mechanism hypothesis includes the possibility that appearance and sensitivity are mediated by physiologically distinct processing streams that encode and adapt identically. In general, distinguishing two separate physiological mechanisms that behave identically is not possible with psychophysical data alone, and an experimental approach to this possibility would require both physiological data and a sharp theory linking that data to performance on the two types of psychophysical tasks.
55. M. Cannon and S. Fullenkamp, "Spatial interactions in apparent contrast—inhibitory effects among gratings," Vision Res. **31**, 1985–1998 (1991).
56. B. Singer and M. D'Zmura, "Color contrast induction," Vision Res. **34**, 3111–3126 (1994).
57. J. Loomis and T. Berger, "Effects of chromatic adaptation on color discrimination and color appearance," Vision Res. **19**, 891–901 (1979).
58. E. N. Pugh and J. Larimer, "Test of the identity of the site of blue/yellow hue cancellation and the site of chromatic antagonism in the π_1 pathway," Vision Res. **20**, 779–788 (1980).
59. J. Walraven, "Perceived color under conditions of chromatic adaptation: evidence for again control by π-mechanisms," Vision Res. **21**, 611–620 (1981).
60. P. Whittle, "Increments and decrements: luminance discrimination," Vision Res. **26**, 1677–1691 (1986).
61. O. Rinner and K. Gegenfurtner, "Time course of chromatic adaptation for color appearance and discrimination," Vision Res. **40**, 1813–1826 (2000).
62. M. A. Webster and J. D. Mollon, "The influence of contrast adaptation on color appearance," Vision Res. **34**, 1993–2020 (1994).
63. M. A. Webster and J. D. Mollon, "Colour constancy influenced by contrast adaptation," Nature **373**, 694–698 (1995).