Current Biology, Volume *21*

Supplemental Information

The Dynamic Range

of Human Lightness Perception

Ana Radonjić, Sarah R. Allred, Alan L. Gilchrist, and David H. Brainard,

Supplemental Inventory

Supplemental Results

Effect of Palette Range on Luminance-to-Lightness Matching Function

Two different matching palettes were available to us: a standard matte palette, consisting of 16 matte papers ranging from Munsell N 2.0/ to N 9.5/ in steps of 0.5 and an extended glossy palette consisting of 19 glossy papers ranging from Munsell N 0.5/ to N 9.5/, also with the 0.5 steps.

The extended version of the palette, which we used for the experiments reported in the paper, differs from the standard palette in both paper material (glossy versus matte) and in the reflectance range: the extended palette papers cover a reflectance range of 2 log units (101:1) while the standard palette papers cover a reflectance range of 1.42 log units (26:1).

In experiment 1 we used the extended palette to maximize the perceptual range over which we could measure perceived lightness. The darker papers on this palette are less reflective than most surfaces encountered in natural viewing [3]. The large compression we observed, however, is not due to the presence of these darker papers. If we consider the standard N 2.0/ to N 9.5/ range of the extended palette, we find that a luminance range of 1263:1 is mapped on to a reflectance range of 28:1, a similarly large compression.

In a separate control study we explored in more detail the effect of palette. To test whether palette paper material had an effect, we compared the luminance-to-lightness matching function for the high range checkerboard (experiment 1) obtained using either the standard matte palette or a version of glossy palette covering the same reflectance range as the standard palette (the darkest three chips were covered with a white cardboard occluder). These data (not shown) indicate that there was no effect of palette paper material: the two matching functions were essentially identical.

To test whether palette range (extended vs. standard) had an effect, we measured the luminance-to-lightness matching function using both palettes for three different checkerboard stimuli: 1) the high range checkerboard used in experiment 1 (~10,000:1 luminance range), 2) a second high range checkerboard with different luminance values (~6,500:1 range) and 3) a low range checkerboard (~30:1 range; different luminance values than the 30:1 checkerboard used in experiment 2). Eight observers who participated in experiment 1 made matches for the 10,000:1 checkerboard; five out of these eight observers made matches for the remaining two checkerboards.

Figure S1 plots the matching functions obtained with the standard palette (gray) and the extended palette (blue) for these three checkerboard configurations. There is only small effect of palette range on the measurements, and this effect is much smaller than those produced by our contextual manipulations (Figure 3). Note that since the palette range affects the context in which the palette papers are viewed, changing the palette range would in fact be expected to have an effect on the appearance of the palette papers. Indeed, the direction of this effect is consistent with the contextual effects of checkerboard dynamic range reported in the paper: the slope of the matching function gets steeper when either the range in the matching context (palette) is increased while keeping the stimulus range constant (Figure S1) or when the checkerboard contextual range is decreased while keeping the matching context constant (experiment 2, Figure 3A).

Luminance-to-Lightness Matching in Stimuli with Spatially Uniform Surrounds

As we discuss in the paper, our data rule out theories that predict perceived lightness from luminance ratios, as proposed by Wallach [8] or Weber contrast. Wallach's proposal was based on data obtained by matching the appearance of tests presented within different spatially uniform surrounds. We verified that if we used our methods to study the effect of changing uniform surrounds on perceived lightness, we could obtain results approximately consistent with ratio matching. Thus, our falsification of Wallach's proposal is not an artifact of our experimental methods, nor does it represent a contradiction of Wallach's experimental data.

We measured luminance-to-lightness matching function for tests embedded within two spatially uniform surrounds: high (2.62 log cd/m^2 / 416.48 cd/m^2) and low (1.3 log cd/m^2 / 19.76 cd/m^2). In both cases, the highest test was equal in luminance to the corresponding surround, the lowest test was approximately $1/30^{th}$ the luminance, and the remaining 22 test luminances were chosen to cover the range between lowest and highest test luminance in equal log steps. The size of the surround was equal to that of the checkerboards used in experiments 1 and 2, and the test was the same size as the test used in those experiments. Five out of eight observers who participated in experiment 1 completed two experimental sessions, one in which they matched test lightness using the extended palette and another in which they used the standard palette. Each session consisted of 3 blocks of trials for the high and 3 blocks of trials for the low surround luminance condition. The order of sessions (standard palette first vs. extended palette first) and blocks within a session (high luminance surround first vs. low luminance surround first) was balanced across observers. To make the data comparable to those obtained using the type of matching paradigm employed by Wallach, Figure S2 plots the average test luminance matched to each palette paper in the high luminance condition against the average luminance matched to the same palette paper in the low luminance condition. The data obtained with the two palettes are in good agreement (extended palette is plotted in blue, standard in gray). More importantly, the data are well-fit by straight lines with a slope near one (r^2 = 0.975 for extended palette, slope 1.13; r^2 = 0.967 for standard palette, slope 1.02), which is the signature of luminance ratio matching (or more generally of matching based on equating Weber contrast). There is perhaps a slight deviation from the line for the lowest luminance tests. Note that Wallach studied a more limited (~10:1) range of test luminances, so such a deviation would not have been apparent in his data.

Numerical Characterization of the Matching Functions

For each stimulus context, we summarized the luminance-to-lightness mapping by computing three values of interest: the white point (luminance matched to N 9.5/), the black point (luminance matched to N 0.5/) and the stimulus luminance range mapped onto the reflectance range of the lightness scale that was used by observers. Table S1 provides these values for each matching function obtained in experiments 1 and 2, as well as the highest and the lowest checkerboard luminance for each context.

Distribution of Out-of-Range Judgments

In the experiments, if none of the palette papers appeared as a proper lightness match for the test stimuli, the observers had an option of choosing one of three out-of-range judgments: "Darker than 0.5", "Lighter than 9.5, but still a surface" or "Glowing". Table S2 summarizes these out-of-range judgments for each test luminance in experiments 1 and 2.

Individual Observer Data

Individual observer data for experiments 1 and 2 are available in tabular form at http://color.psych.upenn.edu/supplements/dynrangelightness, as are plots of these data and corresponding response functions in the same format as Figures 2-4.

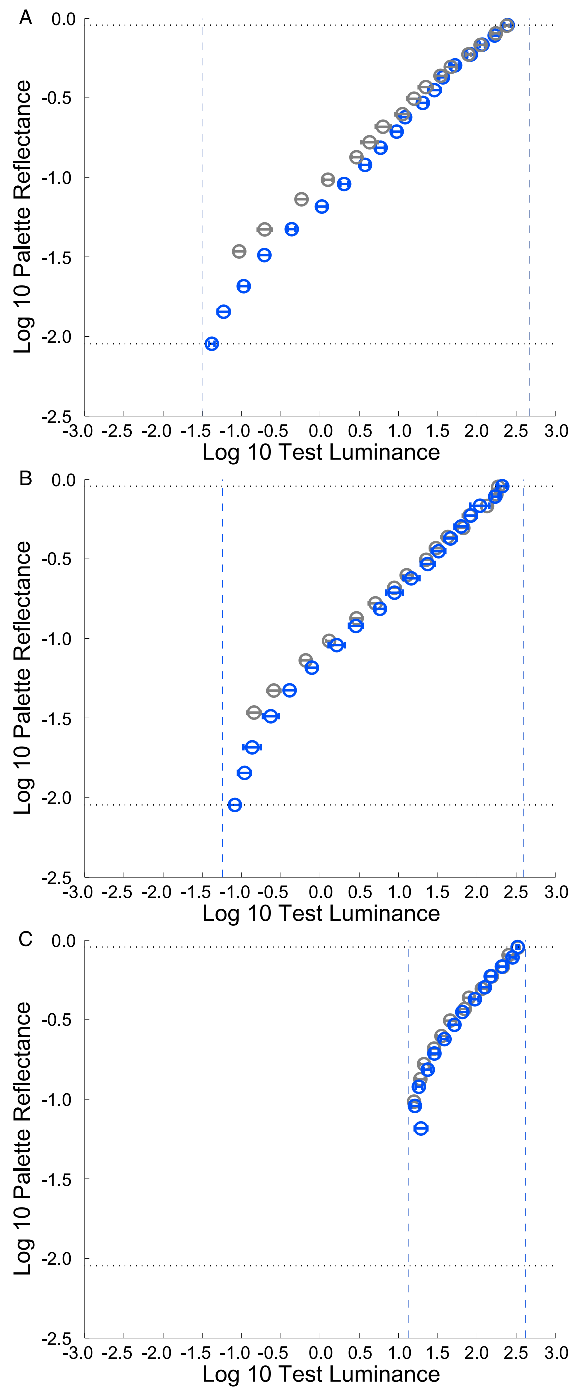

Figure S1. Palette Range Has Little Effect on the Luminance-to-Lightness Matching Functions (Related to Figure 1)

Extended glossy palette condition (in blue) and standard matte palette condition (in gray) for the (A) 10,000: 1 range, (B) 6,500:1 range and (C) 30:1 range checkerboard. In other respects, the figure follows the same conventions as Figure 2. The error bars show +/- 1 SEM computed across observers.

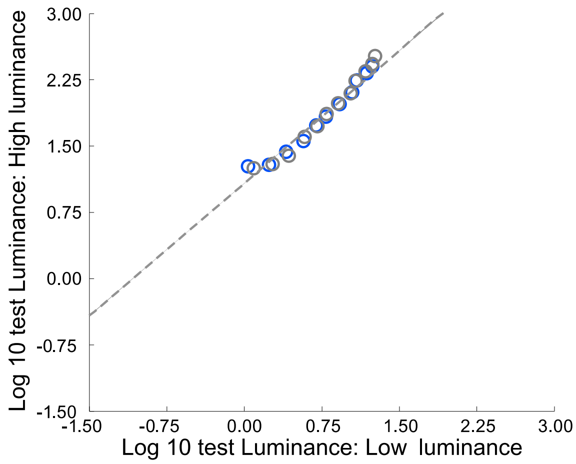

Figure S2. Ratio Matching Replicated for Changes in Uniform Surround (Related to Figure 2)

Average test luminances (log 10 cd/m^2) matched to the same palette paper for the low and high luminance surrounds are plotted against each other. Data from the extended palette in blue; from the standard palette in gray. The dashed line through the data has a slope of 1.

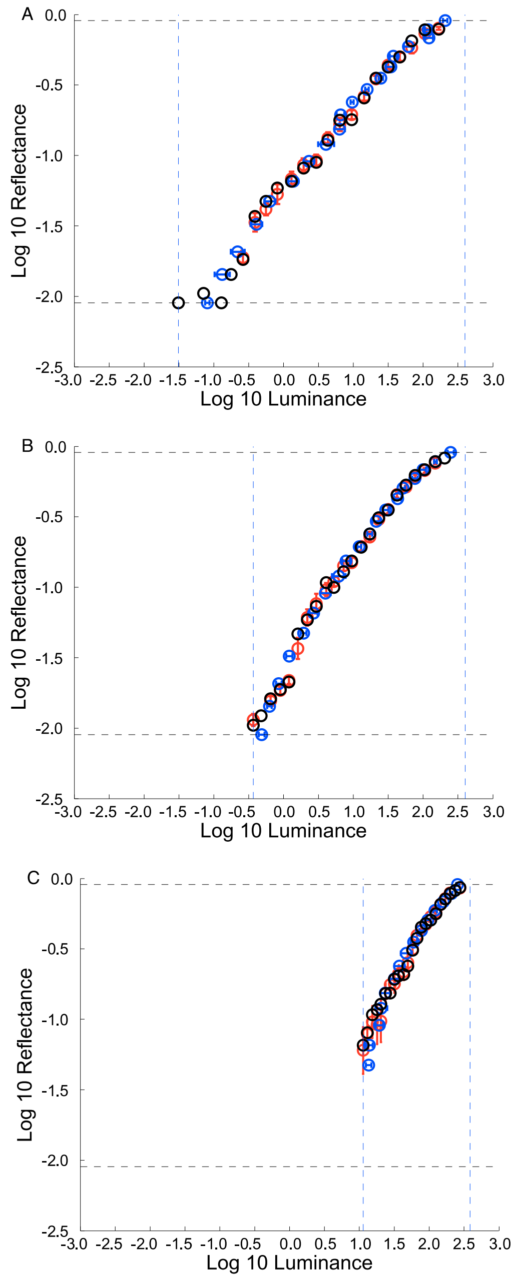

Figure S3. Three Ways of Aggregating the Matching Data Agree

Three different luminance range conditions from experiment 2: (A) 10,000:1, (B) 1,000:1, (C) 30:1. The x-axis represents the log (base 10) luminance of the test square in cd/m^2. The y-axis represents perceived lightness expressed as palette match log reflectance. The red open circles plot the average and the black open circles plot the median log reflectance match for each test log luminance. The blue open circles plot the average log test luminance matched to each palette paper log reflectance. The dotted vertical lines show the upper and lower limits of the test luminances studied, while the horizontal dotted lines show the minimum and maximum palette reflectance (Munsell N 0.5/ and Munsell N 9.5/). Where visible, error bars show +/- 1 SEM computed across observers.

Table S1. Numerical Characterization of the Matching Functions in Experiments 1 and 2

Range/ condition	White point (highest contextual luminance)	Black point (lowest contextual luminance)	Luminance range mapped to Munsell N 9.5/ - N 0.5/ (reflectance range)	Luminance range mapped to fixed common range (reflectance range)
Experiment 1: High range checkerboard				
~10,000 : 1	2.39	-1.38	3.77	
	(2.67)	(-1.50)	(2.00)	
Experiment 2: Contextual range effect				
~10,000 : 1	2.32	-1.09	3.41	2.51
	(2.60)	(-1.51)	(2.00)	(1.45)
~1,000 : 1	2.39	-0.31	2.71	2.11
	(2.60)	(-0.43)	(2.00)	(1.45)
~30 :1	2.41	N/A		1.27
	(2.59)	(1.05)		(1.45)
Experiment 2: Contextual overall luminance effect				
High	2.39	-0.31	2.71	
1,000:1	(2.60)	(-0.43)	(2.00)	
Low	1.73	-1.10	2.83	
1,000:1	(1.78)	(-1.40)	(2.00)	
High	2.41	N/A		1.28
30:1	(2.59)	(1.05)		(1.22)
Low	*1.69	N/A		1.18
30:1	(1.81)	(0.35)		(1.22)
Experiment 2: Contextual luminance distribution effect				
High mean	2.36	-0.75	3.12	
	(2.52)	(-1.01)	(2.00)	
Low mean	2.20	-0.89	3.08	
	(2.52)	(-1.00)	(2.00)	

For experiment 1 and all conditions of experiment 2, the first two columns of the table show the white point (luminance matched to N 9.5/, except as noted below) and the black point (luminance matched to N 0.5/) in log 10 cd/m^2. For comparison, these columns also provide the highest and lowest contextual luminance in parenthesis. The third column provides the log 10 luminance range mapped onto the Munsell palette (Munsell palette reflectance range shown in parenthesis) and the fourth column provides the luminance range mapped to the common reflectance range across compared conditions, for cases where the whole reflectance range of the Munsell palette was not used. The asterisk (*) indicates one case where the white point is equal to the luminance matched to N 9.0/ instead of N 9.5/, as the N 9.5 sample was never chosen as a match for any of the tests by most observers.

Table S2. Distribution of Out-of-Range Judgments in Experiments 1 and 2

Experiment 1		Experiment 2											
10,000:1 range		10,000:1 range		1,000:1 range High luminance		30:1 range High luminance		1,000:1 range Low luminance		3,300:1 range High mean		3,300:1 range Low mean	
Test	Darker than N 0.5/	Test	Darker than N 0.5/	Test	Darker than N 0.5/	Test	Darker than N 0.5/	Test	Darker than N 0.5/	Test	Darker than N 0.5/	Test	Darker than N 0.5/
0.03	1(8)	0.03	1(5)	0.48	1(5)			0.04	3(5)	0.10	1 (5)	0.10	1(5)
	1(24)		2(15)		1(15)				5(15)		2(15)		2(15)
0.05	1(8)	0.07	2(5)					0.06	3(5)	0.16	1 (5)	0.11	2(5)
	1(24)		5(15)						4(15)		1(15)		2(15)
		0.09	3(5)					0.10	1(5)	0.21	1 (5)	0.16	2(5)
			3(15)						1(15)		1(15)		2(15)
		0.13	2(5)					0.13	1(5)	0.40	1 (5)		
			2(15)						1(15)		1(15)		
		0.18	1(5)					0.19	1(5)	1.53	1 (5)		
			2(15)						1(15)		1(15)		
Test	Lighter than N 9.5/	Test	Lighter than N 9.5/	Test	Lighter than N 9.5/	Test	Lighter than N 9.5/	Test	Lighter than N 9.5/	Test	Lighter than N 9.5/	Test	Lighter than N 9.5/
193.05	3(8)	257.39	3(5)	203.34	1(5)	236.83	2(5)	60.32	1(5)	211.23	2(5)	144.24	1(5)
	3(24)		3(15)		1(15)		2(15)		1(15)		4(15)		1(15)
295.48	5(8)	396.00	1(5)	289.60	4(5)	325.15	5(5)			332.15	3(5)	210.37	4(5)
	10(24)		1(15)		6(15)		10(15)				7(15)		5(15)
464.32	6(8)			400.87	1(5)	385.64	5(5)					331.07	1(5)
	9(24)				1(15)		9(15)						3(15)
Test	Glowing	Test	Glowing	Test	Glowing	Test	Glowing	Test	Glowing	Test	Glowing	Test	Glowing
193.05	1(8)	257.39	5(5)	289.60	4(5)	325.15	2(5)			332.15	2(5)	210.37	3(5)
	1(24)		7(15)		5(15)		3(15)				5(15)		4(15)
295.48	1(8)	396.00	5(5)	400.87	5(5)	385.64	4(5)					331.08	4(5)
	3(24)		14(15)		13(15)		6(15)						12(15)
464.32	7(8)												
	14(24)												

For each test luminance (cd/m^2) listed, the number in the first row indicates the number of observers that judged a given test luminance as out of range (out of the total number of observers); the number in the second row indicates the total number of out-of-range judgments for a given test, across observers (out of the total number of judgments). Values for low overall luminance condition for 30:1 range (experiment 2) are not shown because no test appeared as out of range at any trial for any of the observers.

Table S3. Palette Specification

Munsell chip	Nominal Reflectance	Glossy Measured Luminance (cd/m^2)	Glossy Palette Measured Reflectance	Matte Measured Luminance (cd/m^2)	Matte Palette Measured Reflectance
0.5	0.006	0.66	0.009		
1.0	0.012	1.12	0.014		
1.5	0.020	1.71	0.021		
2.0	0.031	2.84	0.033	2.99	0.034
2.5	0.046	4.36	0.047	4.31	0.047
3.0	0.066	6.32	0.066	7.06	0.073
3.5	0.090	9.10	0.091	9.74	0.096
4.0	0.120	12.46	0.120	13.95	0.134
4.5	0.156	16.41	0.153	17.72	0.166
5.0	0.198	21.12	0.194	22.78	0.209
5.5	0.246	26.30	0.239	27.77	0.249
6.0	0.300	32.73	0.294	35.03	0.313
6.5	0.362	39.43	0.354	40.85	0.369
7.0	0.431	46.87	0.425	47.41	0.434
7.5	0.507	54.68	0.505	53.10	0.495
8.0	0.591	62.44	0.592	61.96	0.592
8.5	0.684	70.00	0.684	69.45	0.681
9.0	0.787	77.11	0.780	79.43	0.805
9.5	0.900	85.52	0.906	84.88	0.897

For each palette paper the table shows Munsell value, nominal reflectance of the Munsell standard [S1], measured reflectance, and measured luminance.

Table S4. Test and Contextual Square Luminances

	Contextual squares luminance (cd/m^2)								Test luminance (cd/m^2)	
Range/ Condition (Exp #)	~10,000:1 (1)	~10,000:1 (2)	High 1,000:1 (2)	High 30:1 (2)	Low 1,000:1 (2)	Low 30:1 (2)	High mean (2)	Low mean (2)	High mean (2)	Low mean (2)
1	*0.03	*0.03	0.37	11.33	*0.04	2.23	0.10	0.10	0.10	0.10
2	*0.05	*0.07	0.48	13.08	*0.06	2.53	0.13	0.12	0.16	0.11
3	0.09	0.09	0.65	15.32	0.10	3.00	0.92	0.13	0.21	0.16
4	0.11	0.13	0.90	18.04	0.13	3.57	1.35	0.18	0.27	0.23
5	0.17	0.18	1.19	20.40	0.19	4.03	1.87	0.27	0.40	0.33
6	0.25	0.26	1.61	23.73	0.25	4.63	2.78	0.35	0.55	0.46
7	0.35	0.39	2.19	27.57	0.35	5.38	3.95	0.49	0.78	0.70
8	0.56	0.56	2.94	31.95	0.46	6.36	5.71	0.73	1.08	1.00
9	0.84	0.82	4.08	36.54	0.61	7.31	8.47	0.88	1.53	1.43
10	1.25	1.30	5.25	43.05	0.87	8.32	12.19	1.06	2.15	2.00
11	1.91	1.94	7.24	49.47	1.22	9.57	17.07	1.55	3.06	2.90
12	2.95	2.93	9.50	57.67	1.67	11.28	24.94	1.77	4.30	4.21
13	4.52	4.33	12.90	66.55	2.21	12.93	35.54	2.44	6.16	5.95
14	6.76	6.41	17.23	76.85	3.05	15.03	50.91	3.60	8.85	8.64
15	10.17	9.46	23.12	89.73	4.09	17.66	72.78	5.20	12.30	12.00
16	15.27	14.31	31.63	105.44	5.56	20.14	108.56	7.26	17.20	16.84
17	23.14	21.13	42.30	124.09	7.45	23.31	137.96	10.36	24.51	24.16
18	34.91	31.49	56.98	146.83	10.14	26.96	151.72	14.52	34.55	34.12
19	52.00	46.48	76.97	170.27	13.49	30.94	172.93	20.76	48.67	48.21

20	78.58	68.51	105.19	199.21	18.50	35.21	180.75	25.61	69.04	68.43
21	119.77	106.02	148.58	236.83	25.11	42.21	186.54	36.04	98.36	97.55
22	193.05	167.09	203.34	275.70	33.43	48.17	192.94	51.35	145.06	144.24
23	295.48	257.38	289.60	325.15	45.12	55.09	228.66	214.59	211.23	210.37
24	464.32	396.00	400.87	385.64	60.32	63.91	331.87	325.21	332.15	331.08
Range	*> 10000	*> 10000	1083.43	34.04	*1500	28.66	3318.70	3252.10	3321.50	3310.80
Mean x	0.313	0.313	0.313	0.314	0.313	0.313	0.313	0.312	0.313	0.313
SD x	0.004	0.007	0.004	0.004	0.004	0.002	0.003	0.003	0.003	0.003
Mean y	0.338	0.337	0.340	0.338	0.341	0.339	0.339	0.341	0.340	0.341
SD y	0.004	0.007	0.005	0.002	0.004	0.003	0.004	0.003	0.005	0.004

For all conditions except those in which the mean luminance was varied (high/low mean) the luminance of the 24 test stimuli was sampled from the contextual squares. In the high and low mean condition, the luminance of 24 test stimuli was different than the contextual squares and is shown in the last two columns. For each experimental condition the table shows the luminance range, equal to the ratio of the highest and the lowest luminance square and the mean and standard deviation of x and y chromaticity taken across all measured squares and repeated measures. Asterisk (*) marks estimated luminance values (and corresponding range) that were below PR-650 measuring range. In some cases, there were deviations from the desired maximal and minimal luminance values due to imprecision of our calibration procedures at the low end of the luminance range.

Table S5. Checkerboard Arrangements

CS 7	CS 23	CS 4	CS 10	CS 12
CS 18	CS 6	CS 22	CS 17	CS 21
CS 1	CS 13	TEST	CS 3	CS 14
CS 19	CS 15	CS 16	CS 8	CS 20
CS 5	CS 9	CS 2	CS 24	CS 11

CS 9	CS 24	CS 23	CS 4	CS 22
CS 21	CS 10	CS 5	CS 17	CS 3
CS 7	CS 20	TEST	CS 6	CS 11
CS 18	CS 15	CS 19	CS 12	CS 14
CS 2	CS 8	CS 1	CS 16	CS 13

Arrangement of contextual squares in experiment 1 (left) and experiment 2 (right) indicated by its rank order (1-24 from darkest to brightest).

Supplemental Experimental Procedures

HDR Display

Checkerboard images were presented on custom computer-controlled high-dynamic range (HDR) display, based on a design reported by Seetzen et al. [43; see Figure 1A). The output from a DLP video projector (Panasonic #PT-D7600U) was projected onto a 19" LCD display panel (ViewSonic), through a Fresnel lens and diffuser placed directly against the backside of the panel. Because the LCD panel is a transmissive display it provides a multiplicative attenuation of the projector image, resulting in an overall dynamic range that is nominally the product of the native dynamic ranges of the projector and panel. Both display devices were driven at a pixel resolution of 1280 by 1024 and at a refresh rate of 60 Hz.

The LCD panel was enclosed in a box lined with light absorbing black cloth to minimize the stray light. The observer viewed the LCD panel monocularly from a distance of 73 cm through a circular aperture (6.1 cm in diameter) at the end of the enclosing box. Interposed between the observer and LCD panel was a black reduction screen (44 cm from the viewing aperture, 16 x 16 cm).

To display calibrated high-resolution images on the HDR display, it was necessary both to align the projector image with the LCD panel and to map desired stimulus values to appropriate RGB input settings for the two video cards. These tasks were accomplished using custom software developed in the lab, following the general methods outlined by Seetzen et al. [43]. The experimental software consisted primarily of MATLAB routines. To control the display, we also relied on routines from the Psychtoolbox [S2, S3; http://psychtoolbox.org], mgl (http://justingardner.net/doku.php/mgl/overview), and custom C routines that were called from MATLAB and that accessed the OpenGL API directly.

To control the chromaticity and luminance of the overall display system, we used a spectroradiometer (PhotoResearch, Inc., PR-650) to characterize the properties of the projector and LCD panel separately. This was done in situ, with the radiometer placed at the observer's eye position. Details are provided in a separate technical report [44].

To align the projector image with the LCD panel we used custom software developed in the lab. Specifically, an observer adjusted a 20 by 20 grid presented by the projector so that it aligned with a corresponding fixed grid displayed on the LCD panel. The alignment coordinates were used to create a warping map for the image displayed on the projector so that it is in spatial register with the image on the LCD panel. The warping was performed at the frame rate by processing on the video card. Because the Fresnel lens/diffuser/LCD panel onto which the image was projected had a significant thickness, this alignment was specific to the observer's eye position. The use of an aperture in the display enclosure ensured consistency of this position across sessions and observers.

Matching Chamber

Directly to the observer’s left was a diffusely illuminated chamber (40 cm long, 40.5 cm wide and 40.5 cm deep, constructed from plywood and painted matte gray, see Figure 1B) that contained the matching palette. The extended palette was positioned approximately 64 cm from the observer and consisted of 19 glossy papers ranging from Munsell N 0.5/ to N 9.5/ with a 0.5 step interval. Compared to the standard Munsell palette, the extended version has additional 3 chips on the low reflectance end and allows us to characterize the lightness of very dark stimuli. Because of the diffuse chamber illumination, the palette papers appeared matte. The luminance and reflectance of each palette paper is provided in Table S3. The table also provides the same measurements for the standard Munsell palette used in the control experiments examining the effect of palette range on the measurements (see Figure S1).

An LCD flat panel display was mounted on the back wall of the matching chamber. Observers used a slider to indicate either the number of their selected palette paper on the monitor, or an out-of-range judgment.

Checkerboard Stimulus

The stimulus was a 5 x 5 checkerboard (See Figure 1C for schematic representation). Each checkerboard square subtended 4.1° of visual angle and the whole checkerboard subtended 20.6°. The checkerboard squares had CIE chromaticity x = 0.309, y = 0.338 (approximately) and varied in luminance over the stimulus range in equal log steps. This chromaticity closely matched that of the palette papers in the matching chamber.

The test, which was the center square of the checkerboard, took on 24 different luminances within an experimental block. For all conditions except those in which we manipulated the contextual luminance distribution the test luminance took on the luminance of each of the remaining 24 checkerboard squares.

We measured the luminances and chromaticities of the test and contextual squares directly using our PhotoResearch PR-650 spectral radiometer. These measurements were made for stimuli presented at the center of the checkerboard surround, with the surround set as it was during the experiment. In some conditions, the test luminance was too low to be measured with the radiometer (< 0.07 cd/m^2). We estimated the luminance of these stimuli from their nominal values obtained via our calibration measurements and procedures. These are described in detail elsewhere [44].

To determine the test luminance values for each condition we first chose the highest and lowest test luminance for that condition and then set the remaining 22 tests at equal log steps between the highest and the lowest test. For the highest luminance range conditions (experiment 1; >10 000:1 range condition of experiment 2) the minimal and the maximal test were roughly equal to the minimal and maximal luminance that the HDR display could produce at the desired xy chromaticity. We were not able to directly measure the minimal luminance the display could produce in this condition due to limited sensitivity of PR-650. However, based on the calibration measurements, we estimate the luminance range to be at least 10 000:1.

To manipulate contextual range in experiment 2, we created the ~1000:1 and ~30:1 range checkerboards by setting the highest contextual luminance to that in the corresponding >10000:1 condition (experiment 2); then we obtained lowest luminance by dividing the highest luminance by 1000 and 30, respectively. To manipulate the overall contextual luminance, we reduced the highest luminance in the existing (high overall luminance) checkerboards by ~0.8 log units and divided these values by the desired range to get the lowest luminance.

For the contextual luminance distribution manipulation, we first chose the stimulus luminance range by setting the highest luminance to the maximal display luminance (2.52 log cd/m^2; 327 cd/m^2) and the lowest luminance to -1.01 log cd/m^2 (0.1 cd/m^2). The lowest luminance was sufficiently low to provide a large contextual luminance range (3.53 log units; ~ 3300:1), but high enough to be measureable by the PR-650. For both high and low mean luminance conditions the test luminance values were the same. These were chosen at equal log step between the highest and the lowest contextual luminance. The luminances of the contextual squares, however, differed across the two conditions and were derived in the following manner.

To keep the luminance range constant, the two brightest and the two darkest contextual squares remained the same in both configurations. These were equal to the two brightest and two darkest test luminances. The luminance values of the remaining 20 contextual squares were perturbed upwards and downwards from their corresponding test luminance values by arbitrary amounts, in such a manner as not to change the ordinal relations of the checkerboard luminance values.

Table S4 provides the measured and estimated test luminances for all conditions. There were some deviations from the desired maximal and minimal luminance values due to imprecision of our calibration procedures at the low end of the luminance range. These are, however, relatively small.

We used two different checkerboard arrangements of contextual squares in the two experiments. Table S5 provides the spatial arrangement of the checkerboard squares for each experiment. Both arrangements were pseudo-random with the constraint that neither the two lowest nor the two highest luminances could be immediately adjacent to the center square and that no two adjacent squares had luminances so similar that they appeared to merge together. We checked the X-junction luminance ratios in the checkerboards and verified that there were none that supported an interpretation of an illuminant change [6, 7, 9].

Observers

Eight naive observers, one male and seven female, between the ages of 20 and 38 participated in experiment 1. Five different observers, four male and one female between the ages of 18 and 31, participated in experiment 2. All observers were screened for normal color vision, acuity, and depth perception.

Psychophysical Task

On each trial, observers selected the palette paper that best matched the lightness of the center square of the displayed checkerboard. They were instructed to make their matches as if the checkerboard were made out of paper, and they indicated the chosen paper number by adjusting the slider. Observers also had the option of responding ”Darker than 0.5", ”Lighter than 9.5, but still a surface”, or ”Glowing”. Within a block of trials, the checkerboard context remained the same while the luminance of the center square varied from trial-to-trial in random order. For each experimental condition, each observer completed three blocks of trials. Blocks were spread across experimental sessions; within each experiment the order of conditions run by each observer was randomized separately.

Observers were instructed to take as long as they wished to make a response, and they could look back and forth between the HDR display and the matching booth as many times as they wished. The complete instructions as well as a description of a pre-experimental induction procedure [S4, S5] are provided at http://color.psych.upenn.edu/supplements/dynrangelightness.

Data Aggregation

For each palette paper and observer, we first averaged the log luminances of the tests that were matched to that paper by that observer. Then we averaged these values over observers. There were a few palette papers to which some observers made no match. In these cases, we plot the mean obtained from the remaining observers. Note however that we plot only the data for palette papers that at least half of the observers used to match any test in that condition (at least once, across three blocks of trials).

Our method of data aggregation is unconventional, in the sense that we average over values of the independent rather than the dependent experimental variable. This method extends the range over which the matching functions are computed relative to the more conventional method of finding the average reflectance matched to each test luminance. The difficulty with the conventional method is that it is not possible to compute a meaningful average for test luminances where the test was judged out of range on any trial. Figure S3 shows that over their common range the matching functions computed with these two methods are in good agreement with each other, as well as with the matching function obtained by computing the median luminance matched to each test. In computing the median, tests for which out-of-range judgments were made were included as long as fewer than half of the judgments were out of range.

Model Implementation

To fit the model to the data, we first assigned a fixed response level to each palette paper. This step is somewhat arbitrary, since we assume that there is a fixed function that maps response onto lightness. It is important, however, that the response levels corresponding to the palette papers span the response range. To accomplish this, we chose adaptation parameters for the matching chamber context that spread the responses to the palette papers across the full response range, when the response was computed with palette reflectance rather than luminance as input. To accomplish this, we fixed the value of the exponent n to be 3 and found parameters g and c such that the predicted response to the least reflective palette paper was 0.001 and the predicted response to the most reflective palette paper was 0.999. The remaining response levels were then computed using the resulting response function. These levels are indicated by dots on the y-axis of the plots in Figure 4.

Once the palette response levels were fixed, we could then predict the luminance-to-lightness matching function for any of our checkerboard contexts given a set of adaptation parameters for that context. More specifically, for each palette paper response R we inverted Equation 1 using the adaptation parameters for the checkerboard context and found the test luminance L corresponding to that response. To fit the model, we used numerical parameter search to find the adaptation parameters for each checkerboard context that minimized the prediction error for the corresponding luminance-to-lightness matching function.

Methodological Considerations

It is worth noting three methodological factors that may have affected our results. Importantly, each of these factors would tend to reduce the size of the range effects we measure, and therefore are not of concern with regard to the primary conclusions of our work. First, the reduction screen that surrounded the stimulus checkerboard was itself very dark and may have increased the effective luminance range of our lower range conditions. This would tend to increase the compression seen for the lowest of our stimulus range conditions. Second, the observers' state of adaptation might change as a consequence of sequentially viewing the checkerboard stimulus and the palette in the matching booth. This would tend to reduce any effects of changing the checkerboard stimulus, as the overall performance would represent a mixture of adaptation states for the checkerboard (varying across conditions) and the matching booth (fixed across conditions). Finally, light scatter in the eye could reduce the dynamic range of the retinal image relative to the dynamic range of the image. This would tend to decrease the compression for the highest of our stimulus range conditions.

Supplemental References

S1. Newhall, S.M., Nickerson, D., and Judd, D.B. (1943). Final report of the O.S.A. subcommittee on the spacing of Munsell Colors. Journal of the Optical Society of America 33, 385-412.

S2. Brainard, D.H. (1997). The Psychophysics Toolbox. Spatial Vision 10 (4), 433-436.

S3. Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision 10 (4), 437-442.

S4. Doerschner, K., Boyaci, H., and Maloney, L.T. (2007). Testing limits on matte surface color perception in three-dimensional scenes with complex light fields. Vision Res 47 (28), 3409-3423.

S5. Allred, S.R., and Brainard, D.H. (2009). Contrast, constancy, and measurements of perceived lightness under parametric manipulation of surface slant and surface reflectance. Journal of the Optical Society of America A 26 (4), 949-961.