Published paper

Fotios, S.A. and Cheal, C. (2007) *Evidence for response contraction bias in side-by-side matching tasks*. Lighting Research and Technology, 39 (2). 159 - 169. ISSN 1477-1535

http://dx.doi.org/10.1177/1365782807076490

Evidence For Response Contraction Bias In Side-By-Side Matching Tasks

SA Fotios CEng, MEI, MSLL, MILE, PhD, BEng(Hons), and **C Cheal** MSc, BSc(Hons)

University of Sheffield, School of Architecture

Fotios SA & Cheal C, Evidence For Response Contraction Bias In Side-By-Side Matching Tasks, Lighting Research & Technology, 2007; 39(2); 159-169

Abstract

A recent study of lamp spectrum effects at mesopic levels employed side-by-side matching to investigate brightness. The results revealed an unexpected effect, identified here as a response contraction bias, normally only expected when judging individual stimuli. Response contraction bias causes subjective responses to be biased toward the middle of a response range. Although the bias is small its effect on brightness matching can be significant if the test procedure does not employ appropriate counterbalancing.

1. Introduction

A recent review of street lighting for pedestrians concluded that further investigation was required of the effect of light source spectral power distribution (SPD) on brightness at mesopic conditions.[1] Experimental work has been carried out to fill this gap.[2,3] Three different techniques were used for comparing the brightness of lighting from different types of light source:

- Semantic rating, applied to lighting of a single SPD in a large room, allowing time for chromatic adaptation.
- Brightness ranking, applied to lighting of different SPD and illuminance presented simultaneously in side-by-side booths.
- Brightness matching, applied to lighting of different SPD presented simultaneously in side-by-side booths.

This paper discusses an unexpected bias found in the side-by-side brightness matching tests and the implications for the design of further experimental work.

2. Brightness Matching

Side-by-side matching is a technique that has been frequently employed to compare lighting of different qualities.[4-6] Two stimuli are presented in juxtaposed spaces, these spaces ranging from small bipartite fields,[7] to side-by-side booths,[4] to adjacent symmetrical rooms.[8] The test participant observes both stimuli simultaneously, typically using binocular vision and natural pupils. One of the two stimuli is set by the experimenter to a reference illuminance. The participant adjusts the illuminance of the second stimulus until the two appear equal, as near as is possible, according to the prescribed visual objective. Previous studies have used objectives including equal brightness, equal visual clarity and equal appearance, but recent work found that in side-by-side matching the precise visual objective given to the test participants does not significantly affect the mean illuminance ratio.[9] At the visually matched condition the illuminances of the two stimuli are recorded, the illuminance ratio is established, and the mean illuminance ratio from a sample of participants is tested for departure from unity. In the current work three different reference illuminances were used to determine whether this affected the mean illuminance ratio.

The current study used the side-by-side booths shown in Figure 1 and the light sources described in Table 1. The viewing chamber of each booth is of dimensions 575mm deep x 680mm wide x 660mm high, which presents a visual field of 38^0 wide by 37^0 high from the seated viewing distance of 1.0 metre in front of the partition. The interior surfaces were painted matt grey (Munsell N5) and each contained identical coloured objects, these being four 60mm high pyramids constructed one each from red, green, blue and yellow cardboard. In tests, participants were instructed to examine the whole environment and not concentrate on any one particular item.

The lamps were fitted behind the rear wall of the booths, and hence could not be seen directly. Light was directed into each booth using an internally reflective pipe feeding an integrating box above the viewing chamber. An iris damper was installed in this pipe to permit mechanical dimming. Spot luminances were measured at 14 points inside the booths to compare the spatial distribution of luminance. With an identical lamp (CFL) and illuminance (7.5 lux) in both booths, the mean ratio of the luminances of the 14 corresponding points was 0.997 (std dev = 0.016, maximum ratio = 1.026). Consider the luminances measured with one booth set by the mechanical dimming to 7.5 lux and the other to 5.0 lux, a dimming ratio considerably greater than that suggested by the mean null condition results for equal brightness. With an identical lamp (HPS) in both booths, the standard deviation of the ratios of the 14 corresponding luminance measurements is 0.0047 (mean = 0.663), which suggests negligible differences in luminance distribution at different levels of dimming. Thus it is considered that differences in spatial distribution between different lighting conditions have negligible effect on the brightness matching results.

The aim of the study was to compare the brightness of lighting from a high pressure sodium lamp (HPS) against lighting from lamps of poorer efficacy but higher colour rendering index, as this trade-off is now available for lighting subsidiary streets.[10] These lamps, two types of metal halide (MH1 and MH2) and a compact fluorescent (CFL) were thus individually compared against the HPS. In the following discussion, the CFL, MH1 and MH2 lamps are collectively referred to as the *white* lamps.

One booth was presented at one of three reference illuminances, and the illuminance of the second booth was adjusted by the participant until the two booths appeared, as near as possible, equally bright. At each lamp combination and illuminance condition, each participant provided four matches, counterbalancing both initial illuminance of the comparator (set by the experimenter to be clearly higher or lower than the reference) and application of the dimming control to both sources. The lamps were placed within the left-hand and right-hand booths for an equal number of trials to counterbalance any bias between the booths.

When the HPS lamp was used as the reference, the three reference illuminances were 2.0 lux, 7.5 lux and 15 lux, these being the bottom, middle and top of the S-series specification for lighting in subsidiary streets.[11] However, when the white lamps were used as the reference, the three reference illuminances used were 1.4 lux, 5.0 lux and 10.0 lux, this being an illuminance one class of the S-series lower than when the HPS provided the reference. This is the trade-off between colour rendering and illuminance permitted for lighting in subsidiary streets[10] and, assuming this renders equal brightness, the brightness matches would be made with the visual system at approximately the same adaptation level regardless of whether the HPS or white lamps were used as reference source. In the following discussion, the three reference illuminances are referred to as 2.0 lux, 7.5 lux and 15.0 lux irrespective of whether the HPS or white lamps were used as reference. Illuminances are horizontal illuminances measured at the centre of the floor of the booths using a pair of Konica-Minolta T-10 meters which were calibrated by the manufacturer prior to use.

21 colour-normal participants were used, having ages in the range 18-54 years old, three being in the range 45-54 years old, with an approximate mean of 31 years, and 14 were female. Prior to the commencement of tests the participants were dark adapted for 20 minutes. The results are shown in Table 2. These data show that on average, for equal brightness, the white lamps were set to a lower illuminance than the HPS. The data were confirmed as being drawn from a

normally distributed population using several graphical and statistical tests of normality (histogram, box-plot, skewness, kurtosis, Kolmogorov-Smirnov test and & Shapiro-Wilks test). According to the one-sample *t*-test these mean illuminance ratios are significantly lower than unity ($p<0.01$), which suggests that lamp SPD does affect the illuminance needed for equal brightness.

3. Null Condition Tests

Null-condition brightness matching tests were carried out to allow identification of any bias in the experimental apparatus and procedure. This was done using the same procedure as the main body of tests but using HPS lamps in the two booths. Eighteen colour-normal participants were used, these being in the age range of 18 years to 54 years, although only one participant was aged above 44 years old, and 13 were female.

At each of three reference illuminances (2.0 lux, 7.5 lux and 15.0 lux) each participant made eight brightness matches. This was done to counterbalance allocation of the adjusted lamp between the left-hand and right-hand booths; setting the adjusted lamp to an initial illuminance higher or lower than the reference illuminance; and which of the two identical lamps was nominated to be the test or reference lamp. 'Test' refers to the booth to which dimming was applied by the participant and 'reference' refers to the booth which was set by the experimenter to one of the three reference illuminances.

Table 3 shows the results of the null condition tests as the illuminance ratio of the left-hand booth to right-hand booth at the equally bright condition. A mean illuminance ratio of unity would confirm that there is no unintended experimental bias. The departure from unity was analysed using the one sample *t*-test after the data were confirmed as being drawn from a normally distributed population as described above. At all three illuminances the *t*-test indicates there is no significant departure from unity, suggesting negligible bias between the left-hand and right-hand booths.

4. Response Contraction Bias

Table 4 presents the null-condition results according to whether the lighting was from the booth regulated by the participant or from the fixed reference. There are small but significant departures from unity. At 2.0 lux the illuminance ratio is significantly higher than unity ($p<0.01$, one-sample *t*-test), but at 15.0 lux the illuminance ratio is significantly lower than unity ($p<0.01$). At 7.5 lux the illuminance ratio is not significantly different to unity. The departure from unity and the change of illuminance ratio with illuminance are unexpected since the two stimuli are of identical SPD.

A similar pattern can be seen in previous null condition data.[12] A brightness matching technique was used to compare lighting of identical SPD but different spatial distribution at three different reference illuminances, 300 lux, 500 lux and 700 lux. In the null condition, rooms of identical SPD and spatial distribution were matched, which is the same stimulus condition and task as the null condition of the current study. This null condition data shows that at the lower illuminance (300 lux) the mean illuminance of the adjusted lighting was higher than the reference, although not significantly so,[13] whereas at the higher reference illuminances (500 lux and 700 lux) the mean illuminances of the adjusted lighting were significantly lower than the reference ($p<0.01$).[13] The effect may also be present in other brightness matching studies[4,8] but the published results do not enable it to be identified.

The results of the main brightness matching study (Table 2) are broken down in Table 5 according to application of the dimming control to either the HPS lamp or to the white lamps. These data show the same pattern as do the null condition data in Table 4. At the lower reference illuminance (2.0 lux) the booth to which dimming is applied was set to a higher than average illuminance, whereas at the higher reference illuminances (7.5 lux and 15.0 lux) the booth to which dimming is applied was set to a lower than average illuminance. This trend is consistent for all three lamp combinations. The two-sample *t*-test (two-tailed) suggests these differences are significant ($p<0.05$) at 2.0 lux and 15.0 lux but not at 7.5 lux. Note that for the MH2/HPS lamp combination at 2.0 lux the difference in mean illuminance ratios is significant ($p<0.05$) for a one-tailed test but not for a two-tailed test.

These data suggest a response contraction bias, although this is only normally expected when judging an individual stimulus.[14] Consider judgements of the brightness of a series of stimuli, each stimulus being identical other than luminance, and these are observed individually without the presence of a reference stimulus. A response contraction bias causes the brightness of low luminance stimuli to be overestimated, as shown in Figure 2, and the brightness of high luminance stimuli to be underestimated. The brightness estimates hence converge toward the centre of the particular range used in those tests. Such a bias was not expected in side-by-side matching tests, where brightness judgements are made against a simultaneous reference source, but, nevertheless, it is apparent in the current results.

This can be explained as follows. In the side-by-side matching task the observer has a primary reference – a simultaneous stimulus. Where a number of different reference illuminances are used, the range appears to become apparent to the observer, as does the approximate centre of the range of responses, and this becomes a secondary reference toward which responses are biased. In the current brightness matching tests there were three reference illuminances (2.0 lux, 7.5 lux and 15.0 lux) and these were presented in a random order to participants, who made several matches to each reference illuminance (four in the main tests; eight in the null condition tests).

The null condition data (Table 4) identify the mean effect of this bias, and this can be applied to the main data as a correction factor. For tests carried out at 2.0 lux, the response contraction bias suggests matching illuminances are biased toward a higher level. Illuminances of the booth adjusted by the participant were thus divided by the null condition ratio of 1.05. For tests carried out at 15.0 lux, the response contraction bias suggests matching illuminances are biased toward a lower level. Illuminances of the booth adjusted by the participant were thus divided by the null condition ratio of 0.97. A correction was not applied to data for tests carried out at 7.5 lux.

The results of applying these corrections are shown in Table 6. Firstly consider the data broken down according to whether dimming was applied to the HPS lamp or white lamps. There are now only two pairs of illuminance ratios that are significantly different (CFL/HPS and MH2/HPS, both at 15.0 lux), reduced from six before application of this correction (Table 5). The correction factors used here are based on only a small number of tests and further work would be needed to confirm them. A correction factor of 0.92 would be needed at 15.0 lux to completely eliminate differences. Secondly, consider the overall mean illuminance ratios as corrected for a response contraction bias. These differ only very slightly from the uncorrected mean illuminance ratios (Table 2) which demonstrates that the bias had been successfully counterbalanced.

A previous analysis of null condition data concluded that participants have a tendency to set the adjusted stimulus to an illuminance slightly lower than the average, again a small but significant bias.[13] This effect was recently found again in further data[9] where several visual matches were made but at only one reference illuminance. It appears that when a visual match is made to several different reference illuminances then a response contraction bias is expected, but if there is only one reference illuminance then the adjusted stimulus will be set to a lower than expected illuminance. Further work is needed to investigate the persistence, magnitude and interaction of these two effects.

5. Summary

A significant and unexpected bias noted in the results of side-by-side brightness matching tests is suggested to be a response contraction bias. In the current study it is concluded that the bias does not significantly affect the main results. This is because the bias is small (maximum mean illuminance ratio in null-condition tests is 1.05) compared to the main results (minimum white/HPS mean illuminance ratio = 0.741) and the application of dimming control was counterbalanced to both the HPS and the white lamps. Furthermore, the data are supported by the results of side-by-side brightness ranking tests which also found that booths lit by the white lamps and HPS lamps are equally bright at an illuminance ratio of approximately 0.7. The implications for further work are a reminder of the need to counterbalance the application of dimming to both stimuli in side-by-side matching tests: this has not always been done in previous studies. In the current work the response contraction bias is suggested to have only a small effect, but further work should include null-condition testing to quantify the extent of bias given the specific conditions and procedures being used.

References

1 Fotios SA, Cheal C, Boyce PR. Light Source Spectrum, Brightness Perception and Visual Performance in Pedestrian Environments: A Review. *Lighting Research & Technology* 2005; 37(4): 271-294.

2 Fotios SA, Cheal C. *Designing for white light in subsidiary streets: Proceedings of the Urban Nightscape conference, Athens*. Thission: Hellenic Illumination Committee, 2006.

3 Fotios SA, Cheal C. *Investigations into the white light trade-off for lighting in subsidiary streets: Proceedings of the ILE Annual Conference, Manchester*. 2006.

4 Boyce PR. Investigations of the subjective balance between illuminance and lamp colour properties. *Lighting Research & Technology* 1977; 9: 11-24.

5 Fotios SA. Experimental conditions to examine the relationship between lamp colour properties and apparent brightness. *Lighting Research & Technology* 2002; 34(1): 29-38.

6 Fotios SA. Lamp colour properties and apparent brightness: a review. *Lighting Research & Technology* 2001; 33(3): 163-181.

7 Alman DH. Errors of the standard photometric system when measuring the brightness of general illumination light sources. *Journal of the Illuminating Engineering Society* October 1977: 55-62.

8 Bellchambers HE, Godby AC. Illumination, colour rendering and visual clarity. *Lighting Research & Technology* 1972; 4(2): 104-106.

9 Fotios S, Gado T. A comparison of visual objectives used in side-by-side matching tests. *Lighting Research & Technology* 2005; 37(2): 117-131.

10 British Standards Institution. BS5489-1:2003, *Code of practice for the design of road lighting —Part 1: Lighting of roads and public amenity areas*. London: BSI, 2003.

11 British Standards Institution. BS EN 13201-2:2003, *Road lighting - Part 2: Performance requirements*. London: BSI, 2003.

12 Tiller DK, Veitch JA. Perceived room brightness: pilot study on the effect of luminance distribution. *Lighting Research & Technology* 1995; 27(2): 93-101.

13 Fotios SA. An error in brightness matching associated with the application of dimming. *Lighting Research & Technology* 2001; 33(4): 223-231.

14 Poulton EC. Biases in quantitative judgements. *Applied Ergonomics* 1982; 13(1): 31-42.

Acknowledgement

This work was carried out with support from EPSRC, grant references GR/S25074/01 and GR/T09385/01. Lamps and luminaires were supplied by Philips Lighting and Urbis Lighting; the side-by-side booths were procured through a CIE-UK teaching and research grant.

Lamp type		CCT (K)	CRI (R_a)
HPS	70W SON-T	2000	25
CFL	55W PL-L	3000	82
MH1	70W CDO-TT	2800	83
MH2	70W CDM-T	4200	92

Table 1

Description of lamps used in side-by-side brightness matching tests. CCT and CRI are as reported in manufacturers literature.

White lamp	**CFL**			**MH1**			**MH2**		
Reference illuminance (lux)	2.0	7.5	15.0	2.0	7.5	15.0	2.0	7.5	15.0
Mean illuminance ratio (white/HPS)	0.696	0.718	0.734	0.731	0.733	0.724	0.680	0.724	0.741
Std. Dev.	0.158	0.119	0.172	0.176	0.161	0.161	0.197	0.167	0.203

Table 2

Results of brightness matching tests – mean illuminance ratio at equal brightness. In all cases, n=84 and the mean illuminance ratio departs significantly from unity ($p<0.01$).

Reference illuminance (lux)	**2.0**	**7.5**	**15.0**
Mean illuminance ratio (left/right)	0.99	0.99	1.00
Std dev	0.12	0.10	0.09
n	144	144	144
Departure from unity	n.s.	n.s.	n.s.

Table 3

Results of brightness matching null-condition tests: mean illuminance ratio of left-hand-booth to right-hand-booth at equal brightness. n.s. = not significant ($p>0.05$).

Reference illuminance (lux)	**2.0**	**7.5**	**15.0**
Mean illuminance ratio (test/reference)	1.05	0.99	0.97
Std dev	0.13	0.10	0.08
n	144	144	144
Departure from unity	**p<0.01**	n.s.	**p<0.01**

Table 4

Results of brightness matching null-condition tests: mean illuminance ratio of the test booth (adjusted illuminance) to the reference booth (fixed illuminance) at equal brightness.

Lamp combination			CFL / HPS			MH1 / HPS			MH2 / HPS		
Reference illuminance (lux)			2.0	7.5	15.0	2.0	7.5	15.0	2.0	7.5	15.0
White lamp is dimmed		Mean illum. ratio (white/HPS)	0.737	0.699	0.657	0.793	0.723	0.685	0.718	0.697	0.645
		Std dev	0.154	0.129	0.124	0.184	0.141	0.148	0.199	0.147	0.173
		n	42	42	42	42	42	42	42	42	42
HPS lamp is dimmed		Mean illum. ratio (white/HPS)	0.654	0.737	0.810	0.669	0.744	0.763	0.642	0.752	0.837
		Std dev	0.152	0.105	0.179	0.142	0.179	0.164	0.188	0.180	0.184
		n	42	42	42	42	42	42	42	42	42
Illuminance of adjusted booth relative to overall average.*			H	L	L	H	L	L	H	L	L
Difference between mean illuminance ratios			p< 0.05	n.s.	p< 0.01	p< 0.01	n.s.	p< 0.05	**p< 0.05	n.s.	p< 0.01

Table 5

Results of brightness matching tests broken down according to the application of dimming.

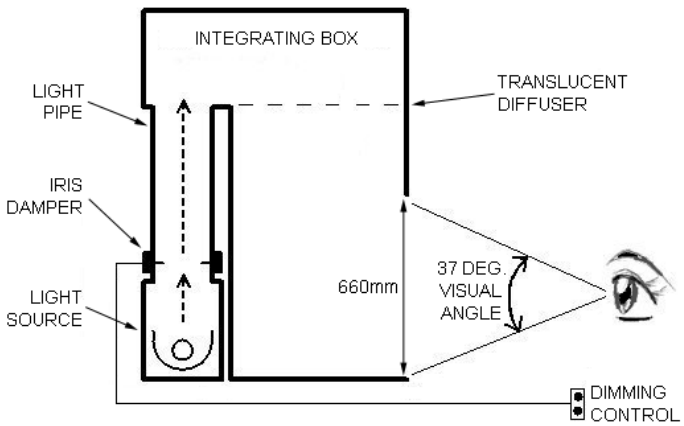
* H = Illuminance of adjusted booth is **higher** than overall average
L = Illuminance of adjusted booth is **lower** than overall average lower

** This significance level is found using the one-tailed test: it is not significant if the two-tailed test is applied.

	Lamp combination	CFL / HPS			MH1 / HPS			MH2 / HPS		
	Reference illuminance (lux)	2.0	7.5	15.0	2.0	7.5	15.0	2.0	7.5	15.0
White lamp is dimmed	Mean illum. ratio (white/HPS)	0.702	0.699	0.678	0.755	0.723	0.707	0.684	0.697	0.665
	Std dev	0.147	0.129	0.128	0.175	0.141	0.152	0.190	0.147	0.179
	n	42	42	42	42	42	42	42	42	42
HPS is dimmed	Mean illum. ratio (white/HPS)	0.687	0.737	0.786	0.702	0.744	0.740	0.674	0.752	0.811
	Std dev	0.159	0.105	0.173	0.149	0.179	0.159	0.197	0.180	0.178
	n	42	42	42	42	42	42	42	42	42
Difference between mean illuminance ratios		n.s.	n.s.	$p < 0.01$	n.s.	n.s.	n.s.	n.s.	n.s.	$p < 0.01$
Overall mean illum. ratio (white/HPS)		0.694	0.718	0.732	0.729	0.733	0.724	0.679	0.724	0.738
Std dev		0.153	0.119	0.162	0.165	0.161	0.157	0.193	0.167	0.193
n		84	84	84	84	84	84	84	84	84
Departure from unity		$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$

Table 6

Results of brightness matching tests corrected for a response contraction bias using the null-condition data.



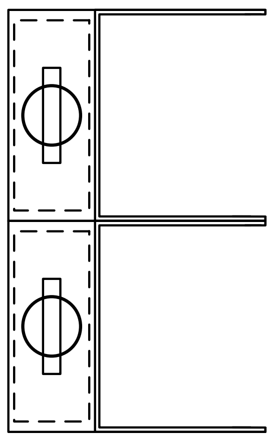


Figure 1

Vertical and horizontal sections through the side-by-side booths.

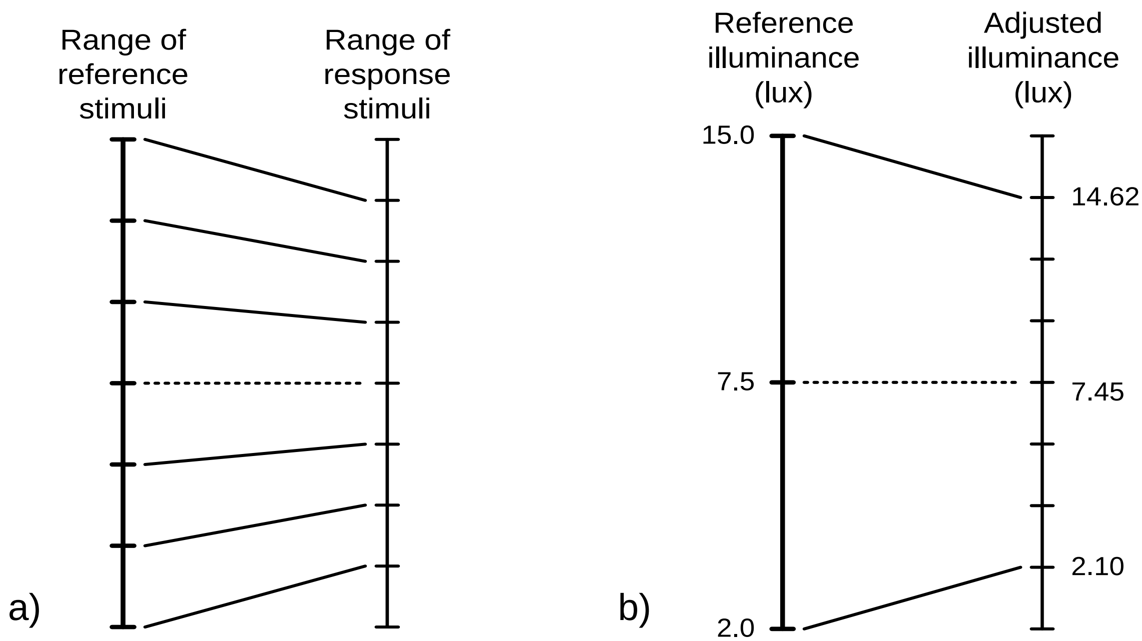


Figure 2

Illustration of the response contraction bias. (a) Responses tend to converge toward the centre of the range. (b) Mean illuminance of adjusted booth from the current null condition results (for illustrative purposes only, the *adjusted illuminance* range is not to scale).