Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 2557–2568
Vision Research
www.elsevier.com/locate/visres

Perceiving colour at a glimpse: The relevance of where one fixates

Eli Brenner *, Jeroen J.M. Granzier, Jeroen B.J. Smeets

Faculty of Human Movement Sciences, Vrije Universiteit, Van der Boechorststraat 9, 1081 BT Amsterdam, The Netherlands

Received 17 January 2007; received in revised form 11 June 2007

Abstract

We used classification images to examine whether certain parts of a surface are particularly important when judging its colour, such as its centre, its edges, or where one is looking. The scene consisted of a regular pattern of square tiles with random colours from along a short line in colour space. Targets defined by a square array of brighter tiles were presented for 200 ms. The colours of the tiles within the target were biased by an amount that led to about 70% of the responses being correct. Subjects fixated a point that fell within the target's lower left quadrant and reported each target's colour. They tended to report the colour of the tiles near the fixation point. The influence of the tiles' colour reversed at the target's border and was weaker outside the target. The colour at the border itself was not particularly important. When coloured tiles were also presented before (and after) target presentation they had an opposite (but weaker) effect, indicating that the *change* in colour is important. Comparing the influence of tiles outside the target with that of tiles at the position at which the target would soon appear suggests that when judging surface colours during the short "glimpses" between saccades, temporal comparisons can be at least as important as spatial ones. We conclude that eye movements are important for colour vision, both because they determine which part of the surface of interest will be given most weight and because the perceived colour of such a surface also depends on what one looked at last.

Keywords: Classification images; Colour vision; Eye movements; Retinal eccentricity; Gaze

1. Introduction

We know a lot about the perceived colour of uniform surfaces that one is fixating. However, in daily life we constantly shift our gaze and objects are seldom uniformly coloured. Eye movements raise two questions for colour vision research: whether where one is looking matters for the perceived colour of a surface of interest, and whether where one was previously looking matters in this respect. The issue of where one's gaze is directed could be particularly relevant for surfaces that are not uniform in colour, because it is not clear how the colour of parts of the *surface* are combined to determine the colour of the surface as a whole, and because the *retina* is not uniform in terms of analysing colour (e.g., Mullen, Sakurai, & Chu, 2005). Since the surfaces of objects around us are seldom completely uniform in reflectance, and the light falling on them is also seldom uniform, these issues are very relevant for how we perceive colours. Fig. 1 shows an example of how different the colour of the light from various parts of an object can be under natural conditions.

If we stare at a scene for some time and then divert our gaze to a large blank surface, we are likely to temporarily see an afterimage of the original scene in complementary colours. This afterimage arises because photoreceptors in the retina have adapted to the light coming from different parts of the original scene and therefore respond differently to exposure to the light from the blank surface (the notion of such a local change in sensitivity is attributed to Fechner by von Kries (1905, pp. 205–206)). Such an afterimage requires prolonged fixation of the first scene, but very short exposure to a coloured surface can also influence the perceived colour of a subsequently presented target surface (Cornelissen & Brenner, 1995). Thus where one was previously looking matters for colour vision. The notion that a

---

[*] Corresponding author.
E-mail address: e.brenner@fbw.vu.nl (E. Brenner).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.06.003

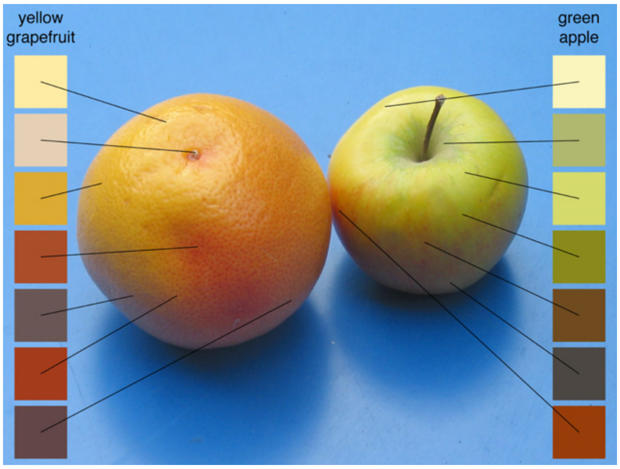

Fig. 1. A grapefruit and an apple on a blue table under natural daylight. The squares show the colours of pixels at various positions across the surfaces, illustrating how the light reaching the eye (or camera) depends on both the surface structure and the illumination (in particular the contribution of light reflected by the table). Why do we consider the grapefruit to be yellow and the apple to be green? (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

combination of eye movements and retinal adaptation can contribute to colour vision is often acknowledged, and there is clear evidence that restricting eye movements can make a large difference in a colour-matching task (Cornelissen & Brenner, 1991). However, eye movements may not only influence the perceived colour by exposing the fovea successively to different parts of the scene. Eye movements also determine precisely which part of the scene the fovea will be exposed to. Very little is known about whether the precise point at which one is directing one's gaze makes any difference for the perceived colour. A recent study by Hansen and Gegenfurtner (2005) suggests that it does.

Hansen and Gegenfurtner used a technique called classification images, in which subjects had to detect coloured targets embedded in noise, to determine human chromatic tuning in various directions of colour space. Apart from their findings regarding chromatic tuning they also found that subjects clearly relied most on the colours near the fixation point. This is somewhat surprising because there are reasons to expect the contrast at objects' edges to be particularly relevant when judging their colours (e.g., Brenner, Ruiz, Herraiz, Cornelissen, & Smeets, 2003; Krauskopf 1963; Shevell & Wei, 1998; Wachtler, Albright, & Sejnowski, 2001; Yund & Armington, 1975), even from studies based on similar methods but for achromatic stimuli (Dakin & Bex, 2003).

A possible reason for the edges not being particularly relevant in Hansen and Gegenfurtner's (2005) study is that the emphasis on border contrast is stronger for uniform surfaces than for colourful ones (Brenner, Granzier, & Smeets, 2007), and Hansen and Gegenfurtner's targets were colourful. However the absence of any indication that the edges were relevant suggests that there may be a more fundamental reason. In their experiment, the position of the edges had to be determined on the basis of the colour. Without independent information about where to separate the image into regions with different colours, there can be no emphasis on the position of the edges (based on image properties alone) until after the colour has been determined. Under such conditions, in terms of strategic decisions, it is probably safer to rely on the target centre.

In the present study we examine Hansen and Gegenfurtner's (2005) conclusion that subjects gave most weight to points near fixation in more detail. In their study the fixation point was at the centre of the target (and of the display). If the colours are averaged within large receptive fields (as proposed in Blakeslee, Pasieka, & McCourt, 2005; Blakeslee & McCourt, 1997) then it is reasonable for the largest response to arise when the scale of the receptive field coincides with that of the target, so one may expect most weight to be given to the points near the centre of the target. Moreover, as mentioned above, if one does not know where the target's edges are, it may be best to rely on the colour far from the edges, and one can certainly not give the edges more weight. We therefore set up similar experiments to those of Hansen and Gegenfurtner (2005) to specifically examine the spatial aspects of colour vision. The main differences between our stimuli and theirs were that we moved the fixation point away from the centre of the target, in order to dissociate an emphasis on the target centre from one on the fixation point, and introduced a clear luminance contrast at the edges of the target so that subjects could clearly identify the borders of the target. Since this meant that the target was always visible we used a colour discrimination task rather than a detection task.

2. Methods of the main experiment

We used the classification images technique (see *Journal of Vision, Volume 2(1)*, for many examples) to identify positions that were particularly important for judging a target surface's colour. The principle of this technique is that random noise is added to an image to mask a target. Subjects are shown many examples of the same target hidden within different patterns of noise, and are asked to respond to the target. The noise patterns that are present when subjects respond correctly are then compared to the noise patterns that are present when subjects respond incorrectly to the same target. A systematic difference between the noise patterns indicates that the corresponding part of the target is critical for the subject's decision; when the noise enhances the critical part of the target the chance of a correct response is larger.

2.1. Stimuli

We presented images at 120 Hz on a computer screen that was 140 cm from the subject. The screen was 40.0 by 29.8 cm (1024 by 768 pixels). The target was a square that was always at the centre of the screen, and was either 4.2° by 4.2° or 2.4° by 2.4° (sides of 10.5 or 5.9 cm; 270 or 150 pixels). The target was at the centre of an 8.1° by 8.1° background (sides of 19.9 cm; 510 pixels). Both the target and the background were composed of square tiles. There were either 2601 (51 by 51) or 289 (17 by 17) tiles, with sides of either 0.16° or 0.47° (0.4 or 1.2 cm; 10 or 30 pixels). There was a black 0.1° (0.2 cm; 6 pixel) diameter fixation point 1.0° (2.3 cm; 60 pixel) below and 1.0° to the left of the target centre.

The tiles' colours were taken from a single line in CIE 1931 colour space: $y = 0.626 - x$. This line has an orientation of $-45°$ and passes through standard illuminant C ($x = 0.310$; $y = 0.316$). The random component of the colours attributed to the tiles (the noise pattern) was determined by finding as many equidistant points along a line segment with length 0.02, centred at standard illuminant C, as there were tiles in the display. On each presentation, each of these points was associated with a randomly chosen tile. The points that were associated with tiles that were part of the target were then shifted along the above-mentioned line in colour space to differentiate between 'red' and 'green' targets. This colour shift was the same for all tiles within the target, but it differed between subjects, and between red and green targets (as will be explained below). If the tile was part of the background, the colour at the associated point on the line in colour space was rendered with a luminance of 20 cd/m^2. If it was part of the target it was rendered with a luminance of 21 cd/m^2. Obviously our 8 bit per gun resolution (calibrated with a Minolta CS-100A Chroma Meter; Minolta Camera Co. LTD., Japan) cannot ensure that we always render exactly the correct values. We rendered each tile as accurately as possible and based all our calculations on the actually rendered values, so the limited resolution does not affect our analysis.

Between target presentations one of three kinds of stimuli was shown: *blank*, *same* or *different*. The first was a *blank* screen with the same luminance and average colour as the target's background ($x = 0.310$; $y = 0.316$; $Y = 20\ \text{cd/m}^2$). The second was the *same* tiled pattern that formed the random component of the colour during target presentation, but without the higher luminance and colour shift that identify the target. Thus when the target appeared (and disappeared) the colour and luminance of the tiles that were not part of the target did not change. The tiles that were part of the target changed luminance (to identify them as the target) and all changed colour by exactly the same amount. The third option was that a *different* random pattern of tile colours was presented before and after target presentation than the one presented during target presentation. In this case the whole pattern of random colours changed when the target appeared, both within and outside the target, although the average colour of the random pattern remained the same.

2.2. Procedure

Red and green targets were presented in random order. Subjects sat with their fingers on the 'r' and 'g' keys of the computer's keyboard. Their task was to indicate the colour of the target (red or green) by pressing the corresponding key. Targets were presented for 200 ms followed by 1000 ms during which the subject could respond (Fig. 2). Trials in which the subject did not respond or responded too late were not repeated. The fixation point remained visible throughout the experiment and was to be fixated at all times except during breaks that were provided after every 100 targets. Subjects could decide when they wanted to proceed after each break. There were 6 sessions that differed in details such as the size of the tiles, the size of the target, and what was seen during the 1000 ms between the targets. When the screen was not *blank* between target presentations the same pattern was shown for 500 ms before and after target presentation, so the pattern of tiles changed half way through the interval.

Each session started with a series of targets that were used to estimate the magnitude of the colour shift that is needed for 67% of the responses to be correct. Within this series the additional shift in the colour of the target changed according to a staircase procedure, whereby errors resulted in a change in the colour shift of 0.002 in the direction of the designated target colour, and correct responses resulted in a change in the colour shift of 0.001 in the opposite direction. There was a separate staircase for 'red' and 'green' as the designated test colour. The two staircases each continued until the steps had changed direction 30 times. The values at which the last 20 changes in direction had occurred were averaged to estimate the shift that would give 67% correct responses. The colour shift was fixed at these estimates (one for red and one for green) for the rest of the session.

Table 1 provides details of the differences between the six sessions, which include the pattern that was visible when there was no target, the size of the target and of the individual tiles, the number of subjects tested, and the number of targets presented to each subject (excluding the initial staircase). In session 2, targets that were preceded and followed by the same pattern were randomly interleaved with ones that were preceded and followed by a different pattern (500 of each). In session 5, the colour shift that defined targets as 'red' or 'green' was applied to all tiles rather than only to those of the target, to check whether the colour contrast that was introduced by the colour shift is in any way critical.

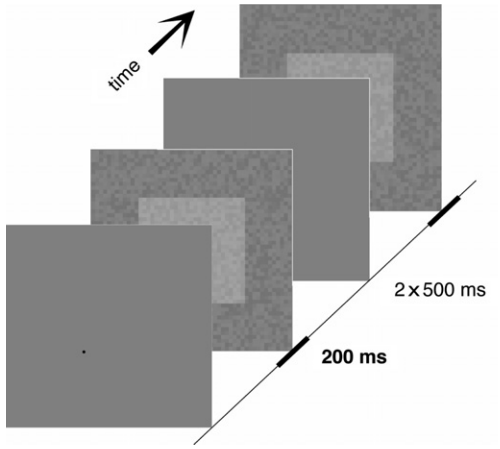

Fig. 2. Schematic representation of two targets in the session with small tiles, a large target, and a blank screen during the interval. In reality, the variability between tiles was in colour rather than luminance, and besides being brighter than the background the target was also slightly reddish or greenish. Targets were presented for 200 ms followed by a 1000 ms interval during which the subject was expected to indicate whether the target had been red or green (in sessions in which the screen was not blank the same pattern was shown for 500 ms before and after each target presentation).

2.3. Subjects

The subjects were the three authors and nine of our colleagues. All subjects had normal colour vision (assessed with Ishihara colour plates). Only the authors were aware of the purpose of the study. The three authors took part in all sessions except session 4, in which only two of the authors took part. Two of the other subjects took part in all sessions, three took part in all sessions except session 5, and the remaining four took part in between 2 and 5 sessions. The ethical committee of the Faculty of Human Movement Sciences approved the experiments.

2.4. Analysis

For each colour shift (towards red or green) and response ('r' or 'g' key) we first determined the average colour of each tile; i.e., the average colour presented at that position across presentations on which that response was given to that kind of target. We did so separately for each subject and session (and separately for the two conditions in session 2). A single number could represent each of these average colours because all the colours that we presented were from a single line in CIE 1931 colour space. We then determined each tile's contribution (C) to the subject's response by subtracting the average colour for incorrect responses from that for correct responses (for each colour shift, subject and condition). Note that because only the random distribution of the colours differed between the targets for which different responses were given, the average colour across all tiles was always the same so the average difference (the average value of C across all tiles) is zero.

Since we were not interested in differences between red and green targets, we averaged the values of C for the two target colours (after inverting the sign of one of the two so that a higher value always corresponded with naming the colour of the shift). We then averaged these values across subjects, and for the left part of Fig. 4 even across conditions. We used weighted averages to compensate for the fact that the reliability of C

Table 1
Experimental details

Session	1	2	3	4	5	6
Pattern	Blank	Same/different	Different	Blank	Different	Different
Width of tiles	0.16°	0.16°	0.16°	0.47°	0.47°	0.47°
Target width	4.2°	4.2°	4.2°	4.2°	4.2°	2.4°
Subjects	10	10	10	10	5	10
Targets each	500	500/500	500	500	1000	1000
Colour shift*	11	8	13	18	15	27
Targets missed (%)	2	9/9	10	3	9	11
Correct (%)	73	72/72	78	70	71	71

* Percentage of range, averaged across subjects and target colours.

depends on the number of responses and their distribution across the two possible choices ('r' or 'g'). The weights were based on the binomial variance. The weight assigned to each contribution C depended on the relative magnitudes of $n_r n_g/(n_r + n_g)$ where n_r and n_g are the number of times that the subject pressed the 'r' and 'g' key, respectively. The weighted averages for all the tiles were plotted as grey-scale images and were used for the further analysis with the help of a simple model that we will present after the initial description of the results.

3. Results of the main experiment

Table 1 gives the colour shift that was estimated from the staircases to be required to obtain 67% correct responses, as well as the percentage of targets that were missed and the actual number of correct responses obtained during the rest of the session. The estimate of the required colour shift was larger for the larger tiles (sessions 4–6), despite the fact that the noise pattern contained the same range of colours for all sessions. A possible explanation for this is that subjects base their judgments on a (weighted) average of the colours within some area. For larger tiles there will be fewer tiles within this area. The average of a few random values from the above-mentioned range is more variable than the average of many such random values, so less of the variability introduced by the noise pattern will be averaged away for large tiles. The estimate of the required colour shift was particularly large for the smaller target (session 6), suggesting that the area that is considered is larger than the smaller target and that tiles within the target are given more weight. As will become evident later this explanation is consistent with our further analysis of the data.

Subjects missed fewer targets when the screen was empty (*blank*) between targets, than when it contained a textured pattern between targets. Apparently the additional variability in local colour and the additional change in pattern half way through the response interval made it more difficult to detect or identify the colour of the targets (masking). The percentage of correct responses (only considering targets that were not missed) was slightly but systematically higher than we anticipated (on average 72% rather than 67%). Apparently subjects performed better during the main part of the experiment than during the initial staircase. Perhaps the fact that the magnitude of the shift was fixed helped them to interpret what they saw. It may also just take some trials to get accustomed to the stimuli at the beginning of each session. It is not likely to simply be a matter of practice because most subjects took part in many sessions, and did not systematically require smaller shifts in later sessions.

Fig. 3 shows the ten subjects' individual classification images for session 4. The brightness indicates the difference between each tile's average colour on trials in which a correct and an incorrect response was given (C as explained in the analysis section). The brighter the colour the larger the value of C. An exceptionally large value of C (bright tiles) suggests that the tile contributes strongly to the subjects' judgments, with a tendency to respond with the colour of that tile. A very negative value of C (dark tiles) also suggests that the tile contributes strongly to the subjects' judgments, but with a tendency to respond with the other colour than that of the tile. For most subjects we see a tendency for tiles near the fixation point to be brighter than the rest.

The choice of session 4 for Fig. 3 is not arbitrary. Having many tiles is an advantage for telling where exactly the tiles that one relies on most heavily are situated, but the contribution of each tile is smaller when there are many tiles, making the overall picture less clear. For our 500 trials per session (or less because missed trials were not repeated) the individual images are not very clear (Fig. 3). Hansen and Gegenfurtner (2005) present somewhat clearer individual data for a similar spatial resolution, presumably because they had 2000 trials per session. Since there were no obvious fundamental differences between the patterns of data for the different subjects—either here or in Hansen and Gegenfurtner's study—there is no reason to believe that different subjects use fundamentally different strategies, so we will present all further data averaged across subjects.

To get the clearest possible picture of the spatial pattern in the classification images the left panel of Fig. 4 shows the average data of all the sessions with small tiles (sessions 1–3). In averaging across these sessions we assume that the overall spatial pattern does not differ fundamentally between the sessions (as is later verified). For the later analysis the sessions are treated separately. The right panel of Fig. 4 shows the average data of session 6 and the left panel of Fig. 5 shows the average data of session 4 (average of the

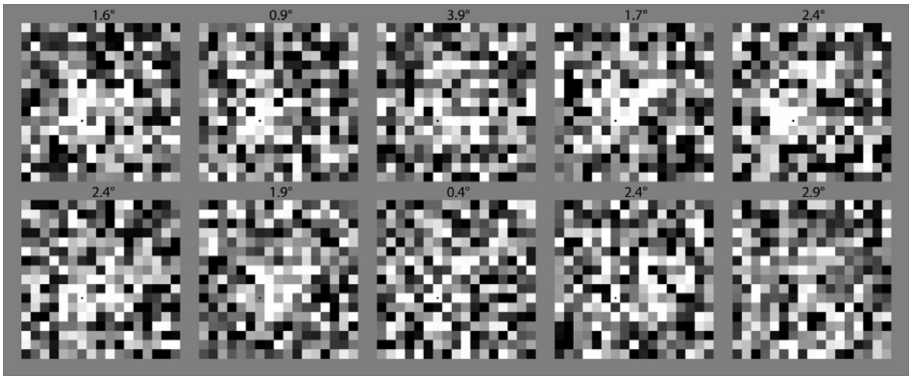

Fig. 3. Classification images for individual subjects in session 4 (large tiles and target; blank screen between targets). Exceptionally bright tiles contribute strongly to the subjects' judgments, with a tendency to respond in the colour of that tile. Exceptionally dark tiles also contribute strongly to the subjects' judgments, but with a tendency to respond in the other colour than that of the tile. The numbers are fit values for σ (for details see text).

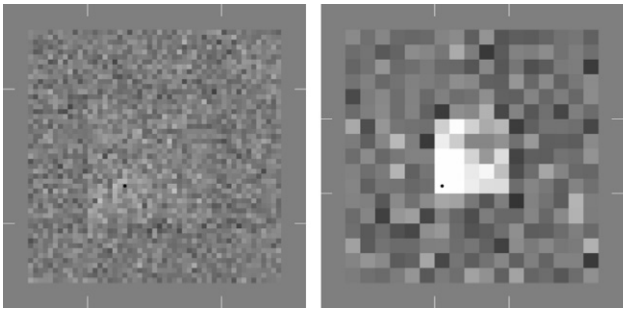

Fig. 4. Classification images for the average of all subjects, for the average of all sessions with small tiles (left; sessions 1–3) and for the session with large tiles and a small target (right; session 6). The extent of the target is indicated by the white lines at the edges.

ten panels of Fig. 3). In all the above-mentioned panels the brightest parts are to be found in the lower left part of the target, near the fixation point. The tiles' contributions to the subjects' responses appear to decrease gradually with their distance from fixation, with an abrupt transition at the target's borders. The contrast at the borders was not given any additional weight, but it is difficult to see whether the relationship between tiles' colours and the response inverts or just stops at the border. In order to determine whether the tiles near fixation also have a stronger influence on subjects' responses when they are outside the target, to estimate the position and spatial extent of the region that subjects rely on most heavily, and to examine whether the colour of the tiles before (and after) the target was presented matters, we fit the data to a simple model.

4. Modelling the results of the main experiment

Although the classification images in Figs. 4 and 5 are clearly not uniform, the spatial bias that we are interested in does not completely overshadow the variability between tiles, despite averaging over very many targets and subjects. In order to get a better grip on the classification images, and in particular to help answer the questions raised in the previous section in a quantitatively justifiable manner, we fit a simple model to the data. The model that we fit to the data starts with a simple two-dimensional normal distribution. The position of the peak is determined by two parameters (x_0, y_0), the width of the peak is determined by a third parameter (σ), and the height of the peak is determined by a fourth parameter (h).

To model the abrupt transition at the target's borders the height of the peak is multiplied by a fifth parameter ($b_{x,y}$) that has a different value for tiles that are inside and outside the target. Due to the way we designed the stimuli and analysed the results, the sum of the contribu-

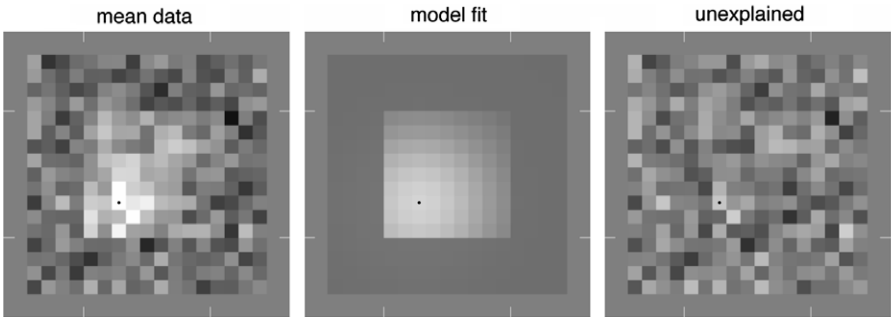

Fig. 5. Classification images for session 4 (large tiles and target; blank screen between targets). The left panel shows the raw data (as in Figs. 3 and 4). The middle panel shows the best fit of the four-parameter (σ, h, a and $b_{x,y}$) version of the model. The right panel shows the variability in the data that is not explained by the model. The model accounts for most of the systematic differences between the tiles.

tions of all the tiles is zero $\sum_{x,y} C_{x,y} = 0$, so we had to include an offset (a) to the model (to compensate for the large positive peak). Since the influence of this sixth parameter is spatially uniform it is of no further interest to us. Finally, whenever there was a different noise pattern between target presentations, we simultaneously fit the tiles' estimated contributions (the classification images) during target presentation and before and after target presentation. To limit the number of free parameters that this introduces we assumed that the spatial properties of the tiles' contributions (x_0, y_0, σ, $b_{x,y}$) were the same for the interval between targets as during target presentation, so we added a single multiplication factor that relates the influence of the preceding colour to that of the colour during target presentation (the seventh parameter c_t). Together we get

$$C_{x,y,t} = c_t\left(a + b_{x,y} h e^{\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}}\right) \tag{1}$$

where x and y are the positions of the tiles, $b_{x,y} = 1$ for tiles within the target (and a fit value for tiles outside the target) and $c_t = 1$ for the tiles' contributions when presented during target presentation (and a fit value when presented before and after target presentation). This equation uses 7 parameters to fit two classification images whenever a different noise pattern was presented during the interval between target presentations, and 6 parameters to fit one classification image in all other cases. We fit Eq. (1) to the data of each session (and condition within session 2) by minimizing the sum of squares of the difference between the model prediction ($C_{x,y,t}$) and the measured data (the values corresponding with the tiles' brightness in Fig. 4 and the left panel of Fig. 5) for all the tiles in the image.

Beside fitting the data with all the above-mentioned free parameters, we also fit the data with several more constrained models, and examined with the help of Akaike's information criteria (AIC$_c$; Motulsky & Christopoulos, 2004) whether the more constrained models are more likely to be true considering the reduction in the number of free parameters (see Appendix A). We first evaluated the trivial model of there being no spatial bias at all ($C_{x,y,t} = 0$). We then evaluated alternative models in which the peak of the normal distribution was confined either to the fixation point or to the target centre. Finally we evaluated a model in which $b_{x,y}$ is zero for tiles outside the target, and a model in which c_t is zero for tiles presented in the interval between targets, to determine whether we could demonstrate any influence at all of the background and of the image presented during the interval between the targets.

5. Modelling results

The trivial model of there being no spatial bias at all was much less likely to be true than our complete model (according to AIC$_c$) for all sessions. In accordance with our impression from the figures, fitting the values of x_0 and y_0 usually gave coordinates that were close to the fixation point. In all cases, the model in which (x_0,y_0) was confined to the *centre of the target* was judged to be less likely to be true than the model in which we fit the values of x_0 and y_0 (for details of this and the following comparisons see Appendix A). In contrast, for sessions 1, 3 and 4, and for the trials with the 'same' pattern in session 2, the model in which (x_0,y_0) was confined to the *fixation point* was more likely to be true than the model in which we fit the values of x_0 and y_0. For session 5, and for the trials with the 'different' pattern in session 2, the model with fit values of x_0 and y_0 was more likely to be true, but barely so (less than 2 times as likely). For session 6, the model with fit values of x_0 and y_0 was clearly more likely to be true, but the optimal position of the peak was far outside the image. Thus, altogether, confining the peak to the fixation point improved the model, so we conducted all subsequent tests with the coordinates of the fixation point as x_0 and y_0.

For sessions 1, 4 and 6, the model with no effect of tiles outside the target (thus $b_{x,y} = 0$ for such tiles) was judged to be about twice as likely to be true than the model in which such tiles did have an effect (i.e., with a fit value of $b_{x,y}$ for such tiles). In the other four cases it was judged to be less likely to be true, although the alternative model (with a fit value of $b_{x,y}$) was only judged to be more than twice as likely to be true for session 3 where it was 8.3 times as likely to be true. For session 3, the model with no effect of tiles presented in the intervals between target presentations (thus $c_t = 0$ for such intervals) was judged to be 2.4 times more likely to be true. For sessions 2, 5 and 6 the model in which tiles presented during the interval did influence the subsequent response was judged to be at least 4.9 times as likely to be true. Thus, altogether, it appears that tiles in the background and ones presented before (and after) the target do influence subjects' judgments, so we did not reduce the number of free parameters any further by fixing the values of $b_{x,y}$ and c_t to zero for tiles that are not part of the target.

The middle panel of Fig. 5 shows the best fit of the final four-parameter (σ, h, a and $b_{x,y}$) model for session 4 (c_t is irrelevant since there was a blank image between targets in this session and (x_0,y_0) is the fixation point). This is one of the sessions in which we found no evidence of a contribution of tiles in the background. That the model captures the general pattern in the underlying classification image (left panel) is evident from examining the difference (right panel) between the data and the model, in which there is no evident pattern.

Having found an acceptable model with which to describe our data, we can turn to the values of the fit parameters to get a better idea of the contributions of tiles at different positions. The values of h (black bars), $b_{x,y} \times h$ (for the background, when $b_{x,y}$ is not 1; grey bars) and $c_t \times h$ (for the interval between the targets, when c_t is not 1; white bars) for each session are shown in Fig. 6b. When a model in which $b_{x,y} = 0$ for tiles outside the target was judged to be more likely to be true, or in which $c_t = 0$ for tiles presented before and after the target presentation

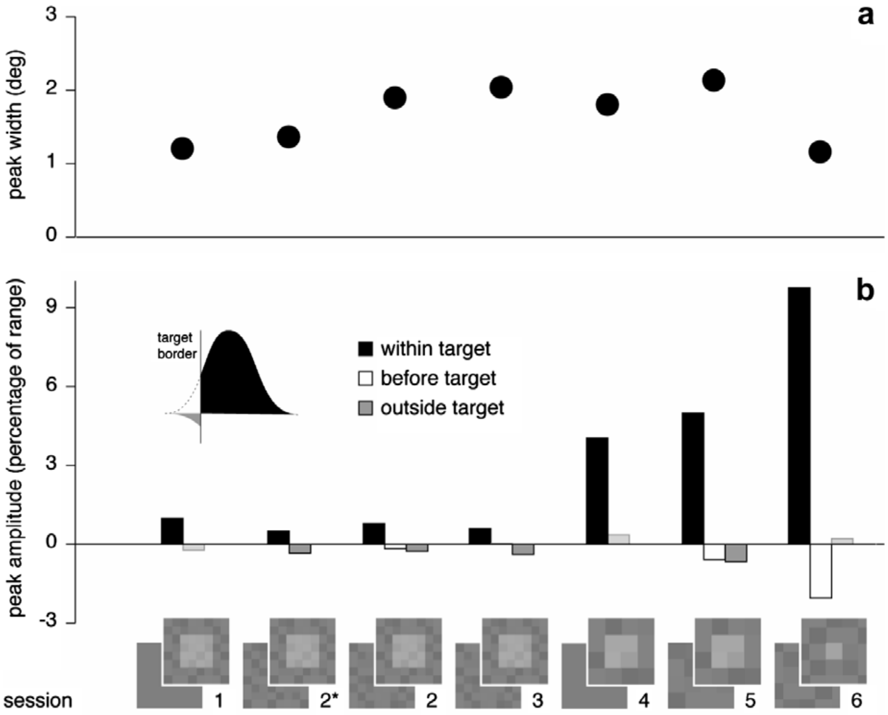

Fig. 6. (a) Values of σ for each fit. (b) Values of h (black bars), $b_{x,y} \times h$ (grey bars) and $c_t \times h$ (white bars) for each fit (as a percentage of the range of random colours). The influence of tiles outside the target and of tiles presented before the target appears is usually negative: a green surrounding, or green tiles near fixation before the target appears, decreases the probability of the target being judged to be green. The influence of tiles outside the target is smaller than that of tiles within the target when they are at the same distance from fixation ($|b_{x,y}| < 1$; see inset). The pale bars indicate that removing this component from the model would give a better fit ($\mathrm{AIC_c}$). In session 2 there was a separate fit for each of the two conditions (the asterisk indicates the condition with the same random pattern of colours during as before and after target presentation).

was judged to be more likely to be true, the values are shown in pale colours. That the value of h is larger when the tiles are larger (sessions 4–6) and when the target is smaller (session 6) is a direct consequence of there being fewer tiles to consider (with the same range of colour values; see Appendix B). The values of $b_{x,y}$ and of c_t are generally negative, indicating that both spatial and temporal colour contrast contribute to the perceived target colour.

The values of h, $b_{x,y} \times h$ and $c_t \times h$ (as shown in Fig. 6b) represent the influences of tiles within, outside and before the target. They are shown as a percentage of the range of the random colours, so the maximal possible value would be about 50% if subjects were to rely on a single tile (because the tiles' colours were chosen at random from within the range), and decreases with the number of tiles that contribute to the decision (see Appendix B). There is no indication that applying the colour shift that determined whether the target was 'red' or 'green' to all tiles, as we did in session 5, rather than only to those of the target, as we did in the other sessions, is critical in any way. We cannot tell whether the impression that the colours of tiles in the background have little influence when there is a blank screen between the targets (sessions 1 and 4) and when the target is small (session 6) is a coincidence or a real finding. The lack of effect of the tiles presented during the interval between the targets in session 3 is presumably a coincidence, because the effect does appear to be present for the identical trials in session 2, as well as in the other two sessions in which a different pattern was presented between targets. Thus Fig. 6b illustrates that spatial and temporal contrast contribute to the perceived colour, but that their effects are modest in comparison with the effect of the colour of the target itself.

Fig. 6a shows the width of the normal distribution that best fits the data. The value of σ varied between 1.2° and 2.1°. The width may be slightly larger when a different pattern of tiles was presented between the targets. It may also be slightly larger for the larger tiles, but it certainly does not scale linearly with tile size. It appears to be smaller for the small target. The numbers in Fig. 3 show the values of σ for individual subjects, assuming that the peak is at fixation. These values are within the same range as estimates of the spatial extent of several other chromatic interactions (e.g., Brenner & Cornelissen, 1991; Granzier, Brenner, Cornelissen, & Smeets, 2005).

One problem with taking the differences between subjects and between conditions too seriously is that we did not control the subjects' eye movements. We asked them to fixate the fixation point, but it is quite possible that they had better fixation when there was no pattern on the screen between target presentations (so that the fixation point was seen in isolation) and when the target itself was small. Similarly subjects may have fixated slightly less accurately when the tiles were large. Thus we cannot conclude that the spatial bias (i.e., the width of the normal distribution in the model) differs fundamentally between the sessions. What we can conclude is that there is a very robust bias towards relying on the colours of tiles near where one is looking (i.e., near fixation).

6. Testing a simple prediction

Subjects clearly mainly relied on the colours of the tiles near where they were looking, rather than on the average colour of the tiles within the target or on the colour contrast at the target's borders. They consider whether the tile is part of the target or not, giving more weight to tiles within the target ($|b_{x,y}| < 1$ for tiles outside the target), and a 'negative' weight to tiles outside the target ($b_{x,y} < 0$ for tiles outside the target; colour contrast). Together, these findings predict that the influence of the colour of the background will depend on where one is looking. Specifically, the shift in the apparent colour of a target away from the colour of the background (chromatic induction) should be larger if one fixates a point near the edge of the target (or even outside the target) than if one fixates the centre of the target, because the background should be given more weight if it is closer to where one is looking. We therefore conducted an additional experiment to examine whether the influence of the background is indeed larger when one fixates further from an object's centre.

6.1. Methods

We used a *nulling* method to quantify the influence of the background: we examined how slightly reddish and greenish backgrounds influenced the (average) colour of the target at which subjects were as likely to indicate that the target was red as that it was green. As in session 1 of the main set of experiments, a field of 0.16° wide tiles—within which the target was recognisable as a 4.2° wide square of brighter tiles—was presented for 200 ms, and only the fixation point was visible on a uniform background ($x = 0.310$; $y = 0.316$; $Y = 20\ \mathrm{cd/m^2}$) during the interval between presentations. However, in this experiment the fixation point was at the centre of the screen and the target was presented at various positions. Moreover, the whole screen (16° by 12°) was tiled during target presentation. The tiles' colours were initially chosen at random from within the same range as in the main experiment. The colours of the tiles in the *background* were then shifted by 0.01, either towards a redder or towards a greener hue (along the same line in colour space as in the main experiment), to obtain reddish and greenish backgrounds on separate trials. The colours of the tiles within the target were shifted (along the same line) by an amount that depended on the subject's performance on previous trials (as will be explained below).

Targets were presented at 4 eccentricities along a diagonal from the lower left to the upper right of an invisible square centred on the fixation point. The distance of the target centre from the fixation point was 0°, 1.4°, 2.7° or 4.1°. The corners of the target were 3.0° from its centre, so there were targets that were centred at fixation (0°), targets with their centre at the same distance from the fixation point as in the main experiments (1.4°), targets that barely included the fixation point (2.7°), and targets that appeared while the fixation point was in the background (4.1°). For each distance the target could be to the lower left or to the upper right of fixation, to discourage subjects from shifting their gaze in one of these directions in anticipation of the target presentation.

Five subjects took part in the experiment, including the authors. They each took part in a single session with 1600 targets: 4 distances between target centre and fixation, 2 directions (lower left or upper right), and 2 background colours (reddish or greenish), each presented 100 times. After each target presentation, subjects pressed the 'r' or 'g' key to indicate whether the target had been red or green. There was no time limit for the response. Once they responded the next target appeared.

In order to quantify the influence of the two different backgrounds on the perceived target colour we determined the target colour for which subjects were as likely to press each of the two keys. This was done with a separate staircase for the red and green backgrounds for each distance of the target (irrespective of the direction). The amplitude of the colour shift in the tiles within the target depended on the subject's responses on previous presentations that contributed to that staircase. It changed in steps of 0.001. On the first trial within each staircase the amplitude was zero. Whenever the subject judged the target to be red, the amplitude changed to make the target slightly greener (or less red) next time. Whenever the subject judged the target to be green, the amplitude changed to make the target slightly redder (or less green) next time.

To estimate the target colour for which subjects were as likely to press each of the two keys from these staircases we fit a sigmoidal function to the percentage of 'r' responses as a function of target colour (least squares with weights based on the square root of the number of presentations) and determined the value of the target colour at which this function crossed 50% 'r' responses. We did this for each subject, colour of the background, and distance of the target centre from fixation. We took the difference between the values for the corresponding red and green backgrounds as our measure of the effect of background colour. We divided this difference by 0.02 (the difference between the colours of the backgrounds) and multiplied the result by 100, to express the influence of the background as a *percent induction*. We did this separately for each subject and distance of the target centre from fixation. A value of zero would mean that the background has no effect. A value of 100% would mean that the target colour has to change by the same amount as the background for it to remain neither reddish nor greenish.

6.2. Results and discussion

The average results of the five subjects are summarised in Fig. 7. The extent to which the background influenced the perceived colour did indeed depend on the position of the target relative to fixation, with the largest effect when subjects were fixating outside the target (repeated measures ANOVA; $p < 0.001$). This supports our interpretation of the results of the main experiments. However, the colour of the background had a stronger influence on the percept (almost 50% induction even when fixating within the target; Fig. 7) than one would expect from Fig. 6b, where the fact that the grey bars are considerably smaller than the black bars indicates that the effect of the background is modest even if one takes the retinal eccentricity into account. Whether the influence was larger in this experiment because of the uncertainty about the upcoming target's position, because the background was chromatically biased, because there is also a component to induction that is not spatially localised and is therefore not revealed by the classification image technique, or for some other reason, remains to be seen.

7. A temporal experiment

Comparing the white and grey bars in Fig. 6b suggests that temporal colour contrast (white bars) may be as important as spatial colour contrast (grey bars). In order to get a better impression of the role of temporal contrast we conducted a final experiment in which we presented a much simpler image, a uniform target within a uniform background, but changed the colours of each of the surfaces on every frame (see Neri & Heeger, 2002, for an example of a similar approach). Again targets were presented for 200 ms and were recognizable by having a higher luminance than the background (21 rather than 20 cd/m^2), and again their colour was shifted so that subjects would respond correctly on about 67% of the trials (as determined on the basis of an initial staircase procedure), and again subjects were given 1000 ms to respond. The variable component of the colours of both the target and the background were chosen at random from within the range used in the main experiments. The analysis was also equivalent to that of the main experiments, except that we now compared the images for the same moment (frame) relative to target onset, rather than for the same positions on the screen. We did so separately for the two surfaces: target and background. The target was a 4.2° square. Ten subjects took part in the experiment, including the three authors.

7.1. Results and discussion

Colours presented up to 300 ms before target onset, at the position at which the target will later appear (leftmost solid symbols in Fig. 8), have a negative contribution, meaning that they increase the chance of choosing the complementary colour. From slightly after target onset the colour of the target has a positive contribution, meaning that subjects are more likely to choose the presented colour (as they should). The decline in the contribution of the colours presented at the end of target presentation is at about the correct moment but it is quite gradual. The difference between the changes at onset and offset may have to do with fast adaptation to the common chromatic component

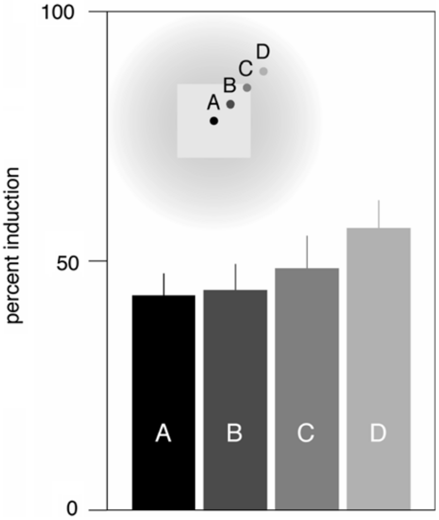

Fig. 7. Testing the prediction that the influence of a coloured background depends on the position of the target relative to fixation. The inset shows the position of fixation relative to the target (actually the position of the target changed rather than that of the fixation point). Bars show mean values across subjects with standard errors. The different positions of the target relative to fixation are represented by different shades of grey.

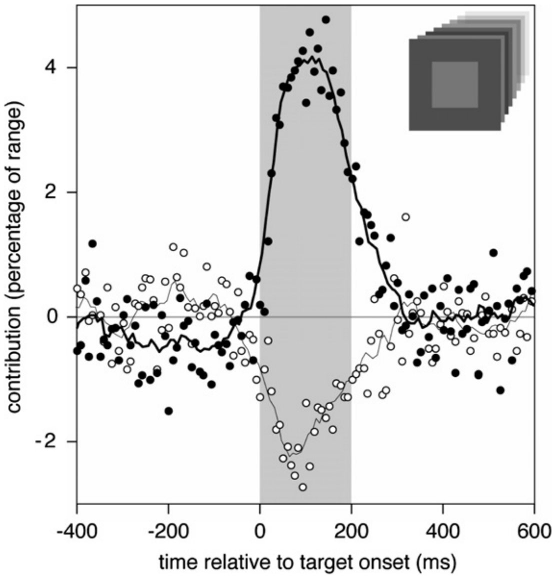

Fig. 8. Average data for the temporal experiment, in which a different colour is shown on every frame (120 Hz). A high value for the contribution is equivalent to a bright tile in Figs. 3 and 4. The grey area indicates the time during which the target is present (it is recognized by being brighter). Solid symbols are for the target and open symbols are for the background.

of the target itself (e.g., Fairchild & Reniff, 1995; Poot, Snippe, & van Hateren, 1997). Colours presented in the background (open symbols) have a smaller, opposite contribution, with more or less the same time course, although there may be a small delay in the contribution of the background (the thin curve in Fig. 8 is not only inverted and smaller, but it may also be shifted slightly to the left with respect to the thick curve).

The fact that we now find a much clearer spatial contrast (clearly negative values of open symbols *during* target presentation) than temporal contrast (only slightly negative values of solid symbols *before* target presentation) can probably be attributed to the difference between the stimuli used. In the main experiments there was a lot of spatial variability in colour, whereas the images were temporally quite stable. In this experiment there was very little spatial variability, but the colour was changing all the time. The finding that spatial contrast is stronger for spatially simpler stimuli is consistent with previous findings (Brenner et al., 2007).

8. General discussion

Our results clearly indicate that where people direct their gaze affects how they evaluate chromatic stimuli. We confirm Hansen and Gegenfurtner's (2005) finding that people are biased by where they are looking, showing that it is really the direction of gaze rather than the centre of the object that is critical, and that this is even so if the edges of the object in question are clearly visible. Our results also show that the classification image technique can be used to analyse chromatic interactions across surfaces' borders and across time.

We modelled the extent to which the contribution of the presence of a colour declines with retinal eccentricity (i.e., with distance from fixation) by a normal distribution with $\sigma \approx 1.5°$. This value suggests that the relevance of various parts of the retinal image simply depends on the number of cones or ganglion cells devoted to that region (see Mullen et al., 2005). It implies that the image on the fovea will generally dominate the percept. However, the perceived colour is not just a spatial average near fixation, because the colours of surfaces outside the target's borders have a smaller, opposite influence on the target's perceived colour.

From our main experiments we may get the impression that the influence of regions outside the target is very small (grey bars in Fig. 6). However, in the temporal experiment we saw that under different conditions the influence of such spatial contrast can be quite large (compare open symbols with solid ones in Fig. 8). Relying on differences in cone excitation ratios between light from different surfaces to determine the surfaces' colours could make colour vision less sensitive both to changes in illumination (Land, 1983; Foster et al., 1997) and to changes in the cones' sensitivity as they adapt to the various regions to which they are exposed. It is known that people do not rely exclusively on such differences (Foster, Amano and Nascimento, 2006), presumably because doing so would make the colour of the surrounding surfaces too important (Brenner & Cornelissen, 1991). We previously proposed that the extent to which the perceived colour is determined by the contrast at a surface's borders depends on various global scene statistics (Brenner & Cornelissen, 2002; Brenner et al., 2007). The present results imply that it also depends on where subjects look (Fig. 7).

If it is true that the extent to which one relies on contrast at the target's borders depends on the scene statistics, then we have identified a shortcoming of the classification image technique, because if so the relative contributions of different parts of the image will be influenced by the noise pattern that is introduced to estimate their contributions. This means that many of the parameters that we obtained from our model (Fig. 6b) are probably quite meaningless outside these experiments. Nevertheless if we are correct in assuming that only the strength of the contribution of contrast at the target's borders is influenced by the global scene statistics, then our conclusion regarding the spatial bias towards fixation (and even its spatial extent) remains valid. The fact that our prediction about the relationship between fixation and chromatic induction was confirmed (Fig. 7) supports this conclusion.

The fact that the perceived colour depends on where one is looking implies that the way we direct our gaze influences how we see objects. This is not only an issue for non-uniformly coloured objects, such as the apple and grapefruit of Fig. 1, but even for uniformly coloured surfaces, because the extent to which the colours of surrounding objects influence the perceived colour depends on where one directs ones eyes (as shown in Fig. 7). Beside the spatial effects, we also found that temporal contrast can influence the perceived colour. Fig. 8 shows that the perceived colour of a target is influenced by the retinal stimulation during the preceding 300 ms at least. In our experiments the changes in retinal stimulation occurred without any shifts in gaze, but in daily life such changes will mainly occur when we shift our gaze by making saccades. Thus both the direction of gaze and the eye movements that shift gaze are undoubtedly relevant for colour vision in our daily life.

Acknowledgment

This work was supported by Grant 051.02.080 of the Cognition Program of the Netherlands Organization for Scientific Research (NWO).

Appendix A. Using AICc to compare various models

We used

$$\mathrm{AIC}_C = n \cdot \ln\left(\frac{\mathrm{RSS}}{n}\right) + 2k + \frac{2k(k+1)}{n-k-1}$$

where n is the number of data points, k is one more than the number of free parameters, and RSS is the residual sums of squares, to calculate the evidence ratio

$$\mathrm{ER} = e^{0.5\cdot\Delta\mathrm{AIC}_C}$$

that indicates how much more likely one model is to be true than the other ($\Delta\mathrm{AIC}_c$ is the difference between the values of AIC_c for the two models). This method can be used to compare any two models with any numbers of free parameters (as long as the scatter of the data points with respect to the model predictions can be assumed to be normally distributed; Motulsky and Christopoulos, 2004).

Table 2 shows the relevant values for the two comparisons in which the number of free parameters was reduced from 7 to 5 by setting the peak at the target *centre* or at *fixation*. ***Bold italic*** values of ER indicate that the model with fewer parameters is more likely to be true (the presented values of ER are based on more precise values of RSS than those given in the tables).

Table 3 shows the relevant values for the two comparisons in which the number of free parameters was further reduced from 5 to 4 by assuming that only tiles *within* the target have an effect or by assuming that only colours presented *during* target presentation are relevant.

Appendix B. Why having more tiles gives less clear data

Let us start with a very simple example. We throw one die and make a certain decision every time the value is above average. For each throw there are six possible outcomes: 1, 2, 3, 4, 5 or 6. The average value is 3.5. If we throw a 4, 5 or 6 we will make a positive decision. Otherwise we will make a negative decision. After many trials the average value thrown when we made the positive decision will be 5. Now what if we throw two dice and base our decision on the sum of the two values. We are interested in the contribution of one of the two dice to the decision (i.e. its average value when a positive decision was made). Now there are 36 possibilities, as shown in Table 4.

If we consider the first digit as representing the contribution that we are interested in, then averaging across all combinations that sum to more than the average value of 7 (bold values) gives: $(2+3+3+4+4+4+5+5+5+5+6+6+6+\ 6+6)/15 = 4.67$. Considering that a value of 7 is neither larger nor smaller than the average, half the throws leading to a value of exactly 7 (italic) could also be considered to be above average. Correcting for this gives a value of 4.47 for the contribution in question. In either case the value is smaller than the value of 5 that we saw for one die.

Our task can be seen as throwing many dice (the number of tiles that are considered when making one's decision), each with 289 or 2601 possible values (the number of colours or tiles). If we ignore the fact that the decision in our task is presumably based on a weighted average and assume—as we did for the dice above—that subjects base their decision on a simple average, we can calculate the average expected value of a single tile when the average of all the ones considered is larger than the overall average (as we did in the example with the dice). For simplicity we will ignore the fact that in our stimuli the values were not completely independent. Fig. 9 shows the average value

Table 4
All combinations of values of two dice

	1	2	3	4	5	6
1	11	12	13	14	15	*16*
2	21	22	23	24	*25*	**26**
3	31	32	33	*34*	**35**	**36**
4	41	42	*43*	**44**	**45**	**46**
5	51	*52*	**53**	**54**	**55**	**56**
6	*61*	**62**	**63**	**64**	**65**	**66**

Table 2
Peak at target centre or fixation point

Session	n	$\mathrm{RSS}_{x,y}$	$\mathrm{RSS}_{\mathrm{centre}}$	$\mathrm{RSS}_{\mathrm{fixation}}$	$\mathrm{ER}_{\mathrm{centre}}$	$\mathrm{ER}_{\mathrm{fixation}}$
1	2601	0.4452	0.4513	0.4453	$>10^6$	***6.3***
2[*]	2601	0.4923	0.4945	0.4925	38.7	***5.4***
2	5202	0.9769	0.9812	0. 9778	$>10^4$	1.9
3	5202	1.1790	1.1804	1.1794	3.4	***2.7***
4	289	0.0425	0.0484	0.0426	$>10^6$	***5.3***
5	578	0.1048	0.1153	0.1057	$>10^6$	1.4
6	578	0.0541	0. 0722	0.0569	$>10^6$	$>10^5$

Table 3
Only considering the target itself

Session	n	$\mathrm{RSS}_{\mathrm{fixation}}$	$\mathrm{RSS}_{\mathrm{within}}$	$\mathrm{RSS}_{\mathrm{during}}$	$\mathrm{ER}_{\mathrm{within}}$	$\mathrm{ER}_{\mathrm{during}}$
1	2601	0.4453	0.4454		***2.0***	
2[*]	2601	0.4925	0.4929		1.3	
2	5202	0.9778	0.9785	0.9788	1.9	4.9
3	5202	1.1794	1.1809	1.1795	8.3	***2.4***
4	289	0.0426	0.0427		***1.9***	
5	578	0.1057	0.1061	0.1068	1.2	8.0
6	578	0.0569	0.0570	0.0597	***2.2***	$>10^5$

[*] Condition with the same random pattern of colours during as before and after target presentation.

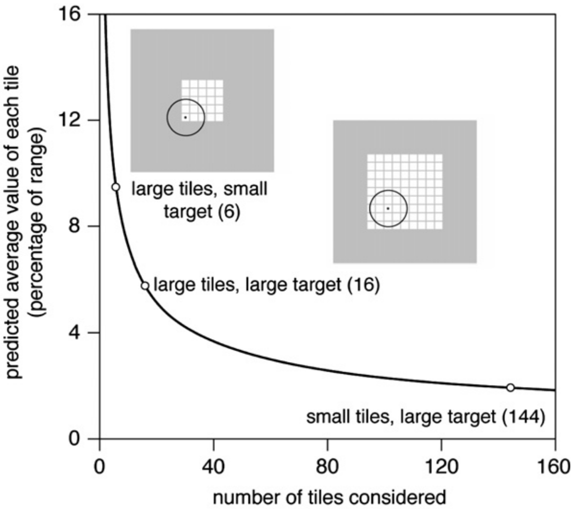

Fig. 9. How the contribution of each tile is expected to contribute to the percept, as a function of the number of tiles that are considered. The points indicate relative values for numbers of tiles that could arise from a fixed spatial extent centred on the fixation point (specified by the numbers within brackets).

of a single tile when the average of the indicated number of tiles is larger than the overall average. The circles indicate values of the number of tiles considered that are consistent with the spatial relationships between our stimuli and would result in approximately the pattern of results that we see in Fig. 6b (black bars).

References

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research, 45*, 607–615.

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research, 37*, 2849–2869.

Brenner, E., & Cornelissen, F. W. (1991). Spatial interactions in color vision depend on distances between boundaries. *Naturwissenschaften, 78*, 70–73.

Brenner, E., & Cornelissen, F. W. (2002). The influence of chromatic and achromatic variability on chromatic induction and perceived colour. *Perception, 31*, 225–232.

Brenner, E., Granzier, J. J., & Smeets, J. B. J. (2007). Combining local and global contributions to perceived colour: An analysis of the variability in symmetric and asymmetric colour matching. *Vision Research, 47*, 114–125.

Brenner, E., Ruiz, J. S., Herraiz, E. M., Cornelissen, F. W., & Smeets, J. B. J. (2003). Chromatic induction and the layout of colours within a complex scene. *Vision Research, 43*, 1413–1421.

Cornelissen, F. W., & Brenner, E. (1991). On the role and nature of adaptation in chromatic induction. In B. Blum (Ed.), *Channels in the visual nervous system: neurophysiology, psychophysics and models* (pp. 109–123). London: Freund publishing house ltd.

Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: an analysis of viewing strategies. *Vision Research, 35*, 2431–2448.

Dakin, S. C., & Bex, P. J. (2003). Natural image statistics mediate brightness 'filling in'. *Proceedings of Biological sciences, 270*, 2341–2348.

Fairchild, M. D., & Reniff, L. (1995). Time course of chromatic adaptation for color-appearance judgments. *Journal of the Optical Society of America A, 12*, 824–833.

Foster, D. H., Amano, K., & Nascimento, S. M. (2006). Color constancy in natural scenes explained by global image statistics. *Visual Neuroscience, 23*, 341–349.

Foster, D. H., Nascimento, S. M. C., Craven, B. J., Linnell, K. J., Cornelissen, F. W., & Brenner, E. (1997). Four issues concerning colour constancy and relational colour constancy. *Vision Research, 37*, 1341–1345.

Granzier, J. J., Brenner, E., Cornelissen, F. W., & Smeets, J. B. J. (2005). Luminance-color correlation is not used to estimate the color of the illumination. *Journal of Vision, 5*, 20–27.

Hansen, T., & Gegenfurtner, K. R. (2005). Classification images for chromatic signal detection. *Journal of the Optical Society of America. A, Optics, image science, and vision, 22*, 2081–2089.

Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America, 53*, 741–744.

Von Kries, J. (1905). Die Gesichtsempfindungen. In W. Nagel (Ed.). *Handbuch der Physiologie des Menschen* (Vol. 3, pp. 109–281). Vieweg und Sohn: Physiologie der Sinne. Braunschweig.

Land, E. H. (1983). Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America, 80*, 5163–5169.

Motulsky, H., & Christopoulos, A. (2004). *Fitting models to biological data using linear and nonlinear regression: A practical guide to curve fitting*. New York, USA: Oxford University Press.

Mullen, K. T., Sakurai, M., & Chu, W. (2005). Does L/M cone opponency disappear in human periphery? *Perception, 34*, 951–959.

Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience, 5*, 812–816.

Poot, L., Snippe, H. P., & van Hateren, J. H. (1997). Dynamics of adaptation at high luminances: adaptation is faster after luminance decrements than after luminance increments. *Journal of the Optical Society of America A, 14*, 2499–2508.

Shevell, S. K., & Wei, J. (1998). Chromatic induction: border contrast or adaptation to surrounding light? *Vision Research, 38*, 1561–1566.

Wachtler, T., Albright, T. D., & Sejnowski, T. J. (2001). Non-local interactions in color perception: Nonlinear processing of chromatic signals from remote inducers. *Vision Research, 41*, 1535–1546.

Yund, E. W., & Armington, J. C. (1975). Color and brightness contrast effects as a function of spatial variables. *Vision Research, 15*, 917–929.