Spatial selectivity of the watercolor effect

Frédéric Devinck, ¹, ², ³ Peggy Gerardin, ¹, ², ³ Michel Dojat, ¹, ³ and Kenneth Knoblauch ², ³, *

¹Université Rennes, 2 CRPCC EA 1285, 35000 Rennes, France
²INSERM U846 Stem-Cell and Brain Research Institute Department of Integrative Neurosciences, 69500 Bron, France
³Université Lyon 1, F-69373 Lyon Cedex 08, France

*Corresponding author: ken.knoblauch@inserm.fr

Received September 19, 2013; accepted October 23, 2013; posted November 1, 2013 (Doc. ID 197898); published December 13, 2013

The spatial selectivity of the watercolor effect (WCE) was assessed by measuring its strength as a function of the luminance contrast of its inducing contours for different spatial configurations, using a maximum likelihood scaling procedure. The approach has previously been demonstrated to provide an efficient method for investigating the WCE as well as other perceptual dimensions. We show that the strength is narrowly tuned to the width of the contour, that it is optimal when its pair of inducing contours are of equal width, and that the strength can be increased by varying the overall size of the stimulus when the width of the inducing contour is not optimal. The results support a neural substrate that has characteristics not unlike double-opponent, color-luminance cells observed in cortical area V1. © 2013 Optical Society of America


http://dx.doi.org/10.1364/JOSAA.31.0000A1

1. INTRODUCTION

The watercolor effect (WCE) is a long range color filling-in phenomenon that appears when a pair of chromatic, wavy contours bound an achromatic surface area [1, 2]. The illusion has generally been demonstrated using a background of higher luminance than the two contours. An assimilation effect is produced when a darker outer contour is flanked by a lighter inner contour [3]. Under these conditions, the lighter color will spread outward over the entire enclosed area [Fig. 1(a)].

The strength of the WCE depends on several stimulus parameters. For example, increasing the luminance contrast between the inner and outer contours leads to a stronger saturation of the fill-in color [3–5]. Also, larger chromatic contrasts between the contour pairs leads to a stronger effect [2, 6]. Although the WCE has been reported for straight contours [2, 7], its strength has been demonstrated to depend on the frequency and amplitude of the contour undulation [8]. In general, these studies indicate that the WCE depends on both local spatial and chromatic properties of the contour of the stimulus that might be accounted for by the selectivity of early visual mechanisms that in themselves would not be expected to be capable of generating a filling-in perception over a large region. For this reason, models that have been proposed to account for the WCE typically invoke processing at several levels in the visual system [2, 9, 10].

In this article, we examine two other stimulus characteristics of the WCE that have received little attention in previous work: the width of the inducing contour and the overall size of the stimulus. We estimate perceptual scales based on paired-comparison judgments and a maximum likelihood estimation (MLE) procedure [11–13] that we have previously shown to provide reliable estimates of the strength of the illusion within the framework of signal detection theory [5].

2. METHODS

A. Observers

A total of six observers were tested in the experiments presented here, although only one (an author) completed the full set of conditions. Each observer returned for several sessions to complete the conditions in which he or she participated. All observers except the author remained naïve over the course of the experiments. All observers had normal color vision as assessed by a Farnsworth Panel D15 and normal or corrected to normal visual acuity. If required, observers wore their corrections during all sessions. Ages ranged from 26 to 39 [mean age (SD) = 28.8 (5.0) years], and both genders were equally represented in the sample. All experiments were performed in accordance with the principles of the Declaration of Helsinki for the protection of human subjects.

B. Apparatus

Stimuli were presented on a NEC MultiSync FP2141sb color CRT monitor driven by a Cambridge Research ViSaGe graphic board (14 bits per gun) (Cambridge Research Systems, Rochester, United Kingdom). Experiments were performed in a dark room. Experimental software was written to generate all stimuli, control stimulus presentation, and collect responses in MATLAB 7.9 (http://www.mathworks.com) using the CRS Toolbox extensions. The monitor was calibrated using an OptiCal photometer with the calibration routines of Cambridge Research Systems. Observer position was stabilized by a chin.
rest so that the screen was viewed binocularly at a distance of 80 cm.

C. Stimuli
All stimuli were displayed on a white background (134 cd/m², CIE xy = 0.29, 0.32). The outer contour was purple (CIE xy = 0.31, 0.11) at 24.1 cd/m², and the inner orange (CIE xy = 0.4, 0.43) with luminance varying between 70.2 and 109.6 cd/m². The stimuli were also specified in the DKL color space \([14]\) with purple and orange contours at azimuths of 320 and 45 deg, respectively. The luminance elevations of the orange contour in DKL space were evenly spaced between 0 and 0.72.

The stimuli were constructed as Fourier descriptors \([15]\) [Figs. 1(b) and 1(c)]. Each stimulus was defined by a circle whose radius, \(r\), was modulated sinusoidally as a function of angle, according to the equation

\[ R(\theta) = r + m \sin(2\pi f \theta), \]

where \(R\) is the stimulus radius at angle \(\theta\), \(r = 1.6\) deg for the first experiment (giving a 3.2 deg average diameter), \(m\) is the modulation, and \(f\) is the frequency in cycles per revolution (cpr). For the experiments described here, the frequency was fixed at a value of \(f = 10\) cpr and the amplitudes of the two ribbons at \(m = 0.11\) and \(m + w\), where \(w\) is the width. Control stimuli were identical, except that the contour ribbons were interlaced and generated little, if any, filling-in [Fig. 1(c)].

D. Procedure
Observers adapted to a blank screen set to the background luminance and chromaticity for 2 min before starting the experiment. A trial consisted of the presentation of a triad, \((a, b, c)\), of WCE patterns with three luminance elevations chosen from a series of 10, with \(a < b < c\) as shown schematically in Fig. 2. Stimulus \(b\) was always the upper stimulus in the middle, and stimuli \(a\) and \(c\) were randomly positioned on the left or right. A fixation cross appeared in the center of the screen. Each WCE pattern was offset 2.7 deg vertically from the fixation cross. The two bottom patterns were offset 2.6 deg laterally (see [5] for more details).

A session consisted of the random presentation of the 10!/(3!7!) = 120 unique triads from the series of 10 elevations of the orange contour for a fixed contour width. At the beginning of each trial, the observer fixated the cross.

With presentation of a triad, (s)he was instructed to fixate each pattern until (s)he could choose which of the two bottom patterns (left or right) was most similar to the upper pattern with respect to the color of its interior region. The observer’s response initiated the next trial. A session typically required about 5 min so that the average duration of a trial was about 2.5 s.

A free viewing procedure was employed to ensure that the observer’s judgments were based on foveal views of the stimuli and not comparisons across different peripheral, retinal regions. A temporal presentation was ruled out to avoid the possibility of memory effects. Free viewing has been used to study the WCE in numerous previous studies \([1–3,6,7,16,17]\). Two previous studies have used durations, respectively, of 1 s \([10]\), not too much shorter than our estimated average duration, and 150 ms to study an achromatic variant of the WCE \([4]\).

To control for responding simply on the basis of the contour instead of the filled-in appearance, sessions were also run with control stimuli [Fig. 1(c)]. Each observer performed 5 sessions each with the test and control stimuli.

For the first set of experiments, the contour pairs that defined the stimuli were tested at total widths of 6, 11, 15, 19 and 24 min with equal widths for the purple and orange ribbons. In the second set of experiments, the total width was fixed at 15 min and three ratios of the two ribbons were tested: 2:1, 1:1, and 1:2. In a third experiment, additional conditions in which both the radius, \(r\), of the modulated circle and the width of the contours at 1:1 ratio were varied.

E. Analyses
A difference scale was estimated from each session using functions from the MLDS package \([12]\) in the open source software R \([18]\), and the scales from the individual sessions were averaged to obtain means and standard errors for the estimated scale values.

The analysis is based on a signal detection model in which the observer’s judgments depend on a decision variable based on comparing perceptual intervals between pairs of stimuli, e.g., as above, \((a, b)\) and \((b, c)\) \([11–13]\). It is assumed that on each trial the decision variable is perturbed by Gaussian noise with variance \(\sigma^2\). Intuitively, for stimulus \(b\), such that the observer is equally likely to choose either stimulus interval, the perceptual intervals between the two pairs are equal.
Scale values and the variance are estimated by MLE so as to predict best the set of the observer’s responses over the course of an experiment. The estimated scale has the property that equal scale differences correspond to equal perceived differences. The scale is unique, however, only up to a linear transformation \([11, 13]\), i.e., adding and/or multiplying all scale values by a constant does not affect the predictions. Thus, we fixed the scales to be zero at the lowest luminance tested and to have \(\sigma^2 = 1\) at each stimulus level tested. Parameterized in this fashion, the estimated scale values of an MLDS experiment are in the same units as the measure \(d’\) from signal detection theory, i.e., in units of the standard deviation of the internal noise (see \([5]\)). We used this parameterization in the graphs throughout the article.

3. RESULTS

A. Contour Width and WCE Strength

Figure 3 shows mean estimated scales of the variation of the strength of the WCE with luminance of the orange contour for test and control conditions (rows) and for the four observers who participated in this experiment (columns). Each symbol type corresponds to a different contour width of the stimulus. When the test stimulus is used, the fill-in response grows with the luminance of the interior contour, as previously demonstrated \([5]\) and tends to asymptote at high luminances. For all four observers, the strongest response is obtained for an intermediate contour width, near 15 min and decreases for both smaller and larger widths. Compared to the test stimuli, the effect is attenuated or absent when the control stimuli are used. This is important in order to be able to argue that the observers responded to the fill-in color on each trial and not the contours themselves. For observer JA there is, in fact, an indication of a small effect for the control stimuli. This is most evident in the data from the two smallest widths (6 and 11 min) for which the WCE would be expected to be smallest. Under these conditions, it is possible that the observer responded on the basis of the contour colors, given that so little filling-in was produced by the contours themselves.

For the purpose of characterizing the width-tuning that is demonstrated in Fig. 3, we found that a Legge–Foley function \([19]\),

\[
R(L) = R_m \frac{L^n}{L^n + \zeta^n},
\]

where \(R\) is the fill-in response, \(R_m\) is the asymptotic maximum response, \(L\) is the luminance elevation, \(\zeta\) is the half-saturation luminance, and \(n\) is an exponent, provided a reasonable description of the increase in response of the WCE with luminance. In fitting the responses by least squares, we let the parameter \(R_m\) vary as a function of contour width for each observer. By evaluating nested models with a likelihood ratio test, we found that the best fit, fixed value of the exponent, \(n = 1.4\), sufficed to describe the data across widths and observers (FD: \(F_{4,43} = 0.14, p = 0.72\); SF: \(F_{4,43} = 0.24, p = 0.63\); BA: \(F_{4,43} = 0.56, p = 0.46\); JA: \(F_{4,43} = 2e - 3, p = 0.96\)), but the value of \(\zeta\) was required to vary across observers (FD: \(\zeta = 0.40, F(4,44) = 7.80, p < 0.01\); SF: \(\zeta = 0.58, F_{4,44} = 11.18, p < 0.01\); BA: \(\zeta = 0.68, F_{4,44} = 19.1, p < 0.01\); JA: \(\zeta = 0.26, F_{4,44} = 49.3, p < 0.01\)). The fitted curves are shown as dashed lines in Fig. 4(a) for each width and observer. The mean response functions for the four observers with standard errors of the mean (SEM) are shown in Fig. 4(b) with the fitted curves of Eq. 2 as solid lines. The symbols correspond to those used in Fig. 3 for the widths. For our purposes, this function adequately characterizes the increase in strength of the phenomenon over the range of luminances tested, even though there is a tendency in several instances to overestimate the strength at the highest luminance tested.

To characterize the dependence of the WCE strength on contour width, we used the fitted curves to predict the fill-in response strength at a criterion luminance elevation of 0.72 and plotted these values as a function of contour width [Fig. 5(a)]. These show the WCE to be quite narrowly tuned as a function of contour width. The bold solid curve is obtained by a local regression used to estimate a smooth average. Over the range of widths that we measured, the full bandwidth at half-height of the average curve is about 8 min.

![Fig. 3](image_url)  
**Fig. 3.** Mean difference scales in units of \(d’\) as a function of the luminance elevation of the interior, orange contour, parameterized by the width of the contour pair (legend on right). Each column shows the results from one observer, with initials indicated in the top strip. The top row shows the results for the test stimulus and the bottom for the control.
The inducing contour of the WCE can be considered as 1 cycle of a chromatic-luminance grating. In Fig. 5(b), we replot the values from Fig. 5(a) as a function of the spatial frequency in cycles/degree (c/deg) of this grating. The results show the average curve to be band pass with a peak response at 4 c/deg and a full bandwidth at half-height over the range that we measured of 1.7 octaves.

B. Contour Ratio and WCE Strength

Figure 6 shows the average response functions and standard errors for three ratios of the widths of the WCE inducing contours for the three observers who participated in this experiment. The three observers showed the same pattern of response to the three width ratios and, as in the preceding experiment, an attenuated response in the control conditions. The inner to outer ratios are indicated as 2:1, 1:1 and 1:2. The total width was fixed at 15 min, the optimal width found in the previous experiment. The results indicate that the WCE is strongest when the outer and inner contours are of equal width. The decline with contour ratio is asymmetric. When the inner contour is half the width of the outer contour the WCE is attenuated by almost a factor of 4 on average, whereas when the contour width ratio is inversed the WCE is only reduced by about 65%.

C. Interactions of Size and Width

We varied the size of the circle whose radius was modulated to form the WCE stimulus by varying its diameter. All three observers who performed this experiment showed the same pattern of responses. Fig. 7 shows the average results of varying the size of the stimulus for three contour widths. The white points indicate the 3.2 deg diameter stimulus and correspond to the mean of the conditions shown in Fig. 4(b) for the three observers, FD, BA, and JA. The previously shown width selectivity for this size stimulus is indicated by the reduction in height of the response curves for 6 and 24 min. At the smaller width, however, when the diameter is reduced by 70% to

Fig. 4. (a) Mean difference scales for the five widths tested for each observer with the fitted curves from Eq. 2 as dashed curves and (b) mean difference scales across observers for each contour width. Error bars of ±1 SEM are used here and in subsequent figures to minimize overlap between points from different conditions. The solid curves are the best fits of Eq. 2.

Fig. 5. (a) Points and dashed lines indicate the estimated strength of the WCE at a criterion luminance elevation as a function of contour width for individual observers. The solid curve is an average obtained from a local regression algorithm and (b) the same data as in panel a but replotted as a function of the spatial frequency of the double contour inducing the WCE (1/width). The meaning of the points and curves is the same as in panel a.

Fig. 6. Average response functions for three observers for each of three ratios of the purple and orange ribbons of the WCE inducing contour. The error bars show ±1 SEM. The line color and type corresponding to each ratio is indicated in the legend above the graph. The numbers in the legend indicate the inner to outer width ratio.
2.3 deg, the WCE strength increases as shown by the gray points. In contrast, with the thicker 24 min contour, increasing the stimulus diameter by 70% to 4.5 deg results in a stronger WCE.

The observed interaction of contour width and stimulus size cannot be considered to be a simple scaling, as for all three sizes, the largest effect is still found for the condition with a 15 min contour width. At this width, there appears to be no effect of increasing the stimulus size, although there is a slight reduction of the WCE for the 2.3 deg diameter stimulus.

4. DISCUSSION
In this study, we quantified three aspects of the WCE using a maximum likelihood, psychophysical scaling paradigm: (1) its selectivity for contour width, (2) its sensitivity to the balance of the widths of the inner and outer contours, and (3) an interaction between overall size of the stimulus and contour width.

We found that the WCE was sharply tuned to a peak width of about 15 min. This is slightly but significantly larger than the value of 12 min reported previously [2], using hand-drawn stimuli and magnitude estimation, and the method of limits. The stimulus configuration was quite different, as well, in this previous study and may have played a role in the slight differences observed. Our results indicate that the underlying mechanism responding to the contour is band pass and tuned to spatial frequencies in the range of 3–5 c/d, near the peak of human luminance contrast sensitivity in the fovea [20].

The phenomenon was also strongest when the widths of the pair of inducing contours were equal. This suggests that the underlying mechanism responds best to spatially balanced contours. The luminance response contours in Fig. 6 indicate that some spatial integration of the luminance and chromatic contrasts to compensate for an imbalance in the contour widths is possible. Thus, an increase of the luminance of the interior orange contour can partially offset a reduction in width of the outer purple contour. Less of a compensation is possible, however, when the inner contour is thinner. Since the purple contour is already fairly dark, there is probably little room for adjusting it to increase the illusion strength. While the WCE occurs for equiluminant contours, it has been reported to be stronger when there is, additionally, some luminance contrast present [3]. It is possible that modifying the ratio of contour widths would result in a systematic variation in the optimal ratio of luminances to produce the WCE, but this would require a more systematic study of the interaction of these variables.

As the first experiment shows, when the contour is not of optimal width, the WCE effect is attenuated. Thus, it is interesting that, when the overall stimulus size is increased or decreased, the WCE is boosted by a respective increase or decrease in the contour width. Nevertheless, the optimal width seems relatively immune to the overall size change over the range studied here, at least, for central viewing. The results would, thus, exclude a single, size invariant mechanism as mediating the WCE.

From its earliest description, the WCE was described as both a coloration (or filling-in) and a figure/ground phenomenon, since the areas over which the filling-in occurs appear as foreground surfaces [2]. By comparing surfaces with real color matched to surfaces with the WCE fill-in color, it was shown that the figure/ground salience was not completely abolished [10], thus demonstrating a dissociation between the two effects. It was suggested that the filling-in is associated with the representation of surface information, while the figure/ground phenomenon is associated with mechanisms related to border ownership. These ideas are implemented in Grossberg’s model [9] as separate pathways for extracting boundary and surface characteristics from images and inspired by neurophysiological evidence from neurons in cortical areas V1 to V4.

The fact that the WCE is strongest for contours that contain both a luminance and a chromatic component [3] suggests that either it depends on an interaction of luminance and chromatic sensitive cells or, more simply, cells tuned to be responsive along both dimensions. Shapley and Hawken [21] describe a large population of double opponent, color-luminance cells in cortical area V1 of the macaque that can be considered as reasonable candidates for the neural substrate for the latter hypothesis. They report a full-width, half-height bandwidth of these cells of 2.05 ± 0.7 octaves. This is wider than the value of 1.7 octaves for our average curve but within their range of variability. Their average tuning curve peaks at 2.6 c/deg, significantly lower than the 4.0 c/deg that we obtain for the WCE. Their measurements, however, were obtained in the vicinity of 5 deg parafoveal while our conditions are for foveal viewing. Thus, given eccentricity as well as possible species differences, it is not unreasonable to hypothesize...
that a similar class of units may underlie the contour processing of the WCE.

ACKNOWLEDGMENTS

This work was supported in part by grant ANR 11 JSH2 0021 from the Agence Nationale de la Recherche to Frédéric Devinck. We gratefully thank John S. Werner for comments.

REFERENCES