Effect of fixation positions on perception of lightness

Matteo Toscani, Matteo Valsecchi, Karl R. Gegenfurtner
Abteilung Allgemeine Psychologie, Justus-Liebig-Universität, Giessen, Germany

ABSTRACT

Visual acuity, luminance sensitivity, contrast sensitivity, and color sensitivity are maximal in the fovea and decrease with retinal eccentricity. Therefore every scene is perceived by integrating the small, high resolution samples collected by moving the eyes around. Moreover, when viewing ambiguous figures the fixated position influences the dominance of the possible percepts. Therefore fixations could serve as a selection mechanism whose function is not confined to finely resolve the selected detail of the scene. Here this hypothesis is tested in the lightness perception domain. In a first series of experiments we demonstrated that when observers matched the color of natural objects they based their lightness judgments on objects' brightest parts. During this task the observers tended to fixate points with above average luminance, suggesting a relationship between perception and fixations that we causally proved using a gaze contingent display in a subsequent experiment. Simulations with rendered physical lighting show that higher values in an object's luminance distribution are particularly informative about reflectance. In a second series of experiments we considered a high level strategy that the visual system uses to segment the visual scene in a layered representation. We demonstrated that eye movement sampling mediates between the layer segregation and its effects on lightness perception. Together these studies show that eye fixations are partially responsible for the selection of information from a scene that allows the visual system to estimate the reflectance of a surface.

Keywords: lightness perception, eye movements, visual saliency, transparency perception.

1. INTRODUCTION

Only a minor portion of the incoming information to the visual system is processed in high detail, thus, in order to appropriately select behaviorally relevant information, selection mechanisms are required. The first gating mechanism takes place in the retina, which sends high resolution visual signals to the brain only from its most central region. Although a scene in the visual field is perceived as colorful, detailed, stable and three-dimensional, this percept is mostly constructed by the visual system from the signals sampled by our retinas. Each retina sends relatively elaborate visual signals to the brain from only about two degrees of its central region, the fovea. The spatial sampling of the retinal image declines with eccentricity, resulting in changes in visual acuity, contrast sensitivity and color sensitivity [12,3] . Peripheral vision is not only characterized by poor resolution, but also the appearance of basic visual features, such as spatial frequency, luminance or chromatic saturation, is distorted in the periphery of the visual field. The visual system contains sensory channels which are sensitive to different ranges of spatial frequencies[4]. Spatial frequencies are overestimated at 8° of eccentricity[5] and this is thought to be due to a failure to adapt the labeling of spatial-frequency channels in the periphery, where spatial pooling with retinal eccentricity causes the peak channel sensitivities to shift towards lower spatial frequencies[6]. Perceived luminance is distorted in the peripheral view, specifically, the perceived luminance of supra-threshold (clearly visible) flashes decreases with eccentricity[7,8]. In order to achieve the impression of the detailed uniform and stable scene we perceive, our visual system has to stitch together its representation of the world from many small samples. The visual system is actively scanning the world through eye movements, therefore some parts of the scene are selected for a finer analysis and some are ignored. It is noticeable that the effect of the scanning pattern goes beyond differential sampling in terms of resolution or biases in basic visual dimensions: Einhauser et al.[9] have found a bidirectional coupling between the eye position and the perceptual switching of the Necker cube. Similarly, Georgiades and Harris[10] found that the position of the fixation point influences the dominant percept of the wife/mother-in-law ambiguous figure. Beyond ambiguous figures, there is a little evidence of a relationship between the perception of surface reflective properties and the fixation behavior. Namely, when observers were asked to adjust a test patch to match its hue and saturation to those of a standard patch, the fixation pattern was fundamentally different from the one produced when they were asked to adjust the test patch to look as if it had been "cut from the same piece of paper" as the standard[11]. Using a classification images paradigm, Hansen and Gegenfurtner [12] showed that the region close to the fixation point was particularly relevant in chromatic judgments. Specifically, they asked their observers to detect colored targets embedded in noise. They found that the observers clearly did not rely on the whole stimulus, but mostly on a small area around the fixation point. In their study the fixation point was at the center of the target, but it has been argued that the center itself already carries more weight, besides the fact that it is the fixation position. Brenner et al.[13] explicitly tested whether the chromatic information presented in the fovea plays a major role in judging the color of a surface. Using a similar image classification technique, they asked their observers to judge whether noisy textures were either red or green. Observers tended to report the colors of the areas within the texture that were close to the fixation point as a global percept of the texture. These results suggest that the foveal information plays a major role in judging the global chromatic properties of a texture, at least where a selection process is explicitly required (observers had to summarize the color noise in the texture in to “green” or “red”). Together these findings suggest that eye movements could serve as selection mechanism, which focuses the system towards particular aspects of a visual scene. There is a debate on which selection strategy is driving the fixation behavior. Yarbus[14] showed that the fixation patterns within the same scene depend on the observer's task, implying a cognitive drive (top down) to fixate at some regions of the scene rather than others. According to a different approach the image properties themselves are prominent in driving the scanning strategies (bottom up). This approach led to the development of models that estimate the fixation distributions (*saliency maps*) on a scene given its image statistics [15,16]. One reason why the saliency maps approach caught much attention was its close relationship to the descriptions of the early visual system available at the time[17]. Nowadays the idea of parallel and independent visual pathways for different visual attributes such as color, form, or motion is not as dominant as it was in previous decades[17]. Furthermore, there is good evidence that the role of a bottom up saliency might be relatively small in explaining fixation patterns, while other more cognitive aspects have been found to be important, like task dependent plans, object recognition, and value[17]. A series of studies[18–20] have shown that during everyday activities observers almost exclusively fixated on task-relevant objects. For instance, when making tea observers fixate on the objects used for the task, but also task-relevant empty areas such as the spot on the table where they wanted to place the cup, which is obviously not predictable in a bottom up fashion. These instances suggest that eye movements are fulfilling an economy principle, despite that fact that there is seldom direct reward or punishment for eye movements and that they cost little metabolic energy[17]. Given that the fixation behavior could change the global interpretation of a scene, eye movements take part in the constructive processes of vision, going beyond a sampling economy principle. On this line the fixation behavior has been proposed to be an accessing method to visual information in an online refreshing fashion, rather than a parsimonious sampling of the visual scene used to rely on visual memory for further processing[21]. For instance, when observers had to copy an arrangement of blocks they repeatedly shifted their gaze to the blocks rather than storing the arrangement in the visual memory[22,23]. This view is consistent with the idea that eye movements are playing an active role in perceiving the surface reflectance. In this view the representations of the fixated areas are weighted more than the ones of the rest of the surface, consistent with the *coherence theory* proposed by Rensink[21]. This theory states that although the whole scene appears detailed, the unattended structures are volatile, and focused attention is needed to reach a level of stability, which is sufficient to guide the behavior in a useful way.

To estimate the reflectance of surface is not a trivial task since the visual system only access to amount of light coming to the eye from the surface and this depends on the material, but also on the illumination, the geometry of the surface and the transmitting medium between the reflecting surface and the eye. When the light reaches the eye, these factors are confounded in the reflected light. Despite this ambiguity, humans are pretty consistent in their judgments on surface color or lightness, a behavior called *color (or lightness) constancy*. The only property that depends on the material itself is the proportion of light that this material reflects, making this coefficient a good property for the visual system to estimate, in order to achieve lightness constancy. This proportion of reflected light (reflectance) depends on its wavelength and on the interaction between the geometry of the surface and the geometrical structure of the illuminant. This relationship is described by a function called *BRDF* (bidirectional reflectance distribution function), which is usually non-trivial and has to be determined empirically, albeit several models have been developed to describe it in a relatively general way. If we consider the simple case of a perfectly matte surface, when the light beam hits a pigment particle it penetrates through the particle; some wavelengths may be absorbed and the remaining light returns into the surrounding medium where it is scattered around all the directions. A material has different reflectance coefficients for different wavelengths. The ratios between these coefficients define the chromatic properties of the material, whereas an average of those coefficients determines its *albedo*, a general measure of the reflectance of a surface. In perceptual sciences, lightness has been defined as the apparent reflectance, as opposite to brightness, which is defined as the apparent luminance[24]. Several factors have been shown to play a role in lightness constancy. Lateral inhibition between retinal neurons filters out the shallow intensity gradients, which are mostly due to illumination effects[25–27]. On this line, Land and McCann[26] proposed a model (*Retinex theory*) which applies a derivative operator to the image to remove the variations due to the illumination, than it reintegrates the edge information over space to reconstruct an image of the reflectance of the scene.

The low level approach is appealing because it relies on a set of established physiological mechanisms and it provides a formal explanation for several well-known visual illusions, but the limits of this approach in explaining other visual effects show that there is more to lightness perception. For instance, Koffka offered an example of how the simultaneous contrast effect can be lowered by contextual cues[28]. More recently Knill and Kersten[29] provided a remarkable demonstration on how an illusion, which is traditionally explained by low level mechanisms, is influenced just by changing the interpretation of the scene. Traditionally lightness constancy has been investigated using flat surfaces under a uniform illuminant or a gradient, as in the O'Brien-Cornsweet display. Nevertheless, the reflected luminance from a surface depends on the relationship between the surface geometry and the geometrical structure of the illumination. Lightness is defined as the perceived reflectance, therefore, in order to perceive lightness, the visual system has to extract one value from the luminance distribution that represents the surface reflectance. When the causes of the variations in the luminance reflected by the surfaces are known, they could be used to compute the surfaces' reflectance from the luminance distributions in a so called *reverse optics* way. This theory implies that visual system produces separate representations of the illuminant, the surface geometry and the surface albedo. However, there is evidence that lightness judgments are directly based on brightness, at least to some degree[30,31]. Contrary to the *reverse optics* approach, one line of research proposed that the visual system estimates the surfaces properties on a luminance histogram basis. This implies that the spatial properties of the shading are largely irrelevant. For instance Motoyoshi et al[32]. found that the skew of the luminance histograms of stucco-like surfaces directly correlates with surface gloss and inversely correlates with surface albedo. The variability in the luminance reaching the eye does not only come from the interaction between the geometry of a surface and the structure of the illumination field, but it also depends on the characteristics of the optical medium. This medium is not necessarily acting uniformly on the whole visual field, but it may increase the variability in the reflected light because of its inhomogeneities. The reflective properties of the surface and the properties of the optical medium are collapsed into a single retinal projection yielding to an ambiguity where infinite solutions are possible. Luminance discontinuities may be the result of different causes, like boundaries of occluding objects or media, or they may be caused by illumination changes such as those along the shadows. Considering simple features of those discontinuities it is possible to recover the most likely cause. For instance, a shadow preserves the contrast polarity along its border, because it is always darker within the shadow. On the other hand, occluding edges often generate contrast reversals because partially occluded surfaces may contain regions that are both lighter and darker than the occluding object. Transparent surfaces and changes in the optical medium generate a constant polarity edge in a similar way to the shadows. Specifically, it has been proposed[33–35] that geometric and luminance relationships within contour junctions induce illusory transparency and lightness perception. Namely, when a border is uniform in contrast polarity and the magnitude of the contrast along the border is discontinuous, the regions whose contrast relative to the background is higher are perceived as belonging to the object surface in a plain view. The lower contrast regions are interpreted as seen through a contrast-reducing medium where its opacity is proportional to the contrast with the highest border contrast region. Using this simple heuristic, the visual system could exploit common layout features to infer the spatial relationships between the objects. Namely, a uniform contrast polarity along an edge is likely to be caused by an occluding object being either darker or lighter. Similarly, a shadow is always darker than the material on which it is projected. Considering these processes allows the observer to distinguish between the surface reflectance and other factors. Specifically, the heuristic suggests which parts of the surface are more informative about its albedo.

Here we test the hypothesis of a link between the local information sampled from individual fixations and the apparent lightness of an object. In a first series of experiments with real surfaces uniform in albedo we first show that observers tend to take in to account heavily the brighter parts of objects when they are asked to match the color or lightness of these objects. We show that observers tend to fixate on the brighter parts of the objects while they make their match. Then we show that this link between fixations and lightness perception is causal. We also show that eye fixations and attention both contribute to this effect, and finally we show that the brighter parts of objects are particularly diagnostic of the object's reflectance. Since people tend to fixate the parts of a surface that are most informative about its albedo, in a second series of studies we aimed to test the flexibility of the fixation strategy and how this is actually driven by heuristics that the visual system is using to select relevant information to estimate the surface reflectances. Therefore in a second series of experiments we presented our observers with a transparency stimulus developed by Anderson and Winaver[35]. According to the heuristic they proposed, the regions whose contrast relative to the background is higher are perceived as belonging to the object surface in plain view whereas those regions whose border contrast is smaller are perceived as belonging to a transparent layer[35,36]. Following this reasoning, with this setup it is possible to change the most informative region about the albedo of the one single surface by shifting the average luminance of the backgrounds.

2. LIGHTNESS PERCEPTION OF REAL OBJECTS

.

2.1 Fixations on real objects during a color matching task

In a first experiment six naïve observers were asked to match the color of six real objects. These objects were placed in front of a computer monitor and a matching disk was displayed on the upper right corner of the screen (Figure 1a). Every object was built with one single material uniform in reflectance (Figure 1b) thus the hue was uniform across the surface, whereas the luminance varied because of shading. Objects thus exhibited a homogeneous color with variations limited to lightness and saturation, as is common in real objects[37]. We therefore focused our analysis on luminance. Observers were instructed to “match the color of the object” by adjusting the color of a 5° visual angle disk. Adjustments were done in CIE L*C*h* color space. This space is basically the cylindrical representation of the CIE L*a*b*[37] color space with $C^* = \sqrt{(a^*)^2 + (b^*)^2}$ and $h^* = tan^{-1}\frac{b^*}{a^*}$. L*a*b* color space has been designed to be approximately percptually linear, thus helping the observers to move in the color space. Observers could independently adjust the hue (h*), saturation (C*) and lightness (L*) of the disk, they provided six matches that were averaged in the analysis. Eye movements were recorded with a head-mounted, video-based eye tracker (EyeLink II; SR Research Ltd., Osgoode, Ontario, Canada) and were sampled at 500 Hz. Observers viewed the display binocularly but only the right eye was sampled. Stimulus display and data collection were controlled by a PC. The eye tracker was calibrated at the beginning of each session. At the beginning of each trial the calibration was checked. If the error was more than 1.5° of visual angle a new calibration was performed, otherwise a simple drift correction was applied. A calibration was accepted only if the validation procedure revealed a mean error smaller than 0.4° of visual angle. To ensure the alignment of spectral data and eye tracking data, we superimposed the objects on a fixed reference grid presented on the computer screen. We measured the spectrum on each cell of the 22x34 grid. This grid cell size was 1.5° of visual angle thus it matched the spatial reliability of the eye tracking system. For the spectral measurements we used a Photo Research PR-650 spectroradiometer, which has a spatial resolution of 1° of visual angle, the instrument was pointed at the center of every grid from the observer’s point of view. Figure 1c shows the color renderings of the spectral images of the six objects.

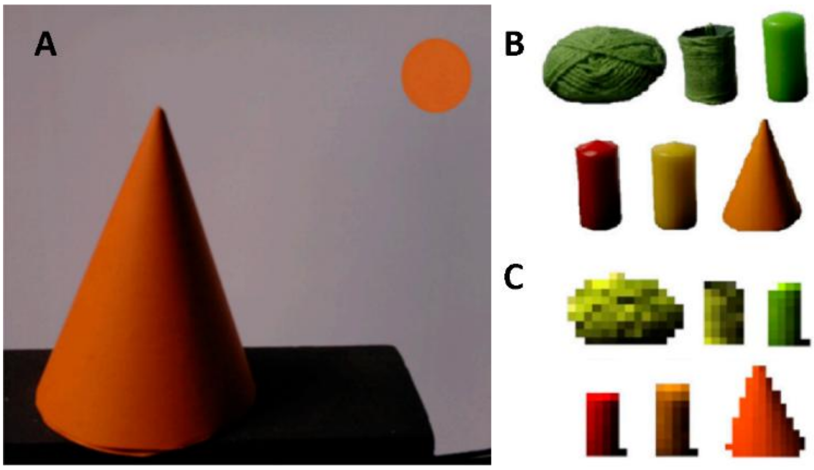

Figure 1. **Experimental setup**. A. The real objects were presented on a black stand placed in front of a computer screen, on the top right corner of the screen a matching disk was displayed. The observers could adjust the color of the disk with no time limit, when they perceived the disk to match color of the surface of the objects the observers pressed the space bar key on the computer keyboard and a new object was placed in front of the screen. B. Photos of the six real objects, a ball of green wool, a cylinder covered of the same wool, a green a red and a yellow cylindrical candle and a cone made of orange paper. C. Spectral images of the same objects measured by pointing a spetroradiometer to the center of each of the cells of a 22x34 grid presented on the computer screen. The spectroradiometer instrument was pointed at the center of every grid from the observer’s point of view. The spectral images are rendered in RGB color for this figure.

2.1.1 Matching results

The luminance matches produced by observers were on average significantly higher than the average luminance of the surfaces (Figure 2a; $t_5 = 11.6084$, $p < .001$). Since the matches closely correspond to the brightest parts of the objects, the brighter regions of the objects are weighted more heavily. Comparable results were obtained when the observers, instead of matching the hue, saturation and luminance of a colored cone, matched only the lightness of an achromatic cone (Figure 2b). Five observers participated in this experiment. The matches are all above the mean of the object and close to the upper border of the cone luminance range. The matches on average are significantly higher than the object's mean luminance ($t_4=4.61$, $p<0.05$). In the right side of Figure 2a are shown the observers' average lightness matches together with samples of the same object (samples from the paper used to build the cone) under different illumination conditions.

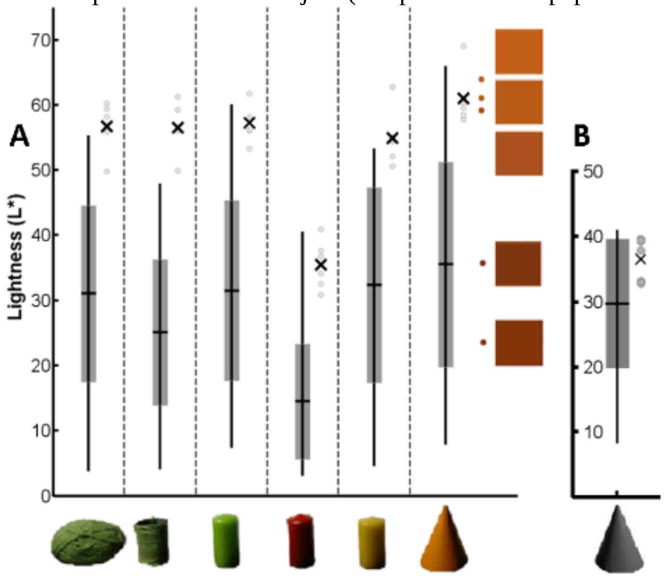

Figure 2. **Lightness matches and objects' distributions**. A) Gray circles represent the mean object matches from each of the six observers. Black crosses represent the mean matches for each object averaged across the observers. The gray bars represent the mean±1 SD and the black vertical lines represent the range of the distribution of L* within the objects. B) Gray circles represent the average matches, the black cross represents the average across the observers. The gray bar represents the mean ± 1 standard deviation and the black vertical line represents the range of the distribution of L* within the object.

From top to bottom the orange squares represent the color of the paper oriented perpendicular to light source, the mean match, the color of the maximum luminance point within the object, the average color within object, the color of the paper mounted on the computer screen. The average match is very similar to the brightest parts of the cone and the piece of paper cut out from the cone such that it is oriented perpendicular to the light source, maximizing its luminance because of the optimal illumination. The average match is also much brighter than the mean luminance across the whole object. This raises the question why observers in perceiving the lightness of a surface weight more the brightest parts of the objects, and whether this strategy is of any advantage.

2.1.2 Eye movement results

We investigated a possible relationship between fixations and lightness perception. Observers based their lightness matches on the highest parts of the objects' luminance distributions, therefore if the perceived lightness of a surface depends on the fixations distribution on this surface, it is reasonable to assume that the fixated luminances are on average higher than the average luminances of the objects. Figure 3 shows the luminance (L*) distributions of the objects and the distributions of the fixated luminances of these objects. In Figure 3a the luminance distributions of the objects are represented by the individual data points by red circles and summarized by box plots (red). The individual fixated luminances are represented by red circles and their distributions are summarized by box plot (blue). Though the object luminance histograms are typically skewed toward darker values, the luminance distributions associated with fixations are centered on values higher than the median. Binomial tests performed for each object revealed that the proportions of fixations on brighter points than the median were significantly higher than chance (all Ps < 0.001). This is visible in Figure 3b where the medians of the fixated luminances of the six objects are plotted function of the medians of luminance distributions of these objects and all the data points fall above the unity line. A similar result was obtained when observers were required to match the lightness of a gray cone, the median luminance of the fixated regions was significantly higher than the median of the luminance distribution (binomial test, 184 out of 307, $p<0.05$).

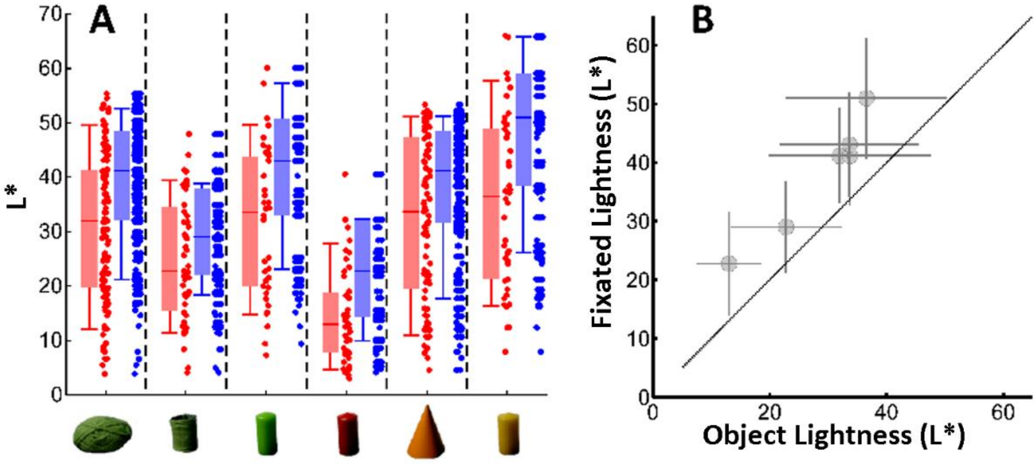

Figure 3. **Luminance (L*) distributions and the distributions of the fixated luminances of the six objects**. A. The luminance distributions of each object is represented by box plots (red), where the red boxes are delimited by the third and the first quartile and whiskers represent the 91st and the 9th percentiles of each distribution. The medians of the distributions are represented by the thick red horizontal lines, whereas the single data points by red circles. The distributions of the fixated luminances on each object is represented by box plots (blue), where the blue boxes are delimited by the third and the first quartile and whiskers represent the 91st and the 9th percentiles of each distribution. The medians of the distributions are represented by the thick horizontal blue lines, whereas the single data points by blue circles. B. The medians of the fixated luminances of the six objects are plotted function of the medians of the luminance distributions of these objects (gray circles). The bars represent a deviation from the median of the half of the interquartile difference for the objects distributions (horizontal bar) and the fixated luminances (vertical bar), for the six objects. All the data points are above the unity line (dotted).

Since in the matching experiment the lightest part of the objects was always nearest to the matching disk because the right side of the objects was exposed to direct illumination from the light source and the matching disk was presented in the right side of the screen. This configuration might have produced the spatial bias in the fixation distribution. We therefore tested whether the object luminance causally drives the observers' gaze. To this aim we repeated the color matching experiment inverting the light gradient of the objects while the mean illumination was kept roughly constant. Seven observers participated in this experiment. When the object was brighter on the left side, the fixations tended to focus on the left as compared to the original experiment, when the light had the opposite gradient (Figure 4). This means that fixations are driven by illumination, and that observers tended to fixate points brighter than the mean of the object.

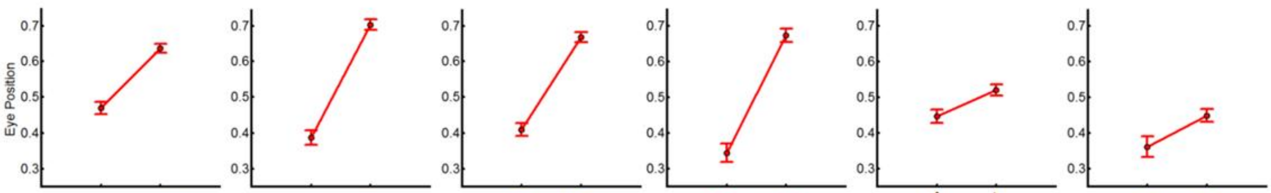

Figure 4. **Fixation shift with the illumination gradient**. Each diagram represents data for one object illuminated from the left and from the right respectively. For each fixation, a horizontal relative position has been computed from the extreme left point to the extreme right point of the object (0 and 1, respectively). Red points represent the mean of all recorded fixations. Vertical bars are the standard errors of the mean. When the object was brighter on the left side, the horizontal fixation position distribution was shifted to the left as compared to the case when the light had the opposite gradient (original matching experiment)

2.1.3 Test of a causal role of fixations

We demonstrated that the light distribution reflected from the surface of the objects influences the fixation behavior of the observers, the next logical step was to test whether the relationship between fixated regions of the objects and matched lightness is causal in nature: that is, whether where we look determines what we see. In order to manipulate the information sampled in central view and test whether this has an impact on the perceived lightness, we forced our observers to look at particular bright or dark regions of the objects. We presented digitized photographs of three of the six objects on a computer screen and monitored eye movements while the observers performed the matching task. In this gaze-contingent setup, we erased the object image from the screen whenever observers tried to fixate a part of the object outside of the predefined fixation region. Whenever this happened, a dot was shown indicating the desired fixation position. The object image was also displayed when the observers fixated on the area of the matching box, so that the parafoveal information was available, as it was in the first two experiments. We selected the forced bright (LF, light fixation) and dark fixation (DF) points based on the spatial distribution of luminance within the objects (average L* difference = 21.09). Namely, we fitted a linear plane to the luminance of every object image and we selected the 25th and the 75th percentile of this plane for the two fixation points (Figure 5B: DF, LF respectively). This choice allowed us to avoid local minima of maxima and to ensure a certain spatial distance between the dark and the light fixation points. To control for a possible effect of the direction of the illumination gradient, images were presented in two versions, illuminated from the left or from the right. The matches were significantly brighter in condition LF, as shown in Fig. 5A-B. A t-Test analysis revealed a significant effect of the fixation condition in both gradient conditions ($t3 = 4.10$, $P < 0.05$; $t3 = 8.72$, $P < 0.05$).

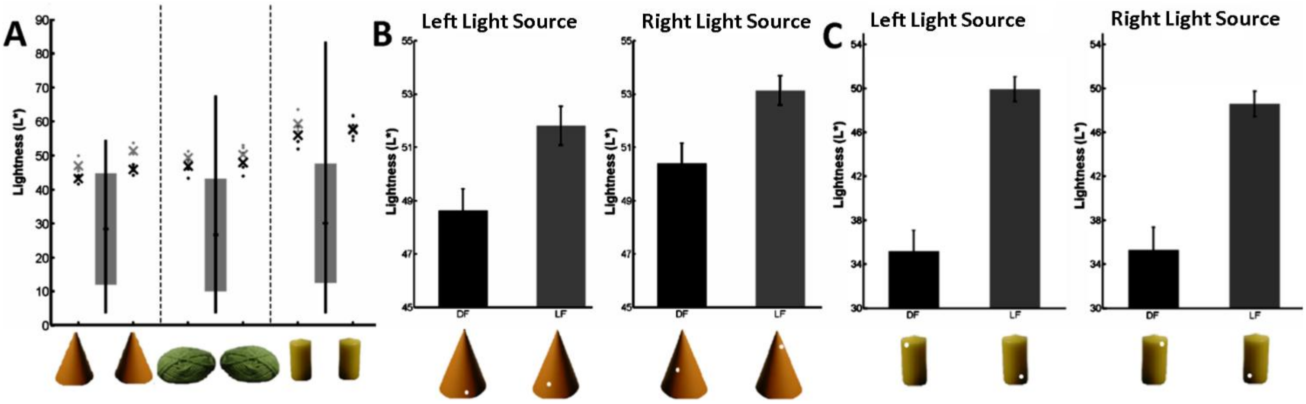

Figure 5. **Forced fixation matching results**. A. Dark gray and light gray circles represent the mean object matches from each of the six observers, respectively, in the DF and in the LF condition. Dark and light crosses represent the mean matches for each object averaged across the observers, respectively, in the DF and LF condition. The gray bars represent the mean ±1 SD and the black vertical lines represent the range of the distribution of L* within the objects. Matches are always within the ranges. B. Lightness matches in the LF and in the DF condition: means and SEs of the matches. Data for images with a light gradient from the left side to the right are presented on the left and data for images with the opposite gradient are presented on the right. Black vertical bars represent the SEs. The four pictures at the bottom represent examples of the stimuli with the white dot indicating the fixated area. C. Retinotopic adaptation. The two panels represent the matches in the two fixation conditions (dark fixation –DF, and light fixation – LF) with the light source coming from left and right respectively. Dark bars represent the observers (N=4) average in the dark fixation condition and light bars in the light fixation condition, black vertical lines represent the standard error.

In the forced fixation experiment observers saw the image only at a constant retinal position. This could cause retinotopic adaptation specific to the foveal area. The fact that observers tended to set a darker color after fixating the darker part of the image could be explained by the fact that they were less light adapted and thus the matching disk appeared brighter. The opposite could be true in the bright fixation condition. To control for this possibility we repeated the gaze contingency experiment with parafoveal matches. The matching disk was presented at the center of the computer screen only when the observers (4 naive undergraduate students) fixated a spot on the right side of the screen. We chose the position of the fixation point and matching disk so that the retinal location of the matching disk was adapted to dark while the object was in the fovea. We chose three objects with different colors but similar shape (candles), so that all objects covered the same retinal area. In order to increase the effectiveness of our experimental manipulation, we chose more extreme points in the objects' luminance distribution (5th and 95th percentile) compared to the previous forced fixation experiment. Percentiles have been taken from the low pass filtered images luminance distributions to avoid local minima. The results qualitatively matched the ones from the previous experiment (Figure 5C). The effect of fixation condition is highly significant (Left light source condition: $p < 0.001$, Right light source condition: $p < 0.001$), thus local adaptation cannot explain the effect of fixation region.

2.2 Lightness versus luminance

It is possible to speculate that observers were not actually judging lightness, the apparent reflectance, but local brightness, the apparent luminance[24]. This seems unlikely based on the results shown in Figure 5A since the luminance difference between the observers' matches under the LF and DF conditions was only 18% of the actual luminance difference between the fixated bright and dark regions. Furthermore, the matches had a considerably higher luminance than the fixated regions, being not compatible with a simple local match. To address this issue directly we explicitly asked our observers to perform a lightness match[38]. Observers were presented with gray scale images of the same objects used in the forced fixation experiment: the cone, the candle and the wool ball. The images were displayed on the computer screen only when the observer was fixating a chosen point (dark or light). The images were presented on the left part of the screen (Figure 6A). 16 real paper chips (Figure 6B) defined by different reflectances from black to white were randomly placed on a board on the right part of the screen. The board was illuminated by a nearby bulb lamp, which produced a strong illumination gradient (about 30 candelas range on the lightest chip).

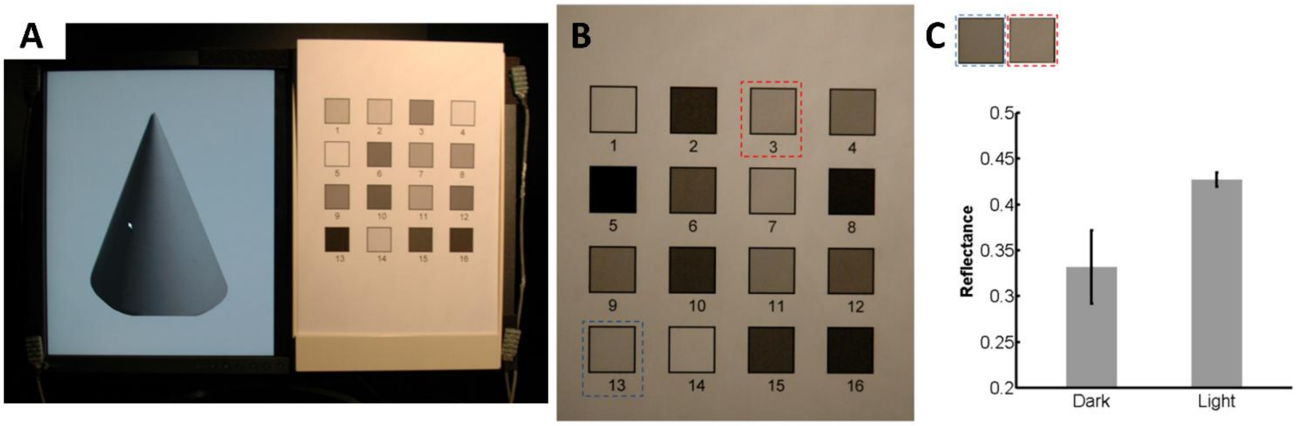

Figure 6. **Lightness matching**. A. The images were presented on the left part of the screen and on the right part 16 real paper chips defined by different reflectances from black to white were randomly placed on a board illuminated by a bulb lamp standing quite close to the board which produced a strong illumination gradient of about 30 candelas range from the darkest to the lightest chip. B. One of 10 random arrangements of the paper chips on the page. Observers had to choose the chip they thought had the same material as the object displayed on the nearby computer screen. B. Set of 16 paper chips. C. Results of the lightness matching experiment. The two bars represent the average matched reflectance for the dark and light fixation conditions. The error bars represent the standard error. In the dark fixation condition, observers chose chips with lower reflectance than in the light fixation condition. Figure B and C also exemplify the dissociation between lightness and brightness. Chip three (red frame) in B looks darker than chip thirteen (blue frame), and it was consistently ranked lower both in terms of lightness and brightness by three observers. However chip three has a higher reflectance than chip thirteen (74% and 60%) although its luminance is lower (29 cd/m^2 and 44 cd/m^2). Indeed, once embedded in a common context (C), chip thirteen looks darker than chip three (upper central squares).

Four naive observers had to "pick the chip made of the same material as the object". The presence of two light fields and the explicit paper-matching task are designed to induce the observers to perform the task in terms of lightness[38]. The chips were presented in 10 different random spatial locations. The random placement of the chips in the illumination gradient added luminance noise, producing dissociations of reflectance and luminance. Under such viewing conditions photometric luminance is typically not perceptually accessible and perceived brightness heavily depends on the context[31,39]. We computed the effect of fixation position on the reflectance selections. In the dark fixation condition, observers chose chips with lower reflectance than in the light fixation condition ($t_3 = 3.3641$, $p < .05$). To ensure that observers were indeed judging reflectance in this experiment, we asked three observers to sort the chips in terms of their brightness and in terms of the lightness of their paint. Observers were practically unable to distinguish between perceived shade of the paint and perceived brightness (Spearman's correlations between the rankings in the two different tasks for the three observers: .991, .998 and .989). Due to the illumination gradient, there were cases where reflectance and luminance would lead to differences in the rankings. When this was the case, observers sorted the chips according to their reflectance in more than 85% of all cases, irrespective of their instruction. Our results indicate that observers discarded the illumination and estimated the chips' lightness before choosing the one matching the fixated area of the object.

Taken together these results prove a direct causal link between how an object is sampled through eye movements and its perceived overall lightness. Observers tend to produce estimates of the global lightness of objects that are above the physical average of the light intensities reaching the eyes and close to the brightest object regions. The matched luminance was more similar to that of the most frequently fixated regions of the objects, as our observers tended to fixate points with above-average luminance. Furthermore, the observers matched a higher lightness when they were forced to fixate a brighter region of the image as compared to when they fixated a darker region. We can thus conclude that the way we perceive the lightness of objects is driven by the way we look at them.

2.3 Optimal lightness estimates

The brightest parts of an object frequently correspond to an area of the surface defined by a high luminance contrast. Luminance contrast is considered to be highly salient for attracting human gaze[15].This way, the viewing strategy could simply be a side effect of more general visual saliency mechanism. However, it could also have a functional role in lightness perception. If this was the case, the fixated regions should yield the best estimate of the object´s reflectance. In other words, these regions should be reliable estimates of the surface reflectance, namely they should be maximally immune to the effects of the viewing conditions on luminance and be maximally sensitive to changes in the object's reflectance. In general, the luminance of diffusely reflecting surfaces is proportional to the cosine of the angle between the surface normal and the direction of the incident light[40]. This implies that differences in surface reflectance have the largest impact on luminance when they occur in regions that are oriented perpendicular to the light source (i.e. in the brightest parts of the objects). Moreover, variations in surface orientation have the smallest effect on luminance when the surface normal is almost perpendicular to the light. This reasoning applies in the simple case where a surface is perfectly convex thus inter-reflections are not occurring and the illumination is represented by a uniform collimated light-field. In nature the light is usually coming from multiple directions at different intensities, directly from a light-source or reflected back from a surface to another or from a surface to itself. In order to test our simple predictions taking in account this complexity we used a physics-based rendering simulation to explore how well the reflectance of real-world objects can be estimated in a natural light field.

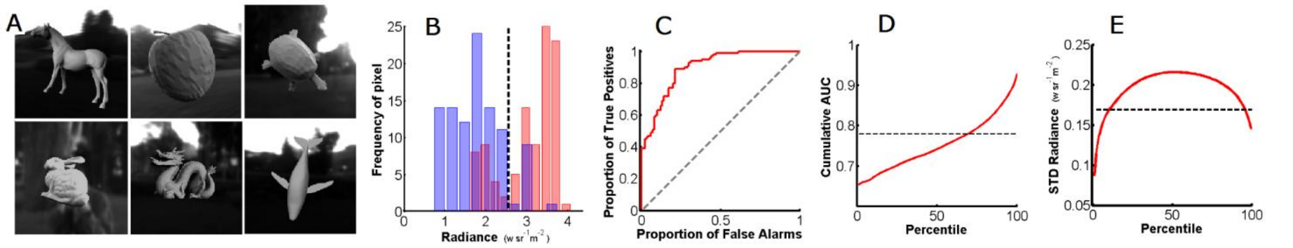

Figure 7. **Physically based rendering simulation**. A. Set of 3D models of objects used in the simulation. B. Single percentile distributions. The histograms represent the two distributions of the radiances (output of the rendering software), 100 different orientations for a certain percentile. These distributions are clearly overlapping. C. ROC curve. The ROC curve is plotted for the two distributions of B. The AUC is close to 1, indicating high discriminability. D. Classification performances for each percentile. The highest percentiles' aggregated discrimination performance is higher than the performance of the discrimination based on the object average luminance (dashed black line). E. SD for each percentile. Extreme percentiles are more stable

Using the software RADIANCE interfaced with a MATLAB toolbox[41], we rendered a set of six virtual models of objects under 100 different viewing and illumination conditions (Figure 7A). Each view was rendered with one of six values of reflectance. We performed two separate analyses for each percentile of the luminance distribution, which we used as our potential estimates of lightness. Our goal was to establish which parts of the luminance distribution are most invariant to object rotation and which parts are most diagnostic of the reflectance changes. Figure 7E shows the distribution of the variances for each estimate. Both the minimum and the maximum luminance would be reliable estimates of lightness. The drop in variability for the lowest percentiles does not easily follow from the cosine low, justifying the choice of such an extensive simulation. Reliability alone is not sufficient for a good estimate, since a constant incorrect estimate is of course perfectly reliable, but entirely useless. A good estimate of lightness also needs to be sensitive to changes in reflectance. 4B-D show the results of an ROC-analysis[42] that indicates to what extent an ideal observer can identify a change in reflectance in the presence of variations in scene geometry given a certain percentile of the luminance distribution of the rendered surfaces. The reflectance values have been chosen to lead to partially overlapping estimate distributions for each pair (Figure 7B). The area under each ROC curve is a measure of criterion independent diagnosticity (Figure 7C). In order to aggregate the different point of view in the light-field, the different objects and the different couples of reflectance, we aggregated the Area Under Curved computed for every percentile in the different conditions into a so called cumulative AUC[43] (Figure 7D). This analysis shows that the discriminability increases with the luminance of the object region that is compared. Performing the same analysis on the average luminance yields worse discrimination performance as compared to the higher percentiles as shown by the dashed line in Figure 7D. The results of both analyses indicate that the luminances of both the dark and the bright regions of objects are comparatively invariant under different views, but only the most illuminated regions are also diagnostic of the object´s reflectance.

2.4 Discussion

The lightness of the object is better estimated by the brighter regions of the object and oversampling the brighter parts thus is a good heuristic to estimate the lightness of the object. The advantage of this heuristic is that no knowledge about any high level visual aspects such as geometry or shape is required. In our case, objects were uniformly colored and not occluded. The luminance gradients could only be caused by the interaction of the surface orientation and the light field. In this case the visual system is still faced with the problem of computing one level of lightness from a non-uniform pattern of brightness. One way the system could use to solve this task is to produce separate representations of the illuminant on the surface and its albedo[44] . However, there is evidence that lightness judgments are at least to some degree directly based on brightness[30,31]. Along these lines, we propose that the brightness of the most illuminated region within the object is the source from which our visual system extracts a global impression of lightness. Coherently, the usage of the simple heuristic described here agrees with the finding that observers are impaired when they have to judge the lightness of a flat object, because they are unable to discount the effect of surface orientation[30]. So far we demonstrated that the free behavior we observed had a functional meaning: our observers preferentially sampled the brightest part of the objects they had to judge and this led to lightness matches which were close to the luminance of the upper parts of the luminance distributions of the objects, these parts are indeed the most informative about the surface reflectance. How this behavior is flexible is still unknown, namely we do not know whether the fixation behavior could functionally be explained by a focus on the surface areas which are most informative about the surface albedo also in the case those areas are not the brightest. Furthermore, the idea of a low level mediator generally acting in lightness perception is appealing per se and could provide additional insights on the nature of the processes revealed by lightness illusions. These last two points will be addressed in a subsequent series of experiments.

3. EYE MOVEMENTS IN A SEGMENTATION DISPLAY

From the experiments described above we proposed that when the visual system is involved in estimating the reflectance of a shaded surface a selection mechanism (lightness heuristic) points out which areas of the surface are particularly informative about its reflectance and these areas are preferentially sampled by fixations, this sampling contributes to the lightness percept. A demonstration of a lightness heuristic causally driving the fixation was missing so far, therefore we selected a particular display that suited the purpose[35]. Differences in contrast magnitude along a border with a uniform contrast polarity (see Figure 8) are thought to be used to segregate the image. According to this heuristic the regions whose contrast relative to the background is higher are perceived as belonging to the object surface in plain view. The regions whose border contrast is smaller are perceived as belonging to a transparent layer[36]. With this display we could investigate whether observers preferentially fixate the regions of the stimuli that are perceived to be more informative on the actual reflectance of the target (i.e. the ones less occluded by the superimposed transparent layer) and whether such a fixation strategy influences the overall perceived lightness.

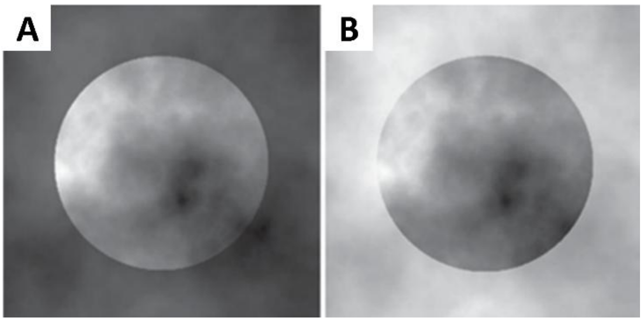

Figure 8. **Anderson's display**. The central circular regions in A and B are the same image, the surrounding noise has a different luminance range in order to change the contrast polarity that is uniform along the disc's border but with a different sign between A and B.

In this stimulus, the lowest contrast region along the border is perceived as being occluded by the foreground layer, and all regions within the object sharing the same brightness are interpreted in the same way. The areas of the object with a different contrast polarity are interpreted as not occluded and representing the actual reflectance of the target. In the dark surround condition (Figure 8A) the smallest border contrast is associated with high luminance, and the brightest parts of the object are interpreted as the least occluded regions. On the contrary, in the light surround condition (Figure 8B), where the highest contrast area along the border corresponds to darker luminance values, the darkest regions of the target are perceived to belong to the object layer. This layered representation leads to the percept of a light disc (Figure 8A) embedded in dark noise and a dark disc (Figure 8B) embedded in light noise, even though the luminance distributions within the two discs are identical. Only a small part of this illusion is explained by simultaneous contrast[35]. We recreated the stimulus starting from a pink noise texture. The noisy texture consisted of low pass filtered normally distributed random noise. The power spectrum of the noise varied as $\frac{1}{f^2}$ .The disc had the whole contrast allowed by the monitor whereas the background contrast was reduced scaling the noise by a certain factor c [0 1]. In order to have a dark and a light surround with the same contrast, the scaled noise was shifted either to the highest or to the lowest luminance value of the screen. The stimuli were presented on a calibrated Samsung SyncMaster 1100mb monitor (40cm x 30 cm; 1280 by 960 pixels; 85 Hz; 8 bits per gun).The monitor chromaticities were measured with PR650 spectroradiometer, CIE xyY coordinates for the upper right corner and for the screen center were R = (0.610, 0.329, 22.2), G = (0.282, 0.585, 76.3), B = (0.156, 0.066, 11.9). Eighteen naïve observers were asked to match the lightness of the central disc with a reference uniform disc displayed on the right side of the same CRT monitor. Observers were asked to adjust the luminance of the matching disk to be as close as possible to the one of the target disk embedded in the texture. They could change the luminance of the disk by pressing the mouse buttons. Each mouse button press corresponded to a change of 0.01 DKL[45] color space units. We represented the size of the illusion as the difference in matched luminance between the two (dark and bright surround) contextual conditions. Observers were required to produce 30 matches for each of the background conditions, for a total of 60 matches. We used the difference between the matches produced in the dark surround condition and the ones produces in the light surround conditions as a measure of the strength of the illusion. The individual stimuli differ from trial to trial because of the random independent generation of the pink noise textures. Eye movements were measured during the matching task. Gaze position signals were recorded with a head-mounted, video-based eye tracker (EyeLink II; SR Research Ltd., Osgoode, Ontario, Canada) and were sampled at 500 Hz. If foveal sampling plays a role in this illusion, it is reasonable to expect that the fixation pattern should change depending on the context. In particular, the fixations should be focused on the regions of the target assumed to be on planar view, avoiding the occluding layers. Therefore the average luminance around the fixated points should be higher in the dark surround condition than in the bright surround condition. We represented the effect of the background on the eye movement strategy as the difference between the average luminance in the regions around each fixated point (circular surround of 1° of visual angle) in the dark surround condition and in the bright surround condition.

3.1 Results

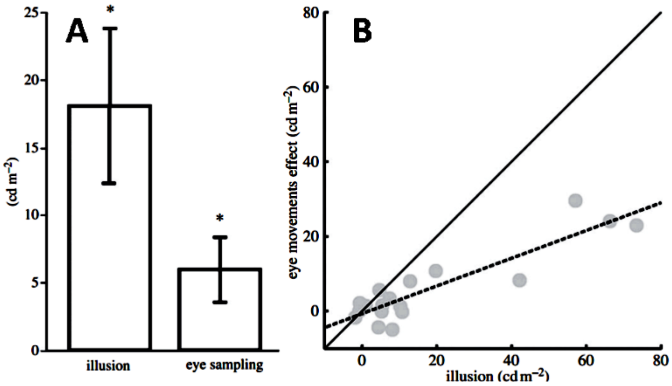

Figure 9. **Matching and eye movement sampling surround effects**. A. The illusion effect is the difference between each observer's average match in the dark and in the light surround condition. The eye movement sampling effect is the difference between the average luminance in the regions around each fixated point for each observer in the dark and bright surround condition. Error bars indicate standard errors. B. Correlation between effects of scission on eye movement sampling and lightness matches. The difference in sampling the scene between the two surround conditions is highly correlated with the perceptual difference in lightness matches.

The illusion effect was highly significant ($t17=3.18$, $p<.05$, Figure 9A) and the background condition also affected the eye sampling strategy ($t17=2.56$, $p<.05$, Figure 9A). As predicted, observers tended to focus on lighter parts of the disc in the dark background condition than in the light background condition. The illusion size was highly correlated with the contextual effect on eye sampling (Pearson's $r = 0.904$, p-value< 0.0001. Figure 9B). The viewing behavior explained 82% of the between observers' variability in the illusion size. Observers with a large perceptual effect also had large differences in eye sampling between the dark and light context condition.

3.2 Causal drive of fixation on lightness perception in a layered segmentation scene.

We actively manipulated the observers' fixation strategy using a gaze contingent display in order to investigate the nature of the relationship between perception and eye sampling. To manipulate the foveal information we used the same gaze contingent paradigm described in 2.1.3. Namely, we used the same stimuli as in the previous free looking experiment, but they were displayed only when the observers fixated a particular region within the target disc. We introduced a dark and a light fixation condition by choosing a dark and a light fixation point on a bright and dark fixation regions. These regions were chosen as the brightest and the darkest point at least 1° of visual angle distant from the edge of the circle. The luminance of the forced fixation region strongly influenced lightness perception: fixating brighter parts of the stimulus led to brighter matches, fixating darker parts led to darker matches (Figure 10A).

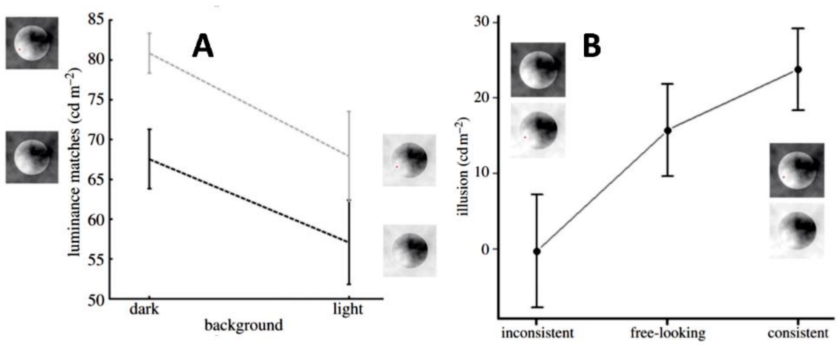

Figure 10. **Forced fixation results**. A. The grey line represents the average matches in the light fixation condition, and the black line represents the average matches in the dark fixation condition. Pictures of the stimuli indicate the fixation conditions with the fixation points indicated by small spots. Error bars indicate standard errors. B. Pictures of the stimuli represent the condition indicated by the label on the x-axis. The illusion effect is represented as the difference between the matches for the pictures above and below. Error bars indicate standard errors.

A repeated measurement ANOVA revealed a strongly significant main effect of fixation condition (F1,14 =12.81, p-value<0.005).The interaction between fixation condition and surround condition was not significant (F1,14 = 23.3915, p-value > 0.25). The main effect of surround condition was just on the edge of significance (F1,14 =4.3589, p=0.05). In order to quantitatively compare the illusion effect in the free looking condition and the forced fixation condition, we computed the illusion effect from the forced looking experiment as the differences between the dark and the light surround conditions. We did this for the case when the forced fixation pattern was consistent or inconsistent with the spontaneous fixation strategy observed earlier (Figure 10B). In the free looking experiment observers tended to focus on darker areas of the surfaces in the light background condition and on lighter areas of the surface in the dark background condition. In the consistent condition observers fixated the bright region in the contextual condition that led to the perception of a light target disc and fixated the dark region in the contextual condition that led to the perception of a dark target disc. In the inconsistent condition the observers fixated the darkest region in the contextual condition that led to the perception of a bright target disc and fixated the brightest region in the contextual condition that led the perception of a dim target. Between the consistent and the inconsistent condition only the fixation area differed, but the overall stimulus was the same, therefore a difference between them could only be explained by an effect of the forced fixation. The illusion effect was significantly higher in the consistent condition than in the inconsistent condition (t14 = 3.5793, p <.005) and in this inconsistent condition the effect disappeared completely (t14 = -0.0436, p >0.95). The illusion effect measured in the free looking condition was significantly higher than in the inconsistent condition (t14 = 3.3545, p <.005), and just lower than in the consistent condition (t14 = 2.6178, p <.05). Forcing the observer to a fixation strategy that emphasized the fixation strategy naturally applied by our observers, we found a stronger illusion effect. Forcing the observers to a fixation strategy at odds with their natural behavior destroyed the illusion.

3.3 Causal role of the segmentation heuristic on the fixation strategy.

These experiments have clearly shown that perceptual effect of the segmentation process is amplified by eye movements, indicating that the foveal input was weighted more than the parafoveal input. The next question is whether eye movements are actually driven by the segmentation process or by some other property of the image, like the test-background contrast, for example. Indeed the luminance-contrast is one of the most important cues for fixation prediction[46–49]. To address this issue we repeated the free looking experiment rotating the test discs 90° (Figure 11A-D)

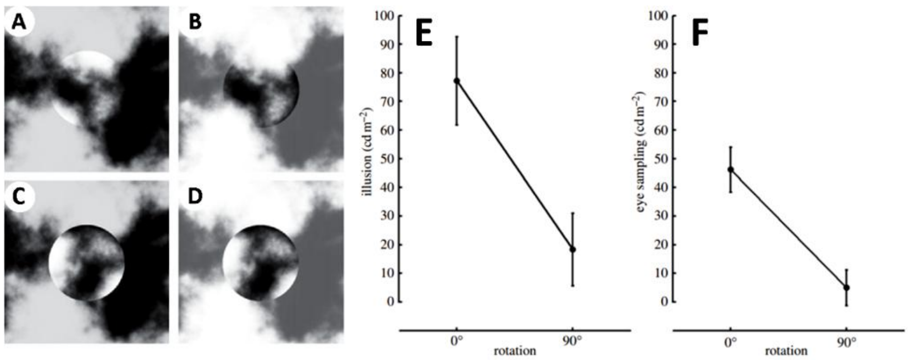

Figure 11. **Rotation Experiment**. A-B. Non-rotated stimuli in the dark and the light surround condition. C-D 90° rotated versions of C-D. After rotation, the contrast polarity along the border is not uniform anymore and the perception of transparency does not occur. The illusion looks vastly diminished in the squares C-D.

This simple manipulation disrupts the segmentation process because the contrast polarity along the border is no longer uniform[34]. The illusion effect (Figure 11E) was highly reduced in the rotated version of the stimulus (t5=3.27, p<0.05). The residual surround effect could be explained by simultaneous contrast, which was found to contribute 11% of the illusion[35]. The difference in the eye sampling strategy (Figure 11F) was so lower in the rotated condition (t5=4.19, p<0.05) to practically disappear. Observers did not fixate more on the sharp borders of the rotated target, showing that the difference in the eye sampling is likely not to be due to the change in the edges achieved by the rotation. This finding indicates that the eye sampling is causally driven by the segmentation heuristic.

3.4 Discussion.

Similarly with section 2, our experiments have shown that eye sampling is closely coupled with the perception of lightness. Fixations are concentrated on the layers of the image that represent the reflectance of the stimulus rather than the occluding layers. In the dark background condition, the lighter areas of the stimulus are preferentially fixated compared to the light background condition where the contrast relationships are reversed and the darker areas are preferred. This shows that the fixation strategy is flexible. From the forced fixation experiments we confirmed the observation that the lightness appearance of a surface causally depends on how the surface is fixated. The last experiment proved that the fixation strategy is causally influenced by the segmentation heuristic. We do not claim that the illusion could depend also on other factors than the fixation pattern on the surface but we conclude that in the causal chain that leads to the lightness appearance of the display we used fixations function as a mediator between the lightness heuristics and the lightness percept.

4. OVERALL DISCUSSION

The information the visual system could exploit to achieve lightness constancy consists of the luminance histogram of reflected light and its spatial configuration within the scene. There are suggestions that the visual system could exploit some non-spatial property of the reflected light distribution, such as the skew of a scene's luminance histogram[32] or the overall brightness or contrast in the scene[31]. The visual system could also make use of the spatial features of the scene, such as an objects' shape[29] or edge contrast[36]. To exploit the spatial properties of a scene, a spatial selection process could be necessary to focus on the relevant features. Here we propose that fixations locally sample a surface focusing on particularly informative aspects of its luminance distribution. The properties which do not have a local spatial representation, such as the skew of the luminance histogram or its average, could be computed indirectly through a local sampling process. The idea we propose is consistent with the idea that the visual system is continuously scanning the visual field to focus on the information that is relevant for the current task[21,23]. Here we provided two examples in the lightness perception domain, where a selection process is used by the visual system to estimate the surface reflectance. This selection process is driven by relatively simple heuristics which determine the parts of a scene that are particularly relevant for estimating the surface reflectance. The selection process occurs by a foveally based sampling process that acts via fixation eye movements. The first series of experiments show that when observers have to match the color of a set of natural objects, they base their judgments on the brightest parts of those objects. The maximum is a local property of the image and therefore well suited for a local sampling mechanism. The same consideration is true for all of the percentile statistics, whereas it is not true for the average of the luminance distribution. With an extensive series of lighting simulations, we proved that the maximum of the luminance distribution is the percentile statistic that is most informative about the reflectance of real matte objects. Thus, it is reasonable to propose that the visual system first extracts a rough sketch of the scene through its peripheral vision, and then it samples the scene, focusing on the local properties that are physically most informative about the reflectance of surfaces. According to this view, eye movements do not only select the information that access the visual system but they weight it differently according to the needs of the current task, as proposed by Rensink's *coherence theory*[21]. This way of thinking about eye movements is consistent with the findings on ambiguous figures described in the introduction, where even though the whole figure was visible parafoveally the information centrally sampled was highly weighted in eliciting a dominant percept[9,10]. In a second study we explored the flexibility of this strategy. We chose a display where a simple heuristic would select regions of the scene that are mostly informative about the surface reflectance of a target disk. Those areas are not necessarily the brightest areas of the stimulus, and by only changing the background (keeping the target stimulus constant), the heuristic selects different areas within the same surface. This indicates that the visual system locally samples lightness in a flexible and adaptive manner and that the same surface could be scanned differently according to what the system finds relevant to estimate the surface reflectance.

5. REFERENCES

[1] Hansen, T., Pracejus, L.., Gegenfurtner, K. R., “Color perception in the intermediate periphery of the visual field,” J. Vis. **9**(4), 26 (2009).

[2] Virsu, V.., Rovamo, J., “Visual resolution, contrast sensitivity, and the cortical magnification factor,” Exp. Brain Res. **37**(3), 475–494 (1979).

[3] Virsu, V., Näsänen, R.., Osmoviita, K., “Cortical magnification and peripheral vision,” JOSA A **4**(8), 1568–1578 (1987).

[4] Sachs, M. B., Nachmias, J.., Robson, J. G., “Spatial-frequency channels in human vision,” JOSA **61**(9), 1176–1186 (1971).

[5] Davis, E. T., Yager, D.., Jones, B. J., “Comparison of perceived spatial frequency between the fovea and the periphery,” JOSA A **4**(8), 1606–1611 (1987).

[6] Thorpe Davis, E., “Modeling shifts in perceived spatial frequency between the fovea and the periphery,” JOSA A **7**(2), 286–296 (1990).

[7] Marks, L. E., “Brightness as a function of retinal locus,” Percept. Psychophys. **1**(10), 335–341 (1966).

[8] Osaka, N., “Brightness exponent as a function of retinal eccentricity in the peripheral visual field: Effects of dark and light adaptation,” Percept. Psychophys. **27**(6), 519–523 (1980).

[9] Einhäuser, W., Martin, K. A.., König, P., “Are switches in perception of the Necker cube related to eye position?,” Eur. J. Neurosci. **20**(10), 2811–2818 (2004).

[10] Georgiades, M. S.., Harris, J. P., “Biasing effects in ambiguous figures: Removal or fixation of critical features can affect perception,” Vis. Cogn. **4**(4), 383–408 (1997).

[11] Cornelissen, F. W.., Brenner, E., “Simultaneous colour constancy revisited: an analysis of viewing strategies,” Vision Res. **35**(17), 2431–2448 (1995).

[12] Hansen, T.., Gegenfurtner, K. R., “Classification images for chromatic signal detection,” JOSA A **22**(10), 2081–2089 (2005).

[13] Brenner, E., Granzier, J. J.., Smeets, J. B., “Perceiving colour at a glimpse: The relevance of where one fixates,” Vision Res. **47**(19), 2557–2568 (2007).

[14] Yarbus, A. L., Haigh, B.., Rigss, L. A., Eye movements and vision, Plenum press New York (1967).

[15] Itti, L.., Koch, C., “A saliency-based search mechanism for overt and covert shifts of visual attention,” Vision Res. **40**(10), 1489–1506 (2000).

[16] Itti, L.., Koch, C., “Computational modelling of visual attention,” Nat. Rev. Neurosci. **2**(3), 194–203 (2001).

[17] Schütz, A. C., Braun, D. I.., Gegenfurtner, K. R., “Eye movements and perception: A selective review,” J. Vis. **11**(5), 9 (2011).

[18] Hayhoe, M., “Vision using routines: A functional account of vision,” Vis. Cogn. **7**(1-3), 43–64 (2000).

[19] Land, M. F.., Hayhoe, M., “In what ways do eye movements contribute to everyday activities?,” Vision Res. **41**(25), 3559–3565 (2001).

[20] Land, M., Mennie, N.., Rusted, J., “The roles of vision and eye movements in the control of activities of daily living,” Percept.-Lond. **28**(11), 1311–1328 (1999).

[21] Rensink, R. A., “The dynamic representation of scenes,” Vis. Cogn. **7**(1-3), 17–42 (2000).

[22] Ballard, D. H., Hayhoe, M. M., Li, F., Whitehead, S. D., Frisby, J., Taylor, J.., Fisher, R., “Hand-eye coordination during sequential tasks [and discussion],” Philos. Trans. R. Soc. Lond. B. Biol. Sci. **337**(1281), 331–339 (1992).

[23] Ballard, D., Hayhoe, M.., Pelz, J., “Memory representations in natural tasks,” Cogn. Neurosci. J. Of **7**(1), 66–80 (1995).

[24] Arend, L. E.., Spehar, B., “Lightness, brightness, and brightness contrast: 1. Illuminance variation,” Percept. Psychophys. **54**(4), 446–456 (1993).

[25] Cornsweet, T., Visual perception, Academic press (2012).

[26] Land, E. H.., McCann, J., “Lightness and retinex theory,” JOSA **61**(1), 1–11 (1971).

[27] Mach, E., “Uber die physiologische Wirkung rumlich vertheilter Lichtreize,” Akad Wiss Wien Sitzber Math-Nat K **54** (1868).

[28] Koffka, K., Principles of Gestalt psychology, Routledge (2013).

[29] Knill, D. C.., Kersten, D., “Apparent surface curvature affects lightness perception,” Nature **351**(6323), 228–230 (1991).

[30] Ripamonti, C., Bloj, M., Hauck, R., Mitha, K., Greenwald, S., Maloney, S. I.., Brainard, D. H., “Measurements of the effect of surface slant on perceived lightness,” J. Vis. **4**(9), 7 (2004).

[31] Robilotto, R.., Zaidi, Q., “Limits of lighness identification for real objects under natural viewing conditions,” J. Vis. **4**(9), 9 (2004).

[32] Motoyoshi, I., Nishida, S., Sharan, L.., Adelson, E. H., “Image statistics and the perception of surface qualities,” Nature **447**(7141), 206–209 (2007).

[33] Anderson, B. L., “A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions,” Percept.-Lond. **26**(4), 419–454 (1997).

[34] Singh, M.., Anderson, B. L., “Toward a perceptual theory of transparency.,” Psychol. Rev. **109**(3), 492 (2002).

[35] Anderson, B. L.., Winawer, J., “Image segmentation and lightness perception,” Nature **434**(7029), 79–83 (2005).

[36] Anderson, B. L., “The role of occlusion in the perception of depth, lightness, and opacity.,” Psychol. Rev. **110**(4), 785 (2003).

[37] CIE 1978., “Technical Report No. Supplement No 2 of CIE Publ,” 15 (1971).

[38] Arend, L.., Reeves, A., “Simultaneous color constancy,” JOSA A **3**(10), 1743–1751 (1986).

[39] Gegenfurtner, K. R.., Sharpe, L. T., Color vision: From genes to perception, Cambridge University Press (2001).

[40] Klett., Eberhard Witwe., Detleffsen., Christoph Peter., IH Lambert... Photometria sive de mensura et gradibus luminis, colorum et umbrae, sumptibus viduae Eberhardi Klett (1760).

[41] Lichtman, D., Xiao, B.., Brainard, D. H., “RenderToolbox Version 2” (2007).

[42] Green, D., “Swets, JA (1966). Signal detection theory and psychophysics,” Psychol. Bull. **75**, 424–429 (6).

[43] Sonego, P., Kocsor, A.., Pongor, S., “ROC analysis: applications to the classification of biological sequences and 3D structures,” Brief. Bioinform. **9**(3), 198–209 (2008).

[44] Adelson, E. H., “Zj^±Lightness Perception and Lightness Illusions,” New Cogn. Neurosci., 339 (2000).

[45] Derrington, A. M., Krauskopf, J.., Lennie, P., “Chromatic mechanisms in lateral geniculate nucleus of macaque.,” J. Physiol. **357**(1), 241–265 (1984).

[46] Jovancevic, J., Sullivan, B.., Hayhoe, M., “Control of attention and gaze in complex environments,” J. Vis. **6**(12), 9 (2006).

[47] Koch, C.., Ullman, S., “Shifts in selective visual attention: towards the underlying neural circuitry,” [Matters of intelligence], Springer, 115–141 (1987).

[48] Parkhurst, D., Law, K.., Niebur, E., “Modeling the role of salience in the allocation of overt visual attention,” Vision Res. **42**(1), 107–123 (2002).

[49] Reinagel, P.., Zador, A. M., “Natural scene statistics at the centre of gaze,” Netw. Comput. Neural Syst. **10**(4), 341–350 (1999).