On the concept of error in visual perception: An example from simultaneous lightness contrast

DANEYKO, Olga and ZAVAGNO, Daniele

Published version

DANEYKO, Olga and ZAVAGNO, Daniele (2008). On the concept of error in visual perception: An example from simultaneous lightness contrast. Teorie & Modelli, 8 (2-3), 175-184.

Teorie & Modelli, n.s., XIII, 2-3, 2008 (175-184)

On the concept of error in visual perception: An example from simultaneous lightness contrast

Olga Daneyko[*] (Trieste), & *Daniele Zavagno* (Milan)[**]

1. Introduction

In vision sciences the multiform category of phenomena known as *illusions* are most often treated as errors or mistakes made by the visual system (Gregory *et al.* 1995). This view is almost always coupled with notion of *veridicality* and *veridical perception* (Rock 1983). The notion of veridical perception implies that in some cases we do not see what we should see, and that in other cases we see exactly what should be seen. Hence, two simple but intriguing questions arise: (1) Given a visual scene, what is it that we should exactly see? (2) How do we know that what we should see is actually what we are seeing? Experimental research on illusions often deals sideways with these questions, and seldom comes up with a straight answer.

The variety of visual illusions is huge, and counting also the numerous versions that exist of a same illusion (Da Pos & Zambianchi 1996), the field comprises a rich bestiary of "curious" phenomena. To frame our investigation we will consider only those illusions that are made of two *objects* that should look alike along some, or all, dimensions – given that their respective physical characteristics are identical – but that appear different because of other surrounding features in the visual scene. For sake of simplicity we shall call this type of phenomena "comparison illusions".[1] In particular, we will be writing about this class of illusions in lightness perception.

[*] Department of Psychology "Gaetano Kanizsa", University of Trieste.

[**] Psychology Department, University of Milano-Bicocca.

[1] The term is used only as shorthand; all illusions are actually of a "comparison" type, since the acknowledgement that one is experiencing an illusion implies a comparison between what is perceived and the knowledge and the expectations of the observer.

Researching the mechanisms or conditions that underlie an illusion usually goes hand in hand with a quantification of its magnitude. This last aspect, however, can be tricky, for some term of comparison is needed. With comparison illusions there are usually two alternatives:

- A measurement of the perceptual differences between two objects that do not appear equal despite their corresponding physical dimensions are equal. This is what we call a "within approach", and the basic question sounds like: *how different are the two objects?*

- A measurement of the differences between the perceptual dimensions of the two objects and their corresponding physical dimensions, to determine the magnitude of the errors made by the visual system. This is what we call a "between approach", and the basic question sounds like: *which of the two physically equal objects shows a bigger departure from their actual (physical) dimension*?

The within approach is epistemologically safe: the researcher stays within the phenomenal domain and eventually uses physical units as a numeric expression of the quantifications carried out, without bothering to ask which of the two objects is *more responsible* for the illusion.

The between approach is instead epistemologically more problematic. Lets just consider one of the problems: which of the two perceived objects of a comparison illusion is closer to its physical model? The *between approach* is closely related to two very different ways of thinking how the visual system works: (1) the visual system has direct access to the structure of the physical scene, in the sense that it sees what is actually there: things appear as they do because they are what they are (Gibson, 1979); (2) the visual system generates a good enough replica of the physical scene by means of inferences on raw sensations based on logical and statistical assumptions driven by cues and past experience (Rock, 1983; Knill & Richards, 1996). Given that both these theoretical views meet large approval in the field of perceptual sciences, it is not surprising that the *between approach* nourishes the interest and indeed the fantasies of many vision scientists

Instead of developing logical arguments against the *between approach* and all the metaphysics that go along with it, we shall try to generate some doubt about the usefulness of the concept of *error* in visual sciences by reporting the results of an experiment we carried out in the field of lightness perception.

2. The "locus of error" paradigm

The phenomenon we examined is *simultaneous lightness contrast* (SLC)[2], which has served as a test-bench for many lightness theories. A question that can be asked is the following: *where does the visual system go wrong when in a SLC display it assigns different lightness values to targets that are photometrically equal?* This question specifically targets the *locus of error* of the illusion and it is at the base of a clever experimental paradigm developed by Gilchrist *et al.* (1999).

The question is not trivial: if we were able to answer similar questions we could easily reject theories that make wrong predictions with reference to both the direction and the magnitude of certain illusions. In fact, we think it is quite fair to claim that in the field of lightness perception this is a core question for different theories, from classic contrast theories to more recent approaches such as the Anchoring theory (Gilchrist, 2006). The Anchoring theory in particular is explicitly based and developed on the notion of errors (Gilchrist et al., 1999), which are to be considered as the trademark of the visual system, as Gilchrist (2006) himself puts it.

Recently Economou, Zdravkovic, & Gilchrist (2007) applied the locus of error paradigm to SLC displays employing a matching method. The method consists in finding the matches of each target on an achromatic Munsell scale. The scale they used ranged from Munsell 2.0 to Munsell 9.5 by means of 16 equally spaced rectangular chips placed from darkest to lightest on a white-black chequered background. Based on the Anchoring theory, the authors predicted a major error for targets on the black background against no or a small error for targets on the white background (Fig. 1). According to their arguments, both classic lateral inhibition models and more recent low level models, such as ODOG (Oriented Differences of Gaussians, Blackeslee & McCourt, 2003), should predict instead different or even opposite results in terms of the magnitudes of the illusion. The results of their experiments largely agree with the predictions based on the Anchoring theory (Figure 2A).

Our intent is not to challenge the findings or the conclusions of Economou *et al.* (2007); rather we want to verify the robustness of the Neutral Value Munsell scale. In fact, the problem is of a methodological

---

[2] The phenomenon is also known as "simultaneous brightness contrast" or just "simultaneous contrast". While the last term is too generic, the first is older and dates back to the days in which many researchers were convinced that simple lateral inhibition models were able to account for the illusion.

nature, but it does not concern the matching method *per se*. Rather, the problem is related to the question that is being asked: *which of the two targets shows the biggest departure from its actual value?* This question requires, in fact, a basic assumption: that the Munsell scale (MS) is capable of showing *actual* lightness values, i.e. *veridical* percepts.

3. Veridicality and the Neutral Value Munsell Scale

Error and *veridicality* are two faces of a same coin. In order to acknowledge that something is wrong, you must assume that there is a truthful or correct answer. According to the locus of error paradigm, the MS provides a way to quantify both the perceptual difference between lightness targets that are physically equal, and also the perceptual difference between a lightness target and its *veridical* lightness value, as displayed on the scale. This last point is critical for theories in lightness perception: if we know what the target should look like, we can find where the error lies, quantify it, and even model hypothetic mechanisms.

At first glance the solution appears ideal. However there is an issue that ought to be addressed: is a Neutral Value Munsell scale actually capable of showing veridical lightness values?

We must confess that we have no idea about what a veridical percept is, how it should look like, and how to distinguish it from a non-veridical percept. In our view, the distinction between veridical and non-veridical is pointless in visual perception. However, from a logical point of view we can speculate that a veridical percept should correspond to some perceptual feature of an object that does not change appearance despite relevant changes in the physical scene.

In lightness literature, it is an established fact that the main issues about lightness constancy concern changes in the level of illumination of a surface target and changes in the luminance pattern surrounding a surface target (background). An achromatic MS is by all means a lightness scale derived by conducting psychophysical experiments (Judd & Nickerson, 1975; Munsell, Sloan, & Godlove, 1933). The idea was to create interval scales to systemize and measure colour perception. However, in order to be able to speak about veridicality, a MS should not be affected by changes either in its level of illumination or in the intensities of its background.

Experiments concerning the robustness of an achromatic MS with uniform illumination modulations have already been carried out (Zavagno & Bressan, in preparation): the scale appears to be quite robust within a certain range of illumination variations.

Here we present the results of an experiment that uses the locus of error paradigm but with three 16 step achromatic Munsell scales that differ only for the backgrounds against which they are seen (Figure 1).

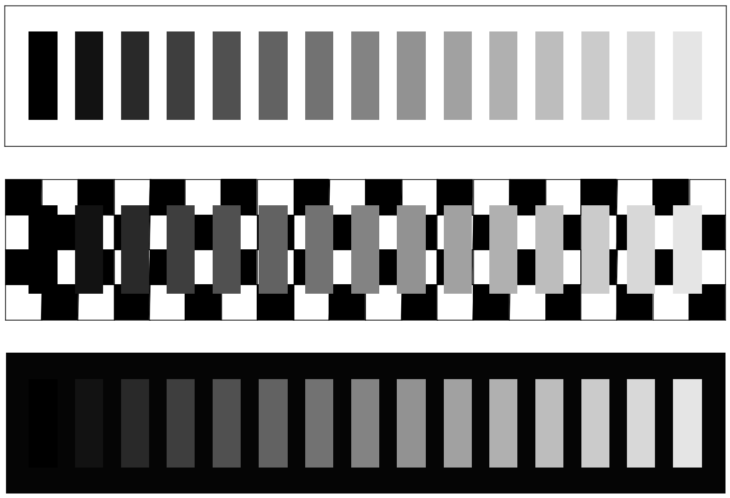

Figure 1. Lightness scales on three different backgrounds.

4. The experiment

Munsell scales are usually employed with backgrounds that are either white-black chequered or completely white. We are aware of only one study that actually employed lightness-matching scales on different backgrounds (Jandò *et al.*, 2003). The aim of such study was to solve the problem of matching increment and decrement targets when a target does not seem to match any of the chips of a lightness scale. Last but not least, such study employed so-called *simulated* Munsell scales presented on a CRT.

In our experiment we used targets and Munsell scales derived from actual Neutral Value Munsell papers. The question that underlies our work is the following: If we change the background against which a MS is seen, will we get the same lightness matches for SLC displays?

To answer such a question we asked 48 observers to carry out lightness matches for a classic SLC display where the targets were two square patches (cm 2.2x2.2) cut from Munsell 6.0 paper, one placed on a white background and the other placed on a black background (cm 20x20). The targets had a luminance of 900 cd/m^2, the background luminances were respectively 2400 and 80 cd/m^2. The observers were assigned to three groups and each group used a MS on a specific background to perform the matching task: a white background, a black background, and a white-black chequered background (Figure 1). The scales ranged from Munsell 2.0 to Munsell 9.5, and consisted of rectangular chips (cm 6 × 2.2) distanced 0.2 cm from each other. MSs were placed 12 cm below the SLC display (Figure 1); the same spotlight illuminated display and scales.

To summarize, we had a within subject variable (target background, TB: black, white) and a between subject variable (Munsell scale background, MSB: white, black, white-black checkers).

Figure 2B shows the mean matches for the three levels of MSB; axes intersect at the actual log reflectance of the targets. An ANOVA for repeated measures, with MSB as between factor, was conducted on the data: both MSB ($F_{2, 45}$=9.52, p<0,0005) and TB ($F_{1,45}$= 96.135, p<0,0001) produced significant effects.

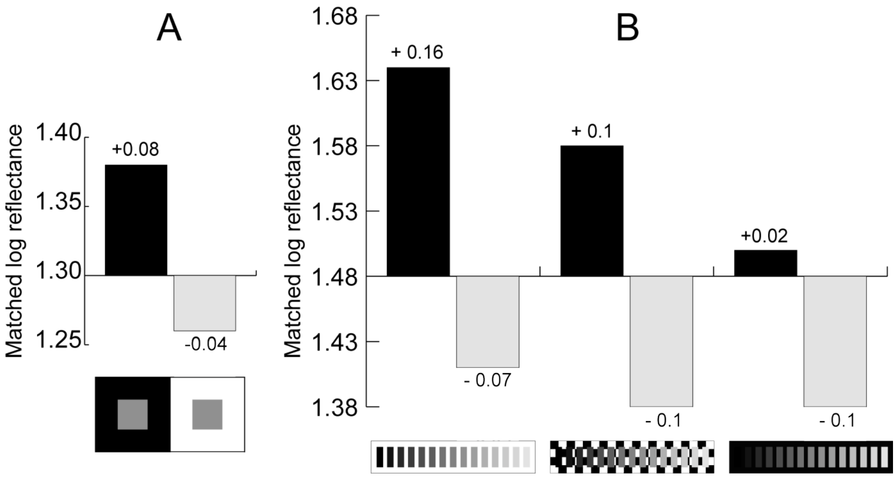

Figure 2. Axes intersect at the actual log reflectance value of the targets. Black and light grey bars stand for targets respectively on the black and on the white background. Numbers above bars indicate the amount of departure from the actual log reflectance value on the Munsell scales. (A) Data extrapolated from Economou, Zdravkovic, & Gilchrist (2007, p. 5, Fig. 3). The graph shows results for the condition “SLC and Munsell scale displayed both on paper”. (B) Results for matches carried out with equal Munsell scales on different backgrounds.

With respect to the target on the black background, unpaired t-tests show that the Munsell scales on the white and on the white-black chequered backgrounds, and Munsell scales on the white and on the black backgrounds, determined matches that are statistically different (respectively: $T_{30}=2.67$, $p=0.012$; $T_{30}=3.84$, $p=0.0006$), while Munsell scales on the white-black chequered and the black backgrounds determined matches that were statistically indistinguishable ($T_{30}=1.09$, $p=0.2$). As for the target on the white background, there is a tendency for a significant difference only between matches carried out with Munsell scales on the white and the black backgrounds ($T_{30}=1.99$, $p=0.055$).

With reference to the question we put forward, the main result of our experiment is that while matches for the target on the white background are relatively stable, those for the target on the black background are highly sensitive to the background against which a MS is seen. Given these results, can we still consider the achromatic MS a reliable instrument capable of showing *veridical* lightness values?

5. Conclusions

Despite our results challenge the Anchoring theory in different ways[3], the aim of this paper is not to question its validity, but to put forward an argument against the widespread use of the concepts of error and veridicality in visual perception sciences.

Our findings, in fact, show that matches performed by using an achromatic MS can change depending on the background against which the scale is seen. In other words, the MS is not a perfect measuring instrument, in the sense that it does not belong to the physical domain: it is by all means a *lightness* scale, i.e. a psychophysical scale that lies entirely inside the phenomenological domain, hence it's appearance can be affected by photo-geometric factors within a given visual scene, just like any other object displaying lightness. Researchers should bear this in mind when they report and discuss their data. It is safe to use matching data to describe the directions of illusions and even to quantify the magnitude of lightness induction between surface targets. However, even in this last case researchers should be aware of the fact that the value of their quantifications are rela-

[3] Just one example: by comparing results reported in Figure 2A-B one can notice that for chequered backgrounds we obtain the same amount of induction for the two targets, while the Anchoring theory predicts a greater induction effect for the target on the black background.

tive to the instrument they used to measure an effect. In fact, our data suggests that lightness scales look less like a ruler and more like a rubber band with the average intensity of the scale's background acting like a stretching the scale either from the black end or the white end.

We nevertheless understand that one might still be tempted to consider the Munsell scale on the white-black chequered background as the only appropriate scale to use. This would be true if we could ultimately say that a white-black chequered background is a neutral background that has no effect whatsoever on the MS. If this were the case, then such scale–background combination would eventually be the only one capable of showing "veridical lightness". This conclusion is however still questionable.

First of all, no background can be considered "neutral", that is having no effect on a target placed on top of it. Neutral displays in visual perception just do not exist. Perception is about ratios and functional relationships among features inside the proximal stimulus; we believe there is very little space left for absolute intensities.

Secondly, such a conclusion leads to an epistemological dead-end: how can we know that the lightness values on the Munsell scale seen against a white-black chequered background are veridical? What makes them more veridical than the values seen against a white or a black background?

If veridicality is a concept devoid of empirical grounds in visual perception, it is obvious that the notion of error is also in trouble. You cannot flip a coin that has only one side to it: such a coin just does not exist in our phenomenal world.

References

Blakeslee B., & McCourt M.E. (2003). A multiscale spatial filtering account of brightness phenomena. In L. Harris & M. Jenkin (Eds.), *Levels of Perception*. New York: Springer, pp. 47-72.

Da Pos O., & Zambianchi E. (1996). *Visual illusions and effects: A collection.* Milano: Guerini & Associati.

Economou E., Zdravkovic S., & Gilchrist A. (2007*).* Anchoring versus spatial filtering accounts of simultaneous lightness contrast*. Journal of Vision*, *7,* 1-15.

Gibson J.J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilchrist A., Kossyfidis C., Agostini T., Li X., Bonato F., Cataliotti J., Spehar B., Annan V., & Economou E. (1999). An anchoring theory of lightness perception. *Psychological review, 106*, 795-834.

Gilchrist A. (2006). *Seeing Black & White*. New York: Oxford University Press.
Gregory R., Harris J., Heard P., & Rose D. (1995). *The artful eye.* Oxford, UK: Oxford University Press.
Jandò G., Agostini T., Galmonte A., & Bruno N. (2003). Measuring surface achromatic color: Toward a common measure for increments and decrements. *Behaviour, Research Methods, Instruments & Computers*, *35,* 70-81.
Judd D.B, & Nickerson D. (1975). Relation between Munsell and Swedish Natural Colour System scales. *Journal of the Optical Society of America*, *65*, 85-90.
Knill D., & Richards W. (1996). *Perception as Bayesian Inference*. Cambridge University Press.
Munsell A.E.O., Sloan L.L., & Godlove I.H. (1933). Neutral value scales. I. Munsell Neutral Value Scale. *Journal of the Optical Society of America*, *23*, 394-411.
Rock I. (1983). *The logic of perception*. Cambrige, MA: MIT Press.

Abstract

This work deals with the concepts of "error" and "veridicality" in visual perception studies by considering the matching paradigm often employed in empirical research to study simultaneous lightness contrast (SLC). Matching paradigms often employ Neutral Value Munsell scales, and there is a strong tendency in the field to consider these scales as a ruler capable of showing veridical lightness values, that is perfect transformations of reflectance values into perceptual values. If this were the case, then Munsell scales should show high constancy to critical changes inside the visual scene. We performed an experiment in which three groups of observers were asked to perform a matching task for a classic SLC display. Each group used the same Munsell scale (MS) but seen against three different backgrounds: white, black, or white-black chequered. Results showed that the background against which the MS is seen heavily influences matches for the target on the black background of the SLC display. Our results support the claim that achromatic Munsell scales are not capable of showing veridical lightness values.

Riassunto

Questo lavoro si occupa dei concetti di "errore" e "veridicità" negli studi sulla percezione visiva considerando il paradigma sperimentale del *matching* spesso utilizzato nelle ricerche sul contrasto simultaneo di bianchezza (CSB). Il paradigma sperimentale del *matching* impiega spesso scale create utilizzando carte Neutral Value Munsell, e vi è una tendenza abbastanza diffusa nel campo a considerare tali scale come un metro in grado di mostrare valori veridici di bianchezza, cioè trasformazioni perfette di valori di riflettenza in valori percettivi. Se questo fosse vero, allora le scale Munsell dovrebbero mostrare un elevato valore di costanza a cambiamenti critici in una scena visiva. Abbiamo condotto un esperimento in cui tre gruppi di osservatori dovevano trovare su una SM il valore di bianchezza corrispondente a quello dei target in una classica configurazione CSB. Ciascun gruppo ha utilizzato la stessa scala ma vista sopra uno dei seguenti sfondi: bianco, nero, a scacchi bianco-neri. I risultati mostrano una forte influenza dello sfondo della SM sulle valutazioni di bianchezza per il target CSB su sfondo nero. I nostri risultati sono di supporto all'affermazione secondo cui una scala Munsell non è in grado di mostrare valori veridici di bianchezza.

Addresses. Olga Daneyko, Department of Psychology "Gaetano Kanizsa", University of Trieste, Via S. Anastasio 12, I-34134 Trieste, Italy.

Daniele Zavagno, Psychology Department, University of Milano-Bicocca, Viale dell'Innovazione 10, I-20126 Milano, Italia.