ON THE PERCEPTION OF ACHROMATIC COLORS: WALLACH’S RATIO PRINCIPLE REVISITED

Jürgen Heller

University of Tübingen

7 October 2014

Running Head: On the perception of achromatic colors

Author Notes: Dr. Jürgen Heller
Department of Psychology
University of Tübingen
Schleichstr. 4
72076 Tübingen, Germany
phone: +49 7071 29-72979, fax: +49 7071 29-3363
email: juergen.heller@uni-tuebingen.de
Abstract

It is commonly assumed that the perception of achromatic colors in center-surround configurations is governed by Wallach’s ratio principle. Within the received view, assuming that perceived achromatic colors form a continuum from black to white, this principle states that the color of the center depends on the center to surround ratio of physical intensities, but not on their absolute values. On the basis of a generalized Fechnerian representation the paper explores the theoretical consequences of various psychophysical invariances: Wallach’s ratio principle and illumination invariance as well as generalizations that may be conceived as near-misses to these conditions. It is shown that these invariances constrain the possible form of the psychophysical functions. Reanalyzing classical data in the light of these theoretical results reveals that only predictions derived from the near-miss to Wallach’s ratio principle and from illumination invariance can account for key features. The implications of these findings are discussed.

Keywords: Ratio principle; near-miss to Weber’s law; illumination invariance; generalized Fechnerian representation; functional equations.
1 Introduction

The perception of achromatic colors has received special interest within research on color perception. Achromatic colors – that is black, white, and all shades of grey – not only form a perceptually distinct subset of colors, in contrast to chromatic colors the characterization of the underlying stimuli seems to be relatively simple. There is a vast literature investigating different percepts, such as brightness or lightness, under a wide variety of conditions ranging from very restricted experimental situations to natural scenes. We will highlight some methodological as well as theoretical aspects of this research that are essential to the subsequent development.

This paper refers to experimental setups using center-surround stimulus configurations which consist of pairs of homogeneous patches, one surrounding the other. The use of center-surround stimulus configurations for studying achromatic color perception dates back at least to Hess & Pretori (1894/1970). They constitute the minimal relational stimuli that allow for evoking color appearances which cannot be produced by single homogeneous patches (Evans, 1964; Mausfeld & Niederée, 1993). Figure 1 (a) shows a display of two center-surround stimuli. Such a side-by-side display is typically presented in an otherwise dark context, and the subject is required to adjust the intensity of one of the patches to match that of the other patch. This procedure is known as asymmetric matching if surrounds differ in intensity. In a variant of the described presentation mode, the left of the pair of center-surround configurations to the left eye only, and the right stimulus to the right eye only. This haploscopic presentation (e.g., Jacobsen & Gilchrist, 1988a, 1988b), keeping an offset between configurations, allows to have different levels of adaptations in both eyes. Another still different haploscopic setup is illustrated in Figure 1 (b). The stimulus configurations form monocular half-images which are presented in such a way that the surrounds are haploscopically superimposed (e.g., Whittle & Challands, 1969). The percept then consists of two patches in a homogeneous surround. It does not come as a surprise that the data obtained with binocular and haploscopic presentation differ substantially. Consider the special case where within each stimulus configuration patch and surround are of identical intensity, but intensities differ across configurations. In a side-by-side display the two patches will most likely look different, while in the haploscopic presentation with superimposed surrounds the percept is a homogeneous surface, and so the two patches look the same. This particular limiting case will be of some importance in the theoretical development exposed below.

The so-called lightness that emerges within center-surround configurations has been investigated in papers which are now considered to be classics of psychology (e.g., Jacobsen & Gilchrist, 1988a; Wallach, 1948), and their results are folklore. Lightness
Figure 1: (a) Simple center-surround stimuli in a side-by-side display. (b) Example of a stimulus configuration for haploscopically superimposing surrounds. Fixation marks aid in fusing the two half-images. The fused binocular image is indicated on the right.

is conceived as the perceived reflectance of the patch, a perceptual property of this stimulus which is independent of the illumination. The most famous observation in this context is attributed to Wallach:

“The colors which come about under these circumstances depend in close approximation on the ratios of the intensities involved and seem independent of the absolute intensity of local stimulation.” (Wallach, 1948, p. 324)

This property is now known as Wallach’s ratio principle. From their data Jacobsen & Gilchrist (1988a, 1988b) conclude that the ratio principle for center-surround stimuli holds over a million-to-one range of intensities. They consider contradicting empirical evidence (e.g. Hess & Pretori, 1894/1970; Jameson & Hurvich, 1961) to be due to experimental artifacts. In the sequel these conclusions are challenged based on a stringent mathematical derivation of the possible forms of the underlying psychophysical functions within a generalized Fechnerian representation, which is known as a subtractive representation (Falmagne, 1985). This will lead to a reappraisal of the empirical evidence provided by Jacobsen & Gilchrist (1988a) which will be presented in Section 2. The subsequently taken route of mathematically formulating a psychophysical theory of lightness differs from the traditional approaches in various respects.

First, the theory explicitly refers to the relevant percepts, and relates them to the empirical observations through a representation for which a measurement-theoretic foundation is available. This is in sharp contrast to the commonly cultivated operationalism, which identifies the lightness of a patch in a given center-surround configuration with the physical stimulus intensity of the matching patch in a white surround. The
theories offered in this context (such as the anchoring theory of lightness of Gilchrist et al., 1999, for instance) consist of plausible, but more or less ad-hoc computations, exclusively formulated on basis of the physical intensities of the experimental stimuli.

Second, the mathematical arguments make explicit all underlying assumptions, which are mostly implicit in the traditional approach, and explores their theoretical implications. In particular, the theory is based on the common belief that perceived achromatic colors form a one-dimensional continuum. Together with the observation that for any surround there is a patch inducing black, white and all sorts of greys in between, this means that asymmetric matches are always possible. Notice that this implication need not hold if perceived achromatic colors have no one-dimensional, but an at least two-dimensional representation.

Third, the proposed theoretical framework allows for investigating empirical laws, such as Wallach’s ratio principle, independent of effects that are due to the presentation order of the stimulus configurations. It is able to take into account what Fechner (1887) called “Einfluss zeitlich-räumlicher Nichtcoincidenz der Reize” [impact of the spatio-temporal non-coincidence of the stimuli]. Dzhafarov (2002) emphasized the importance of these kind of effects, and coined the term observation area to refer to a specific (temporal or spatial) location. Spatio-temporal context effects have been widely neglected in previous research on the perception of achromatic colors.

Developing the formal approach will reveal still another benefit. It will show that Wallach’s ratio principle is closely linked to Weber’s law, and, from a formal perspective, can even be seen to be equivalent to versions of it. This allows for drawing upon already available theoretical results on Weber’s law. In particular, the below introduced terms of a near-miss to Wallach’s ratio principle and to illumination invariance, respectively, are defined in analogy to the well established notion of near-miss to Weber’s law (McGill & Goldberg, 1968). The latter captures the fact that in some empirical contexts the logarithm of the so-called sensitivity function (see Section 5 below) grows linearly with the logarithm of the intensity, just as Weber’s law predicts, but with a slope less than 1 (Falmagne, 1994). The data of Jacobsen & Gilchrist (1988a), that we consider in the subsequent section, shows a similar pattern. A more elaborate discussion of the connections to the near-miss to Weber’s law, which even points to future research, is deferred to Section 5.

2 Data

Jacobsen & Gilchrist (1988a) replicated an experiment by Jameson & Hurvich (1961), but considerably extended the range of illuminations that were used. Figure 2 illu-
trates their stimulus display. Each of the five patches $b$ in surround $t$ presented on the left was to be matched by adjusting patch $a$ in the fixed surround $s$ presented to the right. The configuration to the left was varied by changing illumination over six log units, spanning a million-to-one range of illumination. The configuration was viewed either binocularly, or in (non-overlapping) haploscopic conditions, where each of the two configurations was presented to a single eye only.

Figure 2: Stimulus configuration employed by Jacobsen & Gilchrist (1988a).

Figure 3 (a) replots the data from a haploscopic condition (Jacobsen & Gilchrist, 1988a, Figure 2), which represent the means of the matches obtained for 3 subjects with 4 replications each. It illustrates the dependence of the logarithm of the luminance (in cd/m$^2$) of the matching patch $a$ on the logarithm of the luminance of the referent patch $b$. Data points connected by lines refer to matches obtaind for one of the five patches from the left configuration in Figure 2. With the fixed surround $s$ and the constant ratio $b/t$, Wallach’s ratio principle predicts a constant matching luminance $a$ in these cases.

Jacobsen & Gilchrist (1988a) interpret these data as exhibiting “a very high degree of constancy over the entire range” (p. 1), and they consider the resulting curves to be “so close to horizontal that it seems possible that the residual positive slope could be eliminated by the use of extremely careful procedures” (p. 4). Compressing the scale on the ordinate to match that on the abscissa in their Figure 2 may have facilitated this conclusion. Plotting the data in the plane spanned by log-luminance of the patch versus log-luminance of the surround, as in Figure 3 (b), makes departures from the prediction more pronounced. According to the ratio principle the line segments representing matches should all have unit slope. In particular, the line segments should not cross the diagonal. This, however, is the case in the data of Jacobsen & Gilchrist (1988a), and means that some increments are matched by decrements. The question whether this data pattern reveals systematic deviations from the ratio principle, or may be due to other reasons, will be discussed in more detail below.
3 Formalizing the Received View

This section formalizes the assumptions underlying the traditional approach already sketched above. They constitute what we will refer to as the received view.

3.1 Stimuli

The set of stimuli $S \subseteq I^2$ consists of pairs $(a, s)$, each formed by a patch $a$ in a surround $s$, where $I$ is an open interval $(0, \zeta)$ containing 1. The pair $(\lambda a, \lambda s)$ represents the stimulus $(a, s)$ after changing illuminance by a factor $\lambda > 0$.

Consideration of the following three cases is motivated by empirical results indicating that the visual system distinguishes between incremental and decremental patches. Besides letting the set of stimuli to be the full Cartesian product $I^2$, we consider the subset of incremental stimuli (or, increments) formed by the set $\mathcal{I} = \{(a, s) \mid a, s \in I, a \geq s\}$. The cases $(s, s)$, where the intensity of the patch equals that of the surround, are included for technical reasons. They are of some importance in the subsequent derivations. Additionally, the subset of decremental stimuli (or, decrements) is defined to be the set $\mathcal{D} = \{(a, s) \mid a, s \in I, a \leq s\}$, where equality is included again. In the sequel $S$ can be identified with either $I^2$, $\mathcal{I}$, or $\mathcal{D}$.
3.2 Data and Representation

We refer to situations where two center-surround configurations are presented side-by-side (be it binocularly, or haploscopically) as illustrated in Figure 1. Consider functions $P_{lr}$ and $P_{rl}$ from $S^2$ into the open interval $(0, 1)$, where $P_{lr}(a, s; b, t)$ is interpreted as the probability with which patch $a$ in surround $s$ presented to the left is considered to be more intense as patch $b$ in surround $t$ presented to the right. Consequently, $P_{rl}(a, s; b, t)$ is interpreted as the probability of judging patch $a$ in surround $s$ presented to the right to be more intense than patch $b$ in surround $t$ presented to the left. It is assumed that both $P_{lr}$ and $P_{rl}$ are strictly monotonic increasing in the first and fourth argument, strictly monotonic decreasing in the second and third argument, and continuous in all arguments.

Let us consider $P_{lr}$. Then a subtractive representation assumes the existence of a continuous and strictly increasing function $F$ from the reals into the reals, and of continuous functions $u_l$, $u_r$ from $S$ into the reals, strictly increasing in the first and strictly decreasing in the second argument, such that

$$P_{lr}(a, s; b, t) = F[u_l(a, s) - u_r(b, t)] \quad (1)$$

for all $(a, s), (b, t) \in S$. The functions $u_l$ and $u_r$ constitute the psychophysical functions specifying the perceived color of the patches for the center-surround configurations presented on the left and the right, respectively. In contrast to a Fechnerian representation, where there are functions $F$ and $u$ (with properties as above) such that

$$P_{lr}(a, s; b, t) = F[u(a, s) - u(b, t)] \quad (2)$$

for all $(a, s), (b, t) \in S$, this generalization allows the psychophysical functions to depend on the location where the corresponding stimulus is presented. Sufficient conditions for the subtractive representation are well-known (e.g., Falmagne, 1985, Theorem 7.24).

Within a two alternative forced choice procedure (2AFC) we have

$$P_{lr}(a, s; b, t) + P_{rl}(b, t; a, s) = 1$$

for all $(a, s), (b, t) \in S$. This implies that whenever (1) holds for $P_{lr}$ we obtain a subtractive representation for $P_{rl}$ through

$$P_{rl}(a, s; b, t) = F^*[u_r(a, s) - u_l(b, t)]$$

with $F^*(x) = 1 - F(-x)$ for all $x \in \mathbb{R}$. Any result on $P_{lr}$ thus immediately transfers to $P_{rl}$. This allows to confine consideration to $P_{lr}$ and we drop the index $lr$ in the subsequent sections for convenience.
Before that, however, let us examine the impact of additional constraints that $P_{lr}$ and $P_{rl}$ may meet. In particular, we consider the assumption that responses are independent of the presentation order. It is easily seen that the following symmetry condition (S), requiring that

$$P_{lr}(a, s; b, t) = P_{rl}(a, s; b, t)$$

for all $(a, s), (b, t) \in S$, is equivalent to $P_{lr}$ and $P_{rl}$ both satisfying the so-called balance condition (cf. Falmagne, 1985), i.e.

$$P_{lr}(a, s; b, t) + P_{lr}(b, t; a, s) = P_{rl}(a, s; b, t) + P_{rl}(b, t; a, s) = 1$$

for all $(a, s), (b, t) \in S$. These conditions in turn imply the weak balance condition

$$P_{lr}(a, s; a, s) = P_{rl}(a, s; a, s) = \frac{1}{2}$$

for all $(a, s) \in S$. Within the subtractive representation (1) the weak balance condition forces the functions $u_l$ and $u_r$ to be identical up to an additive constant $k = F^{-1}(\frac{1}{2})$. Given symmetry, we thus arrive at a representation

$$P_{lr}(a, s; b, t) = F'[u(a, s) - u(b, t)]$$

for all $(a, s), (b, t) \in S$, with functions $u(a, s) = u_l(a, s) = u_r(a, s) + k$ for all $(a, s) \in S$ as well as $F'(x) = F(x + k)$ for all real $x$, and $F'(0) = \frac{1}{2}$. Thus, if symmetry (S) holds then a subtractive representation in fact is a Fechnerian representation (2).

4 Results

This section investigates the theoretical consequences of particular empirical laws governing the perception of achromatic colors: Wallach’s ratio principle and illumination invariance as well as generalizations that may be conceived as near-misses to these conditions. It uses functional equation techniques to constrain the possible form of the underlying psychophysical functions. As already indicated above, consideration is confined to $P_{lr}$, and we drop the index in the sequel. Theoretical implications are examined for the subtractive representation (1).

4.1 Wallach’s ratio principle and its near-miss

To formalize the condition, probably the first thing that comes into mind is to require for all real constants $\lambda > 0$

$$P(\lambda a, \lambda s; a, s) = \frac{1}{2}.$$
whenever \((a, s), (\lambda a, \lambda s) \in S\). Here, the stimuli \((\lambda a, \lambda s)\) and \((a, s)\), having the same center to surround ratio of luminances, are presented side-by-side and \((\lambda a, \lambda s)\) is assumed to be judged as more (or less) intense than \((a, s)\) with probability one half (remember that \(P\) stands for \(P_{lr}\)). However, this formalization suffers from a major drawback: The stimuli are presented at different locations (or, observation areas), and thus each of them appears within a different spatial context. In order to avoid confounding with possible effects of the observation areas (may they be perceptual or related to decision processes) the conditions in this section all come in a left and a right version, referring to a single location only. This results in requiring for all real constants \(\lambda > 0\)

\[
P(a, s; b, t) = P(\lambda a, \lambda s; b, t) \quad \text{(RP-1)}
\]
\[
P(a, s; b, t) = p(a, s; \lambda b, \lambda t) \quad \text{(RP-r)}
\]

whenever \((a, s), (b, t), (\lambda a, \lambda s), (\lambda b, \lambda t) \in S\). It is easily seen that these two equations are equivalent to the following property. Consequently, we say that Wallach’s ratio principle holds if for all real constants \(\lambda, \mu > 0\)

\[
P(a, s; b, t) = P(\lambda a, \lambda s; \mu b, \mu t) \quad \text{(RP)}
\]

whenever \((a, s), (b, t), (\lambda a, \lambda s), (\mu b, \mu t) \in S\).

These conditions can be generalized to a situation where there is a real constant \(\rho > 0\) (probably distinct from 1, which characterizes the above treated case), such that for all real constants \(\lambda > 0\)

\[
P(a, s; b, t) = P(\rho^\rho a, \rho^\rho s; b, t) \quad \text{(NMRP-1)}
\]
\[
P(a, s; b, t) = P(a, s; \rho^\rho b, \rho^\rho t) \quad \text{(NMRP-r)}
\]

whenever \((a, s), (b, t), (\rho^\rho a, \rho^\rho s), (\rho^\rho b, \rho^\rho t) \in S\). Integrating these conditions into a single equation again, we say that the near-miss to Wallach’s ratio principle holds if there is a real constant \(\rho > 0\) such that for all real constants \(\lambda, \mu > 0\)

\[
P(a, s; b, t) = P(\rho^\rho a, \rho^\rho s; \rho^\rho b, \rho^\rho t) \quad \text{(NMRP)}
\]

whenever \((a, s), (b, t), (\rho^\rho a, \rho^\rho s), (\rho^\rho b, \rho^\rho t) \in S\).

Notice that, with a slight abuse of language, in this definition we merely require \(\rho > 0\). Strictly speaking, a near-miss to Wallach’s ratio principle is encountered only if \(\rho \neq 1\). However, we prefer to use this lax definition because \(\rho = 1\) (Wallach’s ratio principle) is covered immediately as a special case in the proofs below.

Figure 4 illustrates possible psychometric functions \(P(a, s; b, t)\) for fixed \((b, t)\) according to Wallach’s ratio principle (left) and the near-miss to Wallach’s ratio principle.
with $\rho < 1$ (right). Contour lines of constant probability are indicated on the surfaces and are projected onto the top plane. Within the used log-log coordinates the resulting curves form lines with slope $\rho$. Wallach’s ratio principle (i.e., the special case $\rho = 1$) implies lines of unit slope, while its near-miss with $\rho \neq 1$ generates lines crossing the diagonal. The latter prediction seems to be in line with the matches observed by Jacobsen & Gilchrist (1988a). Notice, however, that the derived properties refer to a single observation area only, while the data of Jacobsen & Gilchrist (1988a), as plotted in Figure 3 (b), do not.

Let us now explore the theoretical consequences of the near-miss to Wallach’s ratio principle within a subtractive representation. Property (NMRP-l) implies that

$$u_l(a, s) = u_l(\lambda^\rho a, \lambda s),$$

and setting $\lambda = \frac{1}{s}$ provides

$$u_l(a, s) = u_l\left(\frac{a}{sp}, 1\right).$$
Since the derivation for $u_r$ referring to the right version (NMRP-r) is analogous, we arrive at the following result.

**Proposition 1.** Assume that a subtractive representation according to (1) exists for $P$. If the near-miss to Wallach’s ratio principle (NMRP) holds for some real constant $\rho > 0$ then the functions $u_l$ and $u_r$ from $S$ into the reals are of the form

$$ u_l(a, s) = h_l \left( \frac{a}{s^\rho} \right) $$

$$ u_r(a, s) = h_r \left( \frac{a}{s^\rho} \right) $$

for all $(a, s) \in S$ and some strictly increasing and continuous functions $h_l$ and $h_r$.

For predicting the data of Jacobsen & Gilchrist (1988a) based on this result, it is assumed that $(a, s)$ matches $(b, t)$ if and only if $P(a, s; b, t) = \frac{1}{2}$. A match is then represented by

$$ h_l \left( \frac{a}{s^\rho} \right) = h_r \left( \frac{b}{t^\rho} \right) + k $$

with $k = F^{-1} \left( \frac{1}{2} \right)$. Concrete predictions are derived from the equation

$$ \log \left( \frac{a}{s^\rho} - \alpha \right) = \log \left( \frac{b}{t^\rho} - \alpha \right) + \kappa $$

with real constants $\rho$, $\alpha$, and $\kappa$, where the latter lumps together $k$ and a potential constant difference between $h_l$ and $h_r$. Figure 5 plots the matches predicted for parameter settings $\rho = 0.96$, $\alpha = 0.08$ and $\kappa = -1.3$. The diagrams are to be compared to the data in Figure 3. To facilitate this comparison, data points from Figure 3 (a) are replotted in light gray in Figure 5 (a).

The parameter $\rho$ turns out to be slightly but clearly distinct from 1, and causes the positive slope of the curves in Figure 5 (a), which is specified by $1 - \rho$. With changing the parameter $\kappa$ the whole bunch of curves is mainly shifted up or down, while the effect of parameter $\alpha$ is best seen in Figure 5 (b). It (and thus the form of the functions $h_l$ and $h_r$) is responsible for the kind of compression that the matches in the fixed surround exhibit. If $\alpha$ were zero then the matches would span the same range as the patches $b$ of the reference stimuli at each of the different illumination levels (see Figure 7). While the curves in Figure 5 (a) predict the observed range of matches very well, the spacing of curves within this range shows clear deviations from the data. Other choices of functions $h_l$, $h_r$ did not improve the fit significantly.

### 4.2 Illumination invariance and its near-miss

Illumination invariance assumes that perceived relations remain the same under changes of illumination, i.e., when all intensities are changed by a factor $\lambda > 0$. Within
Figure 5: Prediction of the data of Jacobsen & Gilchrist (1988a) based on the solutions derived in Proposition 1. (a) Curves represent the prediction of the log-luminance of the matching patch $a$ versus the log-luminance of the referent patch $b$. Data points are plotted in light gray. (b) Points connected by line segments represent predicted matches in the log patch versus log surround plane. See text for details.

In the considered context we say that illumination invariance holds if for all real constants $\lambda > 0$

$$P(a, s; b, t) = P(\lambda a, \lambda s; \lambda b, \lambda t)$$  

(II)

whenever $(a, s), (b, t), (\lambda a, \lambda s), (\lambda b, \lambda t) \in S$. Generalizing this condition the near-miss to illumination invariance requires that there is some real constant $\rho > 0$ such that for all real constants $\lambda > 0$

$$P(a, s; b, t) = P(\lambda^\rho a, \lambda^\rho s; \lambda^\rho b, \lambda^\rho t)$$  

(NMII)

whenever $(a, s), (b, t), (\lambda^\rho a, \lambda^\rho s), (\lambda^\rho b, \lambda^\rho t) \in S$.

Notice that the same remark as above, concerning the slight abuse of language in the definition of the near-miss, applies here too. Notice also that the near-miss to illumination invariance as formulated by (NMII) is a special case of the near-miss to Wallach’s ratio principle (NMRP), which follows by setting $\lambda = \mu$. This means that the constraints set by (NMII) are less restrictive, and additional functional forms may crop up. The subsequent result shows that this is indeed the case.

**Proposition 2.** Assume that a subtractive representation according to (1) exists for $P$ and that the near-miss to illumination invariance (NMII) holds for some real constant $\rho > 0$. 


Then the functions \( u_l \) and \( u_r \) are of one of the following forms:

\[
\begin{align*}
    u_l(a, s) &= h_l \left( \frac{a}{s^\rho} \right) \quad \text{(i)} \\
    u_r(a, s) &= h_r \left( \frac{a}{s^\rho} \right)
\end{align*}
\]

(in particular if \( \rho = 1 \) and \( u_l(a, a) \), \( u_r(a, a) \) constant for all \( a \in I \)), or

\[
\begin{align*}
    u_l(a, s) &= h_l \left( \frac{a}{s^\rho} \right) + \beta \log s \\
    u_r(a, s) &= h_r \left( \frac{a}{s^\rho} \right) + \beta \log s
\end{align*}
\]

for all \((a, s) \in S\), some strictly increasing and continuous functions \( h_l, h_r \), and some nonzero real constant \( \beta \).

**Proof.** From (1) and (NMII) we get

\[
    u_l(\lambda^\rho a, \lambda s) - u_l(a, s) = u_r(\lambda^\rho b, \lambda t) - u_r(b, t).
\]

for all \((a, s), (b, t) \in S\) and \( \lambda > 0 \) such that \((\lambda^\rho a, \lambda s), (\lambda^\rho b, \lambda t) \in S\). Fixing \((b, t)\) the righthand side depends on \( \lambda \) only. Thus, we obtain

\[
    u_l(\lambda^\rho a, \lambda s) = u_l(a, s) + f(\lambda)
\]

with some continuous function \( f \) (independent of \((b, t)\)).

This implies that \( f \) satisfies the logarithmic Cauchy equation

\[
    f(\lambda \mu) = f(\lambda) + f(\mu)
\]

for all real numbers \( \lambda, \mu > 0 \).

If \( f \) is constant then \( f(\lambda) = 0 \) for all \( \lambda > 0 \), which, in particular, is the case for \( \rho = 1 \) and \( u_l(a, a) \) constant for all \( a \in I \). Setting \( \lambda = \frac{1}{s} \) in (3) with \( f \) vanishing provides

\[
    u_l(a, s) = u_l \left( \frac{a}{s^\rho}, 1 \right) = h_l \left( \frac{a}{s^\rho} \right),
\]

a solution of the form (i).

If \( f \) is not constant then the only continuous solution of (4) is

\[
    f(\lambda) = \beta_l \log \lambda
\]

for all \( \lambda > 0 \) and some nonzero real constant \( \beta_l \). Again setting \( \lambda = \frac{1}{s} \) in (3) and plugging in (5) provides

\[
    u_l(a, s) = h_l \left( \frac{a}{s^\rho} \right) + \beta_l \log s
\]
with \( h_l(x) = u_l(x, 1) \). The line of reasoning for constraining the form of \( u_r \) is completely analogous and yields the type (i) solution

\[
u_r(a, s) = h_r \left( \frac{a}{s^\rho} \right)
\]

and

\[
u_r(a, s) = h_r \left( \frac{a}{s^\rho} \right) + \beta_r \log s \tag{7}
\]

with \( h_r(x) = u_r(x, 1) \) and some nonzero real constant \( \beta_r \). Finally, substituting functions (6) and (7) into representation (1) shows that (NMII) holds only if \((\beta_l - \beta_r) \log \lambda = 0\). So, the constraint \( \beta_l = \beta_r = \beta \) results, and we obtain the solutions of type (ii).

Several remarks apply to Proposition 2. First, it is closely related to a result of Falmagne & Iverson (1979). They started from a so-called simple scale representation requiring that

\[
P(a, s; b, t) = F[u(a, s), u(b, t)]
\]

for all \((a, s), (b, t) \in S\). On the one hand this representation is more restrictive than (1), because it considers only a single function \( u \). On the other hand it is more general in the way how \( u(a, s) \) and \( u(b, t) \) are combined. Moreover, Falmagne & Iverson (1979) treat the conjoint Weber law, which is formally equivalent to illumination invariance as introduced in (II). The proof of Proposition 2 is quite distinct from that of the related result by Falmagne & Iverson (1979). This is partly due to the differences in the assumptions, and partly attributable to the fact that some of their arguments do not survive the generalizations needed for treating the near-miss to illumination invariance (NMII).

Second, Proposition 2 is also related to results of Heller (2006), who considered illumination invariance for an algebraic representation based on Plateau's midgray operation (Plateau, 1872; see also Falmagne, 1985, and Heller, 2001) rather than a probabilistic representation as in the present paper. Moreover, there is no account of spatial context effects. However, the resulting functional forms include the type (i) and type (ii) solutions together with still an additional form. This one is similar to the type (ii) solutions, with the additive term depending on the logarithm of the surround being replaced by multiplication with the surround intensity raised to some power.

Third, requiring that \( u_l(a, a) \) and \( u_r(a, a) \) are both constant for all \( a \in I \), a particular case mentioned for (II) (i.e., \( \rho = 1 \)) in the proposition, amounts to the existence of a neutral point (Niederée, 1998). This means that the patches in any two configurations \((a, a)\) and \((b, b)\) look the same. As already mentioned above, however, this will occur if haploscopic presentations with superimposed surrounds are used (as in Whittle & Challands, 1969, for example), but not for the experimental setup of Jacobsen &
Gilchrist (1988a). Under illumination invariance (II) the type (i) solutions thus only arise if there is a neutral point, so that the subsequent consideration can be confined to the type (ii) solutions.

Figure 6: Prediction of the data of Jacobsen & Gilchrist (1988a) based on the type (ii) solutions derived in Proposition 2 with $\rho = 1$. See Figure 5 for a description and text for details.

Figure 6 shows the matches predicted under illumination invariance (II) for the type (ii) solutions. Predictions are derived from the equation

$$\log \left( \frac{a_s - \alpha}{s} \right) + \beta \log s = \log \left( \frac{b_t - \alpha}{t} \right) + \beta \log t + \kappa$$

with parameter settings $\beta = 0.04$, $\alpha = 0.06$ and $\kappa = -1.3$. The resulting curves are highly similar to those illustrated in Figure 5 for the near-miss to Wallach’s ratio principle, and $1 - \beta$ equals the parameter $\rho$ used there.

4.3 Joint Validity of II and NMRP

The previous sections provided evidence that important aspects of the data of Jacobsen & Gilchrist (1988a) can be explained via either illumination invariance or near-miss to Wallach’s ratio principle (with an exponent clearly less than one). So, the question arises whether there exist functions $u_l$ and $u_r$ in a subtractive representation that satisfy both conditions simultaneously. Contrary to first intuition, the following proposition shows that this holds true. Moreover, it proofs that the functions $u_l$ and $u_r$ are uniquely determined up to the choice of some parameters.
Proposition 3. Assume that a subtractive representation according to (1) exists for \( P \) and that illumination invariance (NMII) and the near-miss to Wallach’s ratio principle (NMRP) for some real constant \( 0 < \rho \neq 1 \) both hold.

Then the functions \( u_l \) and \( u_r \) are of the form

\[
\begin{align*}
  u_l(a, s) &= \beta \frac{1}{1-\rho} \left[ \log a - \rho \log s \right] + \alpha_l, \\
  u_r(a, s) &= \beta \frac{1}{1-\rho} \left[ \log a - \rho \log s \right] + \alpha_r.
\end{align*}
\]

for all \( (a, s) \in S \) and some real constants \( \alpha_l, \alpha_r \) and nonzero \( \beta \). In particular, we have \( \beta > 0 \) if \( \rho < 1 \) and \( \beta < 0 \) if \( \rho > 1 \).

Proof. For the type (i) solutions of Proposition 2 satisfying (II) we obtain

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_r \left( \frac{b}{t^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) - h_r^* \left( \frac{b}{t} \right).
\end{align*}
\]

Fixing \( b, t \) and the ratio \( \frac{a}{s} \) this provides

\[
\begin{align*}
  h_l \left( \frac{a}{s^{1-\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) - h_r^* \left( \frac{b}{t} \right) + h_r \left( \frac{b}{t^{\rho}} \right).
\end{align*}
\]

Because of \( \rho \neq 1 \) this implies that \( h_l \) is constant, which contradicts the assumption that \( h_l \) is strictly increasing. Thus, there are no solutions in this case.

For the type (ii) solutions of Proposition 2 satisfying (II) we get

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_r \left( \frac{b}{t^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) + \beta \log s \left( \frac{b}{t} \right) - h_r^* \left( \frac{b}{t} \right) - \beta \log t,
\end{align*}
\]

and rearranging terms provides

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_l^* \left( \frac{a}{s} \right) - \beta \log s &= h_r \left( \frac{b}{t^{\rho}} \right) - h_r^* \left( \frac{b}{t} \right) - \beta \log t.
\end{align*}
\]

Fixing \( b, t \) lets the righthand side be constant, say \( c \), so that selecting \( s = 1 \) yields

\[
\begin{align*}
  h_l(a) - h_l^*(a) &= c.
\end{align*}
\]

This leads to

\[
\begin{align*}
  h_l^* \left( \frac{a}{s^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) + \beta \log s.
\end{align*}
\]

Defining \( x = \frac{a}{s}, \ y = s^{1-\rho} \) and thus \( s = y^{1/(1-\rho)} \) we obtain

\[
\begin{align*}
  h_l^* \left( x \cdot y \right) &= h_l^* \left( x \right) + \frac{\beta}{1-\rho} \log y,
\end{align*}
\]

for all \( (a, s) \in S \) and some real constants \( \alpha_l, \alpha_r \) and nonzero \( \beta \). In particular, we have \( \beta > 0 \) if \( \rho < 1 \) and \( \beta < 0 \) if \( \rho > 1 \).

Proof. For the type (i) solutions of Proposition 2 satisfying (II) we obtain

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_r \left( \frac{b}{t^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) - h_r^* \left( \frac{b}{t} \right).
\end{align*}
\]

Fixing \( b, t \) and the ratio \( \frac{a}{s} \) this provides

\[
\begin{align*}
  h_l \left( \frac{a}{s^{1-\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) - h_r^* \left( \frac{b}{t} \right) + h_r \left( \frac{b}{t^{\rho}} \right).
\end{align*}
\]

Because of \( \rho \neq 1 \) this implies that \( h_l \) is constant, which contradicts the assumption that \( h_l \) is strictly increasing. Thus, there are no solutions in this case.

For the type (ii) solutions of Proposition 2 satisfying (II) we get

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_r \left( \frac{b}{t^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) + \beta \log s \left( \frac{b}{t} \right) - h_r^* \left( \frac{b}{t} \right) - \beta \log t,
\end{align*}
\]

and rearranging terms provides

\[
\begin{align*}
  h_l \left( \frac{a}{s^{\rho}} \right) - h_l^* \left( \frac{a}{s} \right) - \beta \log s &= h_r \left( \frac{b}{t^{\rho}} \right) - h_r^* \left( \frac{b}{t} \right) - \beta \log t.
\end{align*}
\]

Fixing \( b, t \) lets the righthand side be constant, say \( c \), so that selecting \( s = 1 \) yields

\[
\begin{align*}
  h_l(a) - h_l^*(a) &= c.
\end{align*}
\]

This leads to

\[
\begin{align*}
  h_l^* \left( \frac{a}{s^{\rho}} \right) &= h_l^* \left( \frac{a}{s} \right) + \beta \log s.
\end{align*}
\]

Defining \( x = \frac{a}{s}, \ y = s^{1-\rho} \) and thus \( s = y^{1/(1-\rho)} \) we obtain

\[
\begin{align*}
  h_l^* \left( x \cdot y \right) &= h_l^* \left( x \right) + \frac{\beta}{1-\rho} \log y,
\end{align*}
\]

for all \( (a, s) \in S \) and some real constants \( \alpha_l, \alpha_r \) and nonzero \( \beta \). In particular, we have \( \beta > 0 \) if \( \rho < 1 \) and \( \beta < 0 \) if \( \rho > 1 \).
which is a Pexider equation of the form

\[ f_1(x \cdot y) = f_2(x) + f_3(y) \]

with continuous functions \( f_1, f_2, f_3 \), and \( x \in \mathbb{R}^+ \) and \( y \in (0, \zeta^{1-\rho}) \). Its only continuous solution is

\[
\begin{align*}
    f_1(x) &= \alpha_1 \log x + \alpha_2 + \alpha_3, \\
    f_2(x) &= \alpha_1 \log x + \alpha_2, \\
    f_3(x) &= \alpha_1 \log x + \alpha_3,
\end{align*}
\]

with real constants \( \alpha_1, \alpha_2, \alpha_3 \), and \( \alpha_1 > 0 \) since \( f_1, f_2 \) are strictly increasing. Obviously, we have \( \alpha_1 = \frac{\beta}{1-\rho} \) and \( \alpha_3 = 0 \), so that

\[
\begin{align*}
    h_l^*(x) &= \frac{\beta}{1-\rho} \log x + \alpha_2, \\
    h_l(x) &= \frac{\beta}{1-\rho} \log x + \alpha_2 + c.
\end{align*}
\]

This finally gives

\[
u_l(a, s) = \frac{\beta}{1-\rho} [\log a - \rho \log s] + \alpha_l.\]

with nonzero real constants \( \alpha_l, \beta, \) satisfying \( \beta > 0 \) if \( \rho < 1 \) and \( \beta < 0 \) if \( \rho > 1 \). The argument for \( u_r \) is essentially the same.

Figure 7 shows the matches that are predicted on basis of Proposition 3. Predictions are computed through the equation

\[ \log a - \rho \log s = \log b - \rho \log t + \frac{1-\rho}{\beta} \kappa \]

with parameters \( \rho = 0.96, \beta = 0.04, \) and \( \kappa = -0.85 \) (lumping together \( \alpha_l, \alpha_r \) and \( k = F^{-1}(\frac{1}{2}) \)). Notice, however, that the last two parameters are not identifiable. There is a trade-off between \( \beta \) and \( \kappa \) because only their quotient is constrained. The main difference to the previously considered cases lies in the fact that Proposition 3 does not leave much flexibility for modeling the interaction between center and surround. Assuming both (II) and (NMRP) does not allow for capturing the observed compression of matches well enough.

5 Discussion

The previous sections provide a formal analysis of the traditional approach to the perception of achromatic colors. The main features of what we have called the received
Figure 7: Prediction of the data of Jacobsen & Gilchrist (1988a) based on the solutions derived in Proposition 3. See Figure 5 for a description and text for details.

view consist of conceptualizing the space of achromatic colors as a one-dimensional continuum, and of collecting data nearly exclusively by means of asymmetric matches. On this background, Jacobsen & Gilchrist (1988a) claim that the perception of achromatic colors in simple center-surround configurations is governed by Wallach’s ratio principle over a wide range of illuminations. The ratio principle, attributed to Wallach (1948), states that the color of the center depends on the center to surround ratio of physical intensities, and is independent of the absolute values. This interpretation of the data of Jacobsen & Gilchrist (1988a), which was derived by a more or less qualitative inspection, is challenged by fitting to them explicit, mathematically formulated theories. The developed formalization is based on a probabilistic theory that generalizes a Fechnerian representation, and is known as a subtractive representation (Falmagne, 1985). Within this framework, the paper explores the theoretical consequences of various psychophysical invariances, such as Wallach’s ratio principle and illumination invariance, as well as generalizations that may be conceived as near-misses to these conditions. It is shown that these invariances constrain the possible form of the psychophysical functions. Fitting these functions to the data of Jacobsen & Gilchrist (1988a) lead to a reappraisal of the empirical evidence.

As already mentioned above, the route followed here offers various advantages over the traditional approach. The theory explicitly refers to the relevant percepts and avoids the widely endorsed operationalism. It makes explicit all underlying assumptions, and allows for isolating the effects of the psychophysical invariances from spatio-
temporal context effects, may they be perceptual or related to decision processes. For a side-by-side presentation of simple center-surround configurations this is achieved by considering $P_{tl}(a, s; b, t)$, the probability of judging patch $a$ in surround $s$ presented to the left to be more intense than patch $b$ in surround $t$ presented to the right (and $P_{rt}$, respectively). Within the subtractive representation (1) perceptual aspects then are represented by the possibly different location-specific psychophysical functions $u_l$ and $u_r$, while decisional aspects may be represented by the function $F$.

Reanalyzing the Jacobsen & Gilchrist (1988a) data within this framework indicates that the empirical evidence favors the near-miss to Wallach’s ratio principle (NMRP) with an exponent $\rho = 0.96$. The theory is able to predict the overall shape of the curves and the observed range of matches, but there is room for improvement with respect to the spacing of the different curves within this range (Figure 5). Still another new perspective on this set of data emerges from showing that an alternative explanation based on illumination invariance (II) leads to nearly identical predictions (Figure 6). The basic assumption here is that the probabilities in the 2AFC task remain constant under changes of illumination (i.e., when multiplying all intensities by the same factor). The type (ii) solutions in Proposition 2 make clear that there is no complete discount of the surround as in Wallach’s ratio principle, even if $\rho = 1$ holds. This mechanism can account for the nonzero slope of the curves in Figure 3 (a). Combining both assumptions (NMRP) and (II) leads to a complete specification of the form of the psychophysical functions. This form seems to be too restrictive, however, falling short, especially in predicting the observed compression of the range of matches.

Future work may extend the present study in various directions: concerning the experimental setup as well as regarding the theory. It is a particular strength of the presented framework that a direct experimental test of the critical properties is possible. Collecting data that allow for estimating the probabilities in (RP), (NMRP), (II), or (NMII) can provide stronger evidence for or against the theory as performing a post-hoc analysis of asymmetric matches. Observing systematic empirical deviations may even suggest more general conditions. Moreover, all analyses should be based on individual rather than pooled data, since there can be substantial interindividual variability (e.g., Heller, 2001). Corresponding experiments are underway. Another important issue is to avoid the ambiguities inherent in the presentation of center-surround configurations in complete darkness, which does not allow for distinguishing changes of the illumination from changes of the stimuli. These problems can be resolved by using more complex displays as they result from embedding center-surround configurations into Mondrian-like stimuli, consisting of irregular patterns of rectangular patches that span a whole range of achromatic colors from black to white. With this kind of setup, Arend &
Spehar (1993a) found that lightness matches tend to produce curves close to horizontal, but, very consistently over subjects, with a slight negative slope. This is in line with a near-miss to Wallach’s ratio principle (NMRP) with $\rho$ slightly greater than 1, or with illumination invariance (II) and $\beta < 0$. The authors also report a constant offset from settings that would be predicted by a simple ratio rule of intensities. This represents an effect of the observation areas that the theory proposed here can account for. Using simple center-surround configurations, Arend & Spehar (1993b) essentially replicated the pattern showing up in the data of Jacobsen & Gilchrist (1988a) as illustrated in Figure 3.

Theoretical extensions can exploit the analogy to particular formulations of Weber’s law. There are formal equivalences between illumination invariance (II) and Conjoint Weber Law, and between Wallach’s ratio principle (RP) and a version of the Strong Conjoint Weber Law (cf. Falmagne & Iverson, 1979). The corresponding line of research suggests to change the perspective from the probabilities to the so-called sensitivity function which, in the present context, may be defined by

$$\xi_{b,t,p}(s) = a \text{ iff } P(a, s; b, t) = p$$

for $(a, s), (b, t) \in S$. The sensitivity function $\xi_{b,t,p}(s)$ specifies the patch $a$, which in surround $s$ is judged to be more intense than patch $b$ in surround $t$ with probability $p$. With this notion at hand, we may generalize the near-miss to Wallach’s ratio principle (left version) to

$$\xi_{b,t,p}(\lambda s) = \lambda^{\rho(p)} \cdot \xi_{b,t,p}(s)$$

for all $\lambda > 0$. Conceptualizing $\rho$ as a function of the discrimination criterion $p$ leads to a more flexible theory. There is considerable theoretical work on constraining the possible form of this function (Augustin, 2008; Doble, Falmagne, & Berg, 2003; Falmagne, 1985, 1994; Iverson, 2006). Building upon these results Doble, Falmagne, Berg, & Southworth (2006) provide evidence that predictions are in line with data on loudness of pure tones, and that parameters show a systematic pattern of covariation.

To wrap up, the present paper demonstrates the benefits that arise from formulating explicit mathematical theories of the perception of achromatic colors. It exemplifies that one can get a clearer picture of the underlying perceptual mechanisms and the decisional processes involved. As it constitutes nothing but a first step into this direction, further experimental and theoretical work is needed, which may proceed along the lines outlined above.
References


Plateau. J. (1872). Sur la mesure des sensations physiques, et sur la loi qui lie l’intensité de la cause excitante [On the measurement of physical sensations and on the law which links the intensity of these sensations to the intensity of the source]. *Bulletins de L’Academie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique, 2me Série, 33*, 376-388.
