The effect of mean luminance change and grating pedestals on contrast perception: Model simulations suggest a common, retinal, origin

Markku Kilpeläinen*, Lauri Nurminen[2] & Kristian Donner[3]

1) Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland 2) Brain Research Unit, O.V. Lounasmaa Laboratory, School of Science, Aalto University, Espoo, Finland 3) Department of Biosciences, University of Helsinki
*email: markku.kilpelainen@helsinki.fi

Abstract

The percept of a contrast target is substantially affected by co-occurring changes in mean luminance or underlying ("pedestal") contrast elements. These two types of modulatory effects have traditionally been studied as separate phenomena. However, regardless of different higher-level mechanisms, both classes of phenomena will necessarily also depend on shared mechanisms in the first stages of vision, starting with the primary responses of photoreceptors. Here we present model simulations showing that important aspects of both classes may be explained by the temporal dynamics of photoreceptor responses read by integrate-and-fire operators. The model is physiologically justified and all its parameters are constrained by experimental evidence. Although there remains plenty of room for additional mechanisms to shape the exact quantitative realization of the perceptual functions in different situations, we suggest that signature features may be inherited from primary retinal signaling.

Keywords: adaptation; dipper function; psychophysics; cone photoreceptors; retinal ganglion cells; contrast discrimination; modeling

1. Introduction

The neural and perceptual responses to a contrast pattern are always modulated by the spatiotemporal context in which the pattern appears. Two of the most basic "context" attributes are the mean luminance in the relevant part of the visual field and the underlying contrast structure at the very location of the pattern. These two have traditionally been considered as mechanistically separate in psychophysical studies. The purpose of the modeling work presented here was to clarify to what extent two much-studied phenomena, representing each of these two lines of investigation, may be accounted for by shared retinal mechanisms. We find that major features of both the effect of a change in mean luminance and the pedestal effect can be explained by the temporal dynamics and thresholding of photoreceptor signals.

The first phenomenon that we shall consider is how an abrupt change in mean luminance over a wide field affects the perceived contrast of a restricted target (Geisler, 1978; Kilpeläinen, Nurminen & Donner, 2011; Poot, Snippe & van Hateren, 1997). Although the visual system adjusts its sensitivity to the prevailing mean luminance quite efficiently (Rieke & Rudd, 2009; Shapley & Enroth-Cugell, 1984), such adjustment is not instantaneous. As a result, both the contrast response of retinal neurons and psychophysically measured contrast perception are altered.

The second phenomenon is the well known "pedestal effect", where the threshold for detecting a target stimulus is lowered when a low-contrast pedestal stimulus is simultaneously presented under the target, but rises again when pedestal contrast is increased further (resulting in the so-called "dipper function"). The pedestal is usually but not necessarily spatially similar to the target (Goris, Zaenen & Wagemans, 2008). The effect has been studied intensively since it was first described by Nachmias and Sansbury (1974).

In this study, we model responses associated with these two types of stimuli. The model comprises a grid of cone photoreceptors, their corresponding read-out mechanisms, and a simple decision rule. The model units closely replicate experimentally established response parameters of retinal cones plus a simple integrate-and-fire neuron ("ganglion cell"). The model simulations reproduce psychophysical data on the effect of luminance change and on the pedestal effect remarkably well, especially considering that none of the parameters used in the model were optimized to fit the data.

2. Methods and results

2.1 The general architecture of the model

The architecture of the model is very simple. The stimulus is presented to a grid of "cone photoreceptors". For the sake of simplicity, we assume that one cone corresponds to one stimulus pixel (although optically unrealistic, this simplification in the spatial domain has no significant consequences for our modeling). The response of each cone to a luminance step caused by its stimulating pixel is calculated according to the phenomenological "independent activation" model suggested by Baylor, Hodgkin and Lamb (1974) to describe cone responses in turtle retina (equations 1 and 2):

$$\frac{dU}{dt} = \frac{B \times U_L - U(a-(a-1)\times B)}{\tau_L} \qquad (1),$$

where

$$B = \frac{S \times I \times (1 - e^{\frac{-t}{\tau}})^n}{S \times I \times (1 - e^{\frac{-t}{\tau}})^n + a^{-1} \times U_L} \qquad (2),$$

where U may be photovoltage or photocurrent, τ is a time constant related to the time to peak of the response as $\tau = t_p/ln(n)$, I is light intensity, and t is time. The meaning of the other symbols is given in Table 1. Sensitivity S depends on the adapting intensity (background illumination) I_B and was described by the empirical function $S = S_D \times (|d|+0.37bI_B) / ((|d|+0.37I_B) \times (0.37cI_B+1))$ derived from recordings in macaque cones by Dunn, Lankheet & Rieke (2007). The values of all parameters were taken from electrophysiological experiments reported in the literature (Table 1). No free parameters were used, save for a spike threshold in the pedestal modeling tuned to match the absolute sensitivity of a single model cone to the absolute sensitivity of the human subjects, cf. Fig. 4. Most of the parameters stem from photocurrent measurements in macaque cones. Although photovoltage is the visually relevant signal, the difference is likely to be negligible in the early rising phase of responses. The parameters t_p and τ describing response kinetics and b, c and d describing background adaptation, however, are derived from photovoltage recordings in intact retinas, since photocurrent recordings from isolated, superfused photoreceptors tend to indicate too slow response shut-off (Baylor et al., 1983; Schnapf et al., 1990; 1999) and too feeble light-adaptation (Donner et al., 1990). Compared with monkey photovoltage, electroretinogram (ERG) recordings of mass cone responses from the intact human eye have indicated even faster response kinetics (Friedburg, Thomas & Lamb, 2001), but ERG with ganzfeld stimulation is likely to be dominated by peripheral cones, which are faster than the foveal cones relevant here.

Table 1. Parameters used in the cone response model

Parameter	Value	Source
S_D (sensitivity in darkness)	33 x 10^{-3} pA/phot	2)
a (conductance ratio)	1.64	1)
U_L (max amplitude)	19 pA	2)
n (number of filters)	7	1, 3)
t_p (time to peak in moderate light)	32 ms	4, 5)
τ_L (time constant under bright background light)	10 ms	1, 6)
b, c, d (sensitivity scaling)	1.3, 0.00029,100	5)

1) Baylor et al., 1974; 2) Schnapf et al., 1990; 3) Hood & Birch, 1993; 4) Schneeweis & Schnapf, 1999; 5) Dunn et al., 2007; 6) Djupsund et al. 1996

Temporal noise with the amplitude (15 % of maximum light response) and spectral distribution (11 ± 3 Hz) of voltage noise recorded in macaque cones was added to the cone responses (Schneeweis & Schnapf, 1999). No other noise was added, since cone noise is the dominant noise component of the retinal output in cone vision (Ala-Laurila, Greschner, Chichilnisky & Rieke, 2011).

For each cone photoreceptor, the model includes a read-out mechanism, which transforms the graded signal of the cone into spike frequencies and applies a thresholding non-linearity to the signal. The read-out mechanism is a simple leaky integrator, whose membrane voltage (in mV) is determined by equation 3.

$$\frac{dV}{dt} = \frac{1}{\tau_m}(-V + R(U - Inh)) \qquad (3),$$

where τ_m is the membrane time constant (2.7 ms, Weber & Harman, 2005), U is the input current response in pA, R is input resistance (here normalized to 1 GΩ, to simplify the numerical implementation) and Inh is a subtractive static inhibition applied to all cone responses. $Inh = (U \times i) / (i + |U|)$, where i = 3.95 pA. The inhibition component reduces spontaneous firing by a factor of 2.6, in agreement with physiological measurements (Brivanlou, Warland & Meister, 1998). The model produces a spike if $V > r(t_s)$ where r is a function approximating threshold changes during the absolute and relative refractory periods (Trong & Rieke, 2008; Uzzell & Chichilnisky, 2004), described by an inverted cumulative Gaussian (mean = 4.5 ms, SD = 1 ms), where spike threshold decreases from 20 mV (just higher than the highest input signal) to the absolute threshold value (0.13 mV in Case 1 and 0.195 in Case 2) as a function of time since last spike (t_s). We take no stand on the specific neural correlates of the read-out mechanism, but suggest that its behavior resembles that of parasol / Y-type retinal ganglion cells.

In the following, we apply the model to two cases where contrast perception is modulated in a non-linear manner by the context in which the contrast stimulus is presented. In both cases we will first explain the mechanism that we propose to cause the altered contrast percepts, and then compare the model simulation results to corresponding psychophysical data.

2.2 Case 1. The luminance step effect

In a previous study, we observed that an abrupt change in mean luminance attenuates the perceived contrast of a simultaneously presented grating stimulus (Kilpeläinen et al., 2011). The subject's task in the experiment was to match the contrast of a grating presented simultaneously with an upward step in mean luminance (from 185 to 1295 Td) to the contrast of a similar grating presented after a period of adaptation to the higher mean luminance (for full methods, see Kilpeläinen et al., 2011). At the lowest grating contrasts, the absolute amount of attenuation increased to some extent with increasing contrast, but over a wide range of higher grating contrasts, perceived contrast was attenuated by a constant, contrast-independent term (subtractive attenuation). Thus, the attenuation caused by the luminance change became, in relative terms, *weaker with increasing target contrast*.

Fig. 1 summarizes the stimuli and responses in this experimental paradigm as they appear in our model. In the first situation (the "step-up" situation), the model cones are initially adapted to a low photopic luminance level of 185 Td. Then, with the abrupt change in mean luminance (to 1295 Td) and the simultaneous onset of the grating stimulus, each cone is exposed to a new light level which is considerably above the level of adaptation (Fig 1a, bottom). In the second situation, (the "steady-state" situation), the cones are adapted to the higher level (1295 Td). Upon the onset of the grating the cones are exposed to new light levels that depart from the adaptation level only by the amounts

corresponding to the grating contrast (downwards as well as upwards) (Fig. 1A, middle).

Since the simplest definition of contrast, and indeed the simplest signal an observer could plausibly use to determine contrast, is modulation amplitude divided by mean, we simplify the model in this case to only 2 units: one responding to the peak light level (L_{max}) and one to the mean light level (L_{mean}) of the grating. A unit responding to the minimum light level of the grating could, of course, also be of interest, but response families of retinal cones to light decrements accurate enough to provide us with useful response parameters are not currently available.

The cone responses evoked by L_{max} and L_{mean} are illustrated in Fig. 1B, denoted R_{max} and R_{mean}, respectively. By looking at the two response pairs, one might conclude that, due to the well-known compressive relationship between light intensities and cone peak response amplitudes (Baylor et al., 1974; Schnapf et al., 1990), the difference between the responses (and the integral of the difference) would always be smaller in the step-up situation. However, the earliest rise of photoreceptor responses is actually linear against stimulus intensity, and thus the integral of the very earliest segments of R_{max} - R_{mean} will be approximately the same regardless of adaptation level (Baylor et al., 1974; Lamb & Pugh, 1992). (The approximation holds in a moderate luminance range such as here, where the sustained response to the higher luminance does not significantly decrease the dynamic range of the cone, limited from above by the saturation level). This is illustrated in Fig. 1C, which shows integrals of R_{max} - R_{mean} in the step-up situation (blue lines) and the steady-state situation (red lines). When the stimulus contrast is low (the pair of solid curves), the integrals first overlap, but diverge around the time of the crossing of the first spike criterion. The integral in the step-up situation reaches subsequent spike criteria later than that in the steady-state case, thus producing a lower spike frequency. With a higher stimulus contrast (the dashed pair of curves), the two curves overlap nearly up to the fourth spike criterion. Consequently, spike frequencies up to that point will differ only a little between the step-up and the steady-state situation. Although the instantaneous integration of R_{max} - R_{mean} may seem physiologically unrealistic, it is worth noting that parasol/Y-type ganglion cells of the mammalian retina are able to encode *both* light flux linearly summed over the receptive-field centre *and* contrast with finer spatial grain, as the centre itself encompasses smaller non-linear subunits (Enroth-Cugell & Robson, 1966; Crook, Peterson, Packer, Robinson, Troy & Dacey, 2008; Demb, Zaghloul, Haarsma & Sterling, 2001). Importantly, the subunits appear to rely purely on feedforward processing without the delays involved in classical "surround" antagonism.

To summarize, the spike frequency signal for contrast in the step-up situation becomes relatively less suppressed with increasing stimulus contrast, because higher contrast elicits more steeply rising cone responses, whereby the integrating read-out mechanism reaches its threshold criteria faster and consequently uses less and less of the compressed parts of the cone responses. Figure 2A shows simulated spike frequencies (means of responses with 200 different noise samples) as functions of grating contrast for the step-up situation (blue line) and the steady-state situation (red line). The spike responses were produced by using R_{max}-R_{mean} as input to equation 3. The spike frequencies (*SF*) were then calculated as illustrated in Fig. 1C as $(t_x\text{-}t_1)$ / (x-1), i.e. the time between the first and the *x*th spike divided by the number of intervals between them. Here we used x = 3 for the number of spikes carrying the contrast information, but as shown in Fig. 2C, the qualitative results are very similar for x = 4. However, x must be small in order to preserve the central role of the early part of photoreceptor responses.

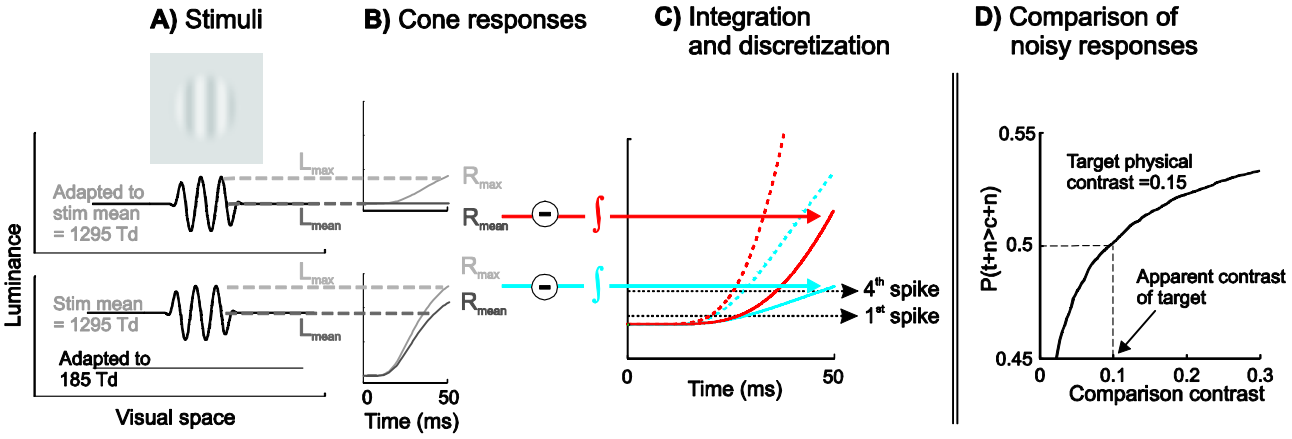

Fig. 1. The logic of the model from stimuli to the comparison of spike responses. A) An illustration of the grating stimulus (top), and the spatial luminance profiles of the stimulus in the steady-state (middle) and the step-up situation (bottom). B) The leading edges of cone responses to light pulses that have the peak (R_{max}) and mean (R_{mean}) intensity of the grating. C) The read-out mechanism integrates R_{max}-R_{mean} for the steady-state stimulus (adapted to 1295 Td, red lines) and the step-up stimulus (adapted to 185 Td, blue lines) and fires a spike every time the integral crosses a fixed criterion (the 1-spike and the 4-spike criteria exemplified by black dashed arrows). Solid curves for low target contrast, dashed curves for high target contrast. For illustrative purposes, responses are here shown without noise. D) The contrast of the comparison stimulus required for a match at a given physical contrast of the target is determined by comparing noisy target responses (step-up responses) against noisy comparison responses (steady-state responses) and finding the comparison contrast at which the comparison response is larger 50 % of the times. Similar results could be obtained with the common method of integrating the ROC-curve (Green & Swets, 1988), but the current approach was adopted for maximal transparency.

The strongly transient nature of parasol/Y-type ganglion cell responses (Lee, Pokorny, Smith & Kremers, 1994; Nirenberg & Meister, 1997; cf. Ludwig, Gilchrist, McSorley & Baddeley, 2005) was approximated by including only spikes from the first 50 ms in the spike rate calculation. Qualitatively, Fig. 2 A immediately shows both that there is suppression (the spike rate for the step-up case is always lower than that for the steady-state case), and that relative suppression decreases with increasing grating contrast. However, a comparison of mean rates is not representative of what really happens in a 2 alternative forced choice experiment involving pair-wise comparisons of single, noisy responses. To simulate this, we calculated for all pairs of target stimuli (i.e., contrast gratings presented together with the step in mean luminance) and comparison stimuli (i.e., contrast gratings presented with steady mean luminance) with added noise the probabilities that the SF produced by the *target+noise* be greater than the SF produced by the *comparison+noise*. The probabilities were calculated according to equation 4.

$$\frac{\sum_{k=1}^{m}\sum_{j=1}^{m}(t+n_j > c+n_k)}{m^2} \qquad (4),$$

where t is the target, c is the comparison stimulus, n is noise and m is the number of noise samples (we used m = 200). The expression $(t + n_j > c + n_k)$ takes the value 1 if the statement in the parentheses is true and 0 if it is false. Observe that since each noise sample is used both with target and comparison stimulus, the number of possible comparisons is m^2. For any given target contrast, the comparison contrast at which the probability is 0.5 (corresponding to random responding by a subject in a 2 alternative task) is the simulated matching contrast of the target. This principle is illustrated in Fig. 1D. The relative suppression due to the luminance step can then be calculated for the results of the simulations in the same way as for results of psychophysical experiments: $(C_{physical} - C_{matching}) / C_{physical}$).

Figure 2B plots relative suppression in the step-up condition according to the model simulations (continuous curve) together with the pertinent psychophysical data from three subjects (Kilpeläinen et al., 2011). The data is the same as in Fig. 1A of Kilpeläinen et al. (2011) with the addition of one new data point for the lowest grating contrast for each subject. Interestingly, some of the inter-subject differences might be explained by differences in the number of spikes used (cf. Fig. 2 C) and/or the spike threshold voltage. Such differences might occur e.g. as a consequence of differences in overall levels of intrinsic noise between subjects. Here, the spike threshold voltage was set to 0.13, so that the average spike frequency in response to noise alone was approximately 20 Hz, corresponding to spontaneous activity of macaque parasol cells at photopic light levels (Croner, Purpura & Kaplan, 1993; Trong & Rieke, 2008).

2.3 Case 2. The pedestal effect

One of the most intensely studied phenomena related to context-dependent contrast perception is the so-called pedestal effect, which refers to the fact that the threshold for

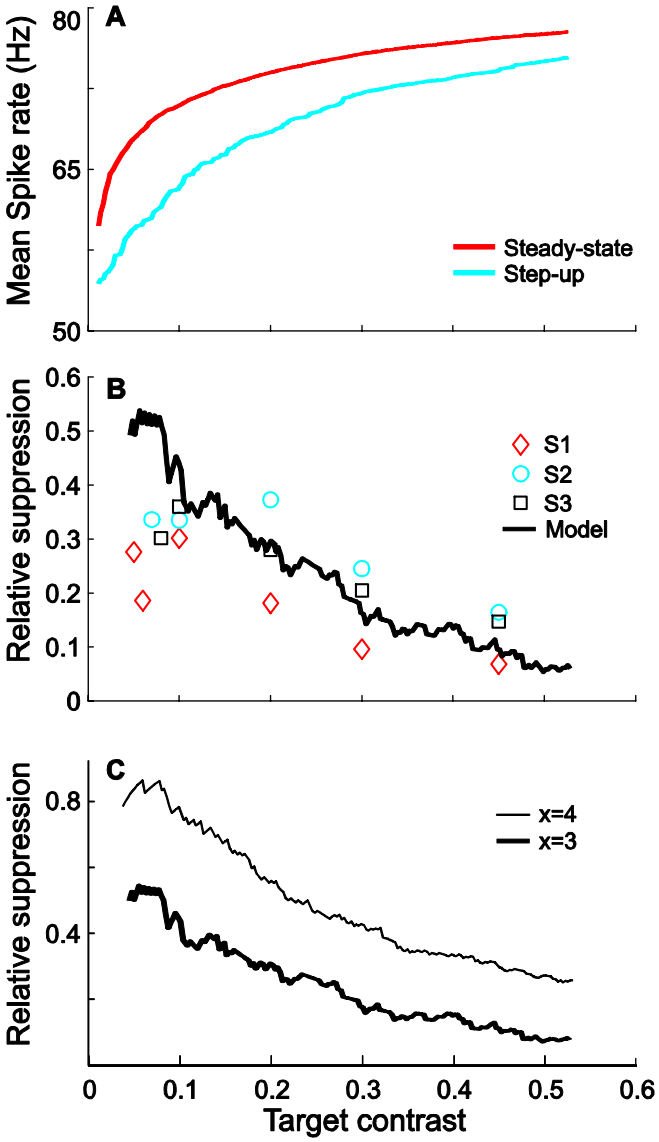

Fig. 2 A) Spike frequencies produced by model simulations as functions of grating contrast for the step-up situation (blue line) and the steady-state situation (red line). B) The continuous curve is the simulated relative suppression caused by the luminance step. The data points are derived from Fig.1A of Kilpeläinen et al. (2011), re-plotted as *relative* suppression of perceived contrast. The symbols for the three subjects are the same as in the original figure. C) Comparison of simulated relative suppression when using x=3 (thick line) of 4 (thin line).

discriminating a contrast stimulus added to a low baseline contrast (the pedestal) can be lower than the absolute *detection* threshold for the same stimulus. From a certain point, increasing the baseline contrast further causes the discrimination threshold to rise above the detection threshold, creating the characteristic threshold vs. pedestal contrast function commonly known as the "dipper function".

We investigated whether the pedestal effect could, in principle, also arise from the simple retinal mechanisms described above. The model parameters were kept the same as in case 1, but 32 "cone" units were now used. The rationale is that when an observer decides in which of two stimulus locations or intervals the target grating is, it is not necessary or efficient to determine the difference between peak and mean luminance. Instead, the photoreceptor responses to *any one pixel* of the two gratings have to differ sufficiently for the spike frequency in response to

pedestal+target+noise to be significantly larger than the spike frequency for *pedestal+noise*. In modeling, the 32 model units were deployed to cover a quarter cycle of a sinusoidal grating, from peak to mean. In most experiments, grating stimuli produced on a monitor have much fewer than 32 pixels per quarter cycle, but due to retinal point spread (Navarro, Artal & Williams, 1993; Westheimer & Campbell, 1962), the cone mosaic is actually receiving a smooth continuum of light intensities.

Discrimination thresholds were simulated in the same way as the matching contrast in Case 1. Spike frequencies were calculated from the first three spikes (x=3). Then the model calculates for all values of pedestal and target contrast the probability that the SF produced by *pedestal + target + noise* is greater than the SF produced by *pedestal + noise*. Thus, substituting the relevant stimuli into equation 4 we get equation 5.

$$\frac{\sum_{k=1}^{m}\sum_{j=1}^{m}(p+t+n_j > p+n_k)}{m^2} \qquad (5),$$

where p is pedestal, t is target, n is noise and m is the number of noise samples (again we used m = 200). The simulation applies the same principles as that described under Case 1 (above), repeating the procedure for every combination of pedestal and target. Figure 3 illustrates the route from comparison of noisy responses to the threshold-versus-pedestal-contrast function. First, the probabilities are calculated in the manner presented, separately for different pedestal contrasts (Fig. 3A). Then the threshold contrast for any desired criterion level can be determined as the target contrast where the probablity reaches the criterion (75 % in this example). Figure 3B plots the target contrasts at which the probability reaches the 75 % threshold criterion as a function of pedestal contrast. This can be done for any threshold criterion and for each of the 32 model units. Figure 3C shows functions for threshold criterion 75 % for 5 different model units. It can be seen that the units processing pixels further removed from the peak of the contrast pattern reach the lowest discrimination threshold at higher pedestal contrasts and are thus most useful at these contrasts. We emphasize that discrimination is here implied to benefit from the pixels close to mean luminance only at high contrasts, where responses to those pixels are just crossing the absolute threshold. A trend in favour of this idea is in the data of Kingdom and Whittle (1996), where discrimination thresholds relative to the absolute threshold are mostly higher for square wave gratings than sine wave gratings, although there is much more of the peak (less of the pixels close to mean) in the square wave gratings.

Figures 3B and 3C clearly show that the model does produce the general characteristics of dipper functions. When the target is presented on top of a low-contrast pedestal, the discrimination threshold initially falls below the absolute detection threshold.

An intuitive understanding would be that the pedestal "helps" the target response to exceed the spike threshold. As pedestal contrast is raised further, the discrimination threshold starts rising again, primarily because of the decelerating intensity-response function generated by the model thresholding mechanism (cf. Donner 1989), but also due to the refractoriness of the integrate-and-fire neuron.

Given the discrimination data from many different model units, it would be a trivial exercise to devise rules that combine the unit responses in a way that would allow the model to predict almost any data quite well. We shall refrain from developing a number of (arbitrary) rules and only consider the arguably simplest (possibly unrealistically simple) principle: the discrimination threshold in each situation is equal to the lowest threshold of any model unit. For example, in Fig. 3C, the unit that is the least sensitive one at zero pedestal would be the one that determines the discrimination threshold at pedestal contrasts higher than 0.07 (the curve reaching its minimum around pedestal contrast 0.13).

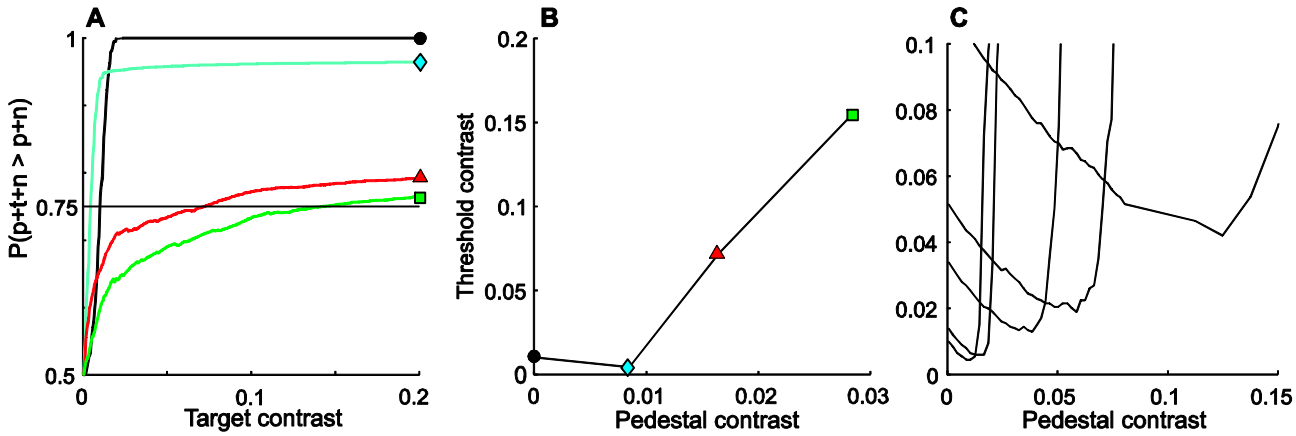

Fig. 3. A) The probability that the *SF* for pedestal + target is greater than the *SF* for pedestal alone as a function of target contrast, simulated for one model unit. The black curve is for no pedestal, the blue, red and green curves for pedestal contrasts 0.009, 0.017, 0.025, respectively. B) Threshold contrast for a criterion of 75% correct as a function of pedestal contrast, extracted from the curves in A as the target contrast where the corresponding curve crosses the threshold criterion (the horizontal line at 0.75 in A). Colours of markers correspond to colours of lines in A. C) Threshold contrasts for criteria of 75% correct in different model units, based on simulations similar to those in A. Each curve represents one of the model units (only 5 model units included to avoid clutter). Unit 1 (leftmost curve) corresponds to the peak of the sine wave; the other units are nos. 16, 26, 28, and 30.

In Fig. 4, we compare predictions based on this "most sensitive unit" rule with psychophysical data of Henning and Wichmann (2007). Panel A presents the psychophysical data and the model simulation for the 75% threshold criterion. Here, the spike threshold was increased to 0.195 mV to match the mean of the psychophysical subjects' detection thresholds (with pedestal contrast 0). Even with this change, the model still operates with the early rise of the photoreceptor signals (times of the 3 spikes relative to "stimulus onset" with a 10 % contrast (in ms) are -20, 1, 15 in case 1 and 4, 19, 27 in case 2, negative spike times caused by noise just before stimulus onset). While the model curve does not closely reproduce any of the three data sets, the shape of the curve is quite similar. Rather than reproducing any single data set as such, the emphasis here is in grasping the larger patterns of the data as whole. Such aspects of model performance are illustrated by Fig. 4B, where discrimination thresholds for three different % correct criteria (60%, 75% and 90%), shown separately for each of the three subjects, are compared with corresponding model curves (here, all absolute thresholds have been normalized to unity). The model correctly predicts two salient features of the data: lowering the criterion deepens the dip trough and moves it to higher pedestal contrast. Also the slopes of the subsequent rising parts of the functions are also generally well reproduced. The main shortcoming is that the model produces a somewhat too shallow dip in the 60% criterion function.

3. Discussion

The first steps of vision are shared by all visual input and must therefore constrain any visual function. The common stages include at least phototransduction and the formation and early transmission of photoreceptor signals. It is a common assumption in current psychophysical literature that these stages provide an approximately linear representation of natural stimuli and can therefore be of little help in the analysis of perceptual phenomena currently in the focus of vision research. We think, on the contrary, that the modelling of phenomena observed in psychophysics and cortical electrophysiology would generally benefit from building on a solid foundation of known retinal physiology. In this study we have shown that central properties of two fundamental nonlinearities in the processing of contrast patterns, the effects of abrupt changes of mean luminance and the pedestal effect, may be inherited from early retinal signalling.

3.1 The performance of the model: Luminance change

The fact that an abrupt increase in mean luminance attenuates the perceived contrast of a simultaneously presented grating is not surprising as such. There is a compressive relationship between light intensity and cone response amplitudes (see Schnapf et al., 1990), and the step from 185 to 1295 Td is likely to transiently consume most of the cones' dynamic range. From such strong compression, however, one would expect attenuation to be stronger than the higher the grating contrast. What is observed is quite the opposite: roughly subtractive attenuation, implying that relative attenuation *decreases* substantially with increasing grating contrast (Kilpeläinen et al., 2011). The present model predicts this counter-intuitive result correctly, capturing the overall pattern of the data correctly (Fig. 2),

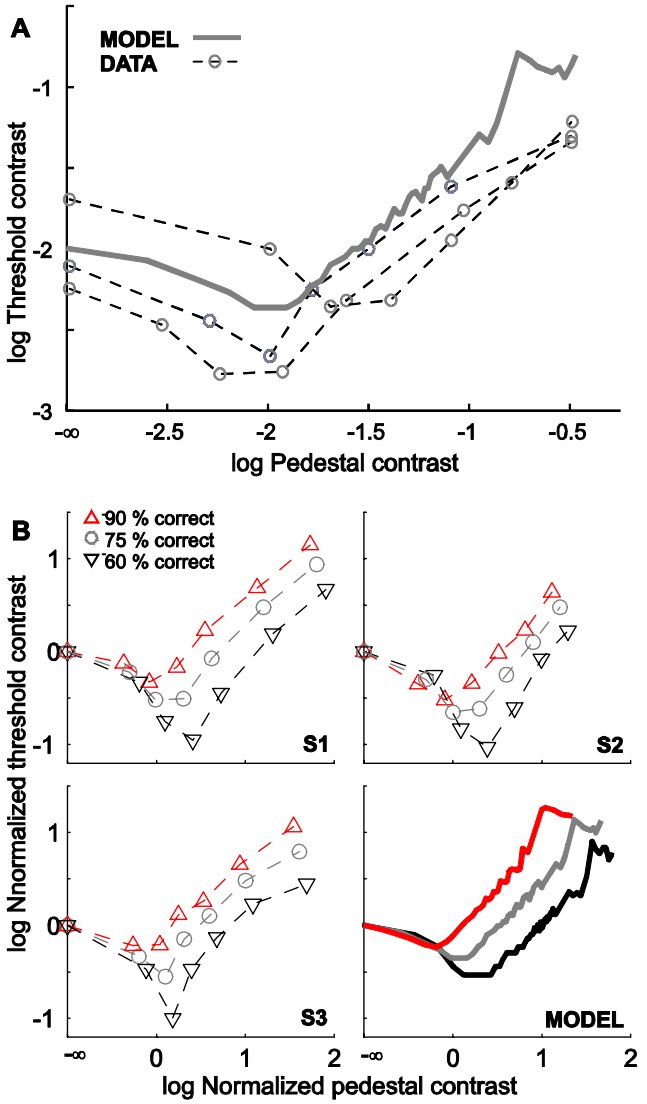

Fig. 4. The pedestal effect. A) Comparison of model simulations and data for three subjects at 75 % correct threshold criterion (data from Henning and Wichmann, 2007). B) Comparison of simulations and data at three different % correct criteria (see legend) on normalized axes. Each of the panels S1-S3 shows data for one subject; the fourth panel (bottom right) gives the model curves for the same levels of % correct. The data markers as well as the order of subjects are as given by Henning & Wichmann (2007). Note: the authors do not give information on observer pupil sizes. Here, retinal illuminance has been estimated by assuming an average pupil diameter of 5.6 mm at the mean luminance used in the experiments, 50 cd/m^2 (cf. Winn, Whitaker, Elliott & Phillips, 1994).

with no free parameters. Admittedly, a certain amount of quantitative inter-subject variation remains unexplained. However, although analysis of differences between subjects might be interesting, it would require individual estimation of sensitivity and noise beyond the scope of the present study.

Why does the present model give such a different result compared with a straightforward correlation of photoreceptor intensity-response data with psychophysics? The main reason is that it largely ignores the final (peak) amplitudes of the cone responses. Instead, it considers the early parts of the cone signal, where the relation between response amplitude and light intensity is closer to linear. With increasing contrast, the "thresholding point" moves

successively closer to the earliest rise of the photoresponse (which is the most linear part). This is certainly functionally realistic. It would be inefficient for the system to wait for the final amplitude of responses before making decisions (the noisy photoresponses would, in fact, need to be analyzed beyond the peak to determine their final amplitudes in the first place). This idea is supported by psychophysical studies suggesting linear summation of visual responses to the onset of even very high contrasts or light intensities (Alpern, Rushton & Torii, 1970; Vassilev, Mihaylova & Bonnet, 2002), which is best understood as resulting from the linearity of the early rising part of photoreceptor responses (Donner, 1989; Donner & Fagerholm, 2003). Functions relating reaction time (latency) and perceived contrast to stimulus contrast at different adapting luminances are also consistent with modelling based on this idea (Djupsund, Fyhrquist, Hariyama & Donner, 1996; Donner, 1989). When transmitted through the retina, signal components derived from the early rise of photoreceptor responses, including initial spike frequencies of ganglion cells, are comparatively robust against differentiation, surround antagonism, or lateral inhibition, which may strongly modulate or suppress later parts (Donner, 1981; Donner, 1989). In addition, the earliest rising phase of photoreceptor responses is little affected by adapting luminances that do not cause sustained responses that significantly decrease the dynamic range of the cells (Friedburg et al., 2001; Heikkinen et al., 2009). This partial invulnerability of the early photoreceptor signal should probably be taken into consideration in the characterization of the adaptation mechanisms that restore the steady-state responsivity of the system (e.g., Hayhoe, Benimoff & Hood, 1987; Hayhoe, Levin & Koshel, 1992; von Wiegand, Hood & Graham, 1995)

The plausible neural code of retinal ganglion cells has recently been elegantly explored by Jacobs et al. (2009). Two of their conclusions are relevant to the present study. Firstly, they found that spike count is an implausible neural code. This is in line with our emphasis on the initial spike rate. For example, our luminance step simulation with spike count coding would produce nearly complete suppression, regardless of target contrast. Secondly, they found that a spike rate code would also fall somewhat short of satisfactory performance and suggested that a more complex spike time correlation code might be necessary. It would be interesting to see whether the "initial spike rate code" we present here would prove to be biologically plausible in such an analysis.

3.2 The performance of the model: Pedestal effect

The pedestal data is more complex than the data on luminance change, consisting of entire threshold-versus-contrast functions measured for three different threshold criteria (Henning & Wichmann, 2007). In view of the complexity of the data, our simple model with parameters strongly constrained by physiology performs quite well. It reproduces qualitatively several salient effects in different conditions, e.g., the relative magnitudes and locations of the dips as well as the shapes and the relative locations of the entire functions, although inter-individual differences remain unexplained. One weakness of our "pedestal" simulation is that we needed to raise the spike threshold of the leaky integrator from what was used in the "luminance step" simulation (where the threshold was based strictly on the spontaneous firing rate of macaque ganglion cells). To match the absolute thresholds in the pedestal experiments (cf. Fig. 4), we may have assigned to our retinal spike threshold effects of some mechanisms that actually operate on post-retinal levels. For example, neurons in macaque thalamus are known to transmit spikes with probability increasing with increasing spike rate (Carandini, Horton & Sincich, 2007).

3.3 Relation to other models of the pedestal effect

The psychophysical dipper function has most often been treated as an expression of an underlying sigmoid-shaped contrast-response function (Legge & Foley, 1980). According to this description, the initial dip in thresholds reflects the initial acceleration of the contrast-response function (when going from zero contrast to low contrasts) and the subsequent rise in thresholds correspondingly reflects deceleration of the function at higher contrasts. This general framework has produced much fine-tuning of models with excitatory and inhibitory, linear and non-linear operators and clever experimental studies measuring interactions between pedestals and other contextual elements (Chen & Tyler, 2001; Chen & Tyler, 2008; Foley, 1994). The idea has since been complicated by the notion that the amount of internal noise will also have an effect on thresholds, and that this noise may be signal dependent (Geisler & Albrecht, 1997; Georgeson & Meese, 2006, see also, Meese & Baker, 2011).

The sigmoidal contrast-response function is a phenomenological construct, i.e., without strong claims about underlying mechanisms. Thus there is no substantive contradiction with our model. It may be noted that a thresholding mechanism reading the early rise of cone responses will not quite produce a sigmoid contrast-response function (cf. Donner, 1989), but on the other hand there is no data with sufficient resolution to judge between such subtle shape differences. Our threshold assumption is of course originally derived from ganglion cell physiology (see e.g., Barlow & Levick, 1969), but is also in agreement with the conclusion of Kontsevich and Tyler (1999) that, besides a non-linear transducer function, a hard threshold is a factor in the pedestal effect.

Approaches that differ qualitatively from the ones presented above include the idea that the dip is caused by a reduction in observer uncertainty (Klein & Levi, 2009; Pelli, 1985) or by off-frequency looking, i.e., involvement of units tuned to spatial frequencies other than the nominal spatial frequency of the target grating (Goris, Wichmann & Henning, 2009, see also Chirimuuta & Tolhurst, 2005).

Our approach does not categorically exclude any of these theories. For example, if spike frequency responses are more consistent and less variable when the stimulus is presented on a pedestal (Fig. 3), this may be seen as a specification of "decreased observer uncertainty" (Pelli, 1985). Similarly, the possibility of off-frequency looking (Goris et al., 2009) is inherent in the idea of units that read the "cone" responses at different parts of the grating and a discrimination decision built on combination of these unit responses.

Indeed, in view of the extremely different situations where "dipper functions" may arise (for a review, see Solomon, 2009), it may be wisest to accept that almost any reasonable mechanism that has been suggested may in some situation surface as dominant. The modelling presented here operates on a different level, however, dealing with

physiological preconditions "below" the alternative theories discussed above. We have shown that basic retinal mechanisms, which undoubtedly are at work, can generate functions similar to those observed in psychophysical experiments. A claim that the proposed mechanism is irrelevant would need to be supported by an explanation of where and how the retinal effect *is lost* and a similar function *recreated* by other means at a higher level.

3.4 Links between the two cases

Although the modelling of the two phenomena is explicitly based on the same mechanisms and parameters, it may be useful to point out how the main features operate in the two cases. Firstly, the increasing importance of the earliest, near-linear part of the photoreceptor signal with increasing contrast is essential for both. When applied to this signal, the thresholding read-out mechanism produces intensity-response functions of the type illustrated in Fig. 2A: monotonically increasing, non-saturating, but decelerating (cf. Donner 1989). In Case 1, this allows the grating response in the step-up situation to "catch up" with that in the steady-state situation as grating contrast increases. In Case 2, this determines, for a single model unit, the rise of the discrimination threshold with increasing pedestal contrast after the dip (cf. Mansfield 1976). The post-receptoral thresholding mechanism in itself is of course essential. For both cases, it is the mechanism that implements the move towards ever earlier parts of the cone signal with increasing contrast. For Case 2 specifically, it is the mechanism chiefly responsible for the dip in the dipper function.

3.5 Conclusions

We investigate how far certain perceptual phenomena may be explained by retinal mechanisms that are necessarily present, using simple modelling tightly constrained by physiological evidence and with a minimum of arbitrary elements. We show that one such model is quite successful in accounting for central features of two basic context modulation effects in contrast perception: contrast attenuation by an abrupt change in mean luminance, and the pedestal effect. By increasing *mechanistic* understanding, this type of model may be useful in at least three ways. Firstly, it suggests new electrophysiological experiments. For example, do spiking responses of (some) monkey ganglion cells to stimuli similar to those considered here carry components conforming to the model predictions? Secondly, it helps to dissect relevant psychophysical phenomena into retinal and cortical components. There is a large body of psychophysical studies using paradigms similar to those considered here which may be amenable to analysis on similar lines, e.g., studies involving various forms of overlay of contrast patterns. Thirdly, it suggests perceptually relevant properties of retinal circuitry that need to be implemented in retinal models designed as tools for such analyses (see eg., Wohrer & Kornprobst, 2009).

Acknowledgements

This study was funded by the Finnish National Programme of Psychology (MK), the Finnish Graduate School of Neuroscience (LN) and the Academy of Finland (grant 128081 for KD and 13464 for LN). The authors thank Dr. Petri Ala-Laurila for valuable comments on the manuscript.

NOTICE: this is the author's version of a work that was accepted for publication in Vision Research. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in VISION RESEARCH, [VOL#, ISSUE#, (DATE)] DOI: 10.1016/j.visres.2012.02.002

References

Ala-Laurila, P., Greschner, M., Chichilnisky, E. J., & Rieke, F. (2011). Cone photoreceptor contributions to noise and correlations in the retinal output. *Nature Neuroscience 14*, 1309–1316.

Alpern, M., Rushton, W.A.H., & Torii, S. (1970). Size of rod signals. *Journal of Physiology-London, 206*, 193-208.

Barlow, H.B., & Levick, W.R. (1969). Changes in the maintained discharge with adaptation level in the cat retina. *Journal of Physiology-London, 202*, 699-718.

Baylor, D.A., Hodgkin, A.L., & Lamb, T.D. (1974). Electrical response of turtle cones to flashes and steps of light. *Journal of Physiology-London, 242*, 685-727.

Brivanlou, I. H., Warland, D. K., & Meister, M. (1998). Mechanisms of concerted firing among retinal ganglion cells. *Neuron, 20*, 527-539.

Carandini, M., Horton, J.C., & Sincich, L.C. (2007). Thalamic filtering of retinal spike trains by postsynaptic summation. *Journal of Vision, 7*.

Chen, C.C., & Tyler, C.W. (2001). Lateral sensitivity modulation explains the flanker effect in contrast discrimination. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*, 509-516.

Chen, C.C., & Tyler, C.W. (2008). Excitatory and inhibitory interaction fields of flankers revealed by contrast-masking functions. *Journal of Vision, 8.*

Chirimuuta, M., & Tolhurst, D. J. (2005). Does a bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research, 45*, 2943-2959.

Croner, L.J., Purpura, K., & Kaplan, E. (1993). Response variability in retinal ganglion-cells of primates. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 8128-8130.

Crook, J.D., Peterson, B.B., Packer, O.S., Robinson, F.R., Troy, J.B., & Dacey, D.M. (2008). Y-cell receptive field and collicular projection of parasol ganglion cells in macaque monkey retina. *Journal of Neuroscience, 28*, 11277-11291.

Demb, J.B., Zaghloul, K., Haarsma, L., & Sterling, P. (2001). Bipolar cells contribute to nonlinear spatial summation in the brisk-transient (Y) ganglion cell in mammalian retina. *Journal of Neuroscience, 21*, 7447-7454.

Djupsund, K., Fyhrquist, N., Hariyama, T., & Donner, K. (1996). The effect of background luminance on visual responses to strong flashes: Perceived brightness and the early rise of photoreceptor responses. *Vision Research, 36*, 3253-3264.

Donner, K. (1981). Receptive-fields of frog retinal ganglion-cells - response formation and light-dark-adaptation. *Journal of Physiology-London, 319*, 131-142.

Donner, K. (1989). Visual latency and brightness: An interpretation based on the responses of rods and ganglion-cells in the frog retina. *Visual Neuroscience, 3*, 39-51.

Donner, K., Copenhagen, D.R. & Reuter, T. (1990). Weber and noise adaptation in the retina of the toad Bufo marinus. *Journal of General Physiology 95*, 733-753.

Donner, K., & Fagerholm, P. (2003). Visual reaction time: Neural conditions for the equivalence of stimulus area and contrast. *Vision Research, 43*, 2937-2940.

Dunn, F. A., Lankheet, M. J., & Rieke, F. (2007). Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature, 449*, 603-606.

Enroth-Cugell, C., & Robson, J.G. (1966). Contrast sensitivity of retinal ganglion cells of cat. *Journal of Physiology-London, 187*, 517-552.

Foley, J.M. (1994). Human luminance pattern-vision mechanisms - masking experiments require a new model. *Journal of the Optical Society of America A-Optics Image Science and Vision, 11*, 1710-1719.

Friedburg, C., Thomas, M.M., & Lamb, T.D. (2001). Time course of the flash response of dark- and light-adapted human rod photoreceptors derived from the electroretinogram. *Journal of Physiology-London, 534*, 217-242.

Geisler, W.S. (1978). Adaptation, afterimages and cone saturation. *Vision Research, 18*, 279-289.

Geisler, W.S., & Albrecht, D.G. (1997). Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Visual Neuroscience, 14*, 897-919.

Georgeson, M.A., & Meese, T.S. (2006). Fixed or variable noise in contrast discrimination? The jury's still out. *Vision Research, 46*, 4294-4303.

Goris, R.L.T., Wichmann, F.A., & Henning, G.B. (2009). A neurophysiologically plausible population code model for human contrast discrimination. *Journal of Vision, 9*.

Goris, R.L.T., Zaenen, P., & Wagemans, J. (2008). Some observations on contrast detection in noise. *Journal of Vision, 8*.

Green, M.D., & Swets, A.M. (1988). *Signal detection theory and psychophysics* Peninsula publishing, Los Altos.

Hayhoe, M.M., Benimoff, N.I., & Hood, D.C. (1987). The time-course of multiplicative and subtractive adaptation process. *Vision Research, 27*, 1981-1996.

Hayhoe, M.M., Levin, M.E., & Koshel, R.J. (1992). Subtractive processes in light adaptation. *Vision Research, 32*, 323-333.

Heikkinen, H., Nymark, S., Donner, K. & Koskelainen, A. (2009). Temperature dependence of dark-adapted sensitivity and light-adaptation in photoreceptors with A1 visual pigments: A comparison of frog L-cones and rods. *Vision Research, 49*, 1717-1728.

Henning, G.B., & Wichmann, F.A. (2007). Some observations on the pedestal effect. *Journal of Vision, 7*.

Hood, D.C. & Birch, D.G. (1993). Human cone receptor activity: The leading edge of the a-wave and models of receptor activity. *Visual Neuroscience, 10*, 857-871.

Jacobs, A. L., Fridman, G., Douglas, R. M., Alam, N. M., Latham, P. E., Prusky, G. T., et al. (2009). Ruling out and ruling in neural codes. *Proceedings of the National Academy of Sciences, 106*, 5936-5941.

Kilpeläinen, M., Nurminen, L., & Donner, K. (2011). Effects of mean luminance changes on human contrast perception: Contrast dependence, time-course and spatial specificity. *PLoS ONE, 6*, e17200.

Kingdom, F. A. A., & Whittle, P. (1996). Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. *Vision Research, 36*, 817-829.

Klein, S.A., & Levi, D.M. (2009). Stochastic model for detection of signals in noise. *Journal of the Optical Society of America A-Optics Image Science and Vision, 26*, B110-B126.

Kontsevich, L.L., & Tyler, C.W. (1999). Nonlinearities of near-threshold contrast transduction. *Vision Research, 39*, 1869-1880.

Lamb, T.D., & Pugh, E.N. (1992). A quantitative account of the activation steps involved in phototransduction in amphibian photoreceptors. *Journal of Physiology-London, 449*, 719-758.

Legge, G.E., & Foley, J.M. (1980). Contrast masking in human vision. *Journal Of The Optical Society Of America, 70*, 1458-1471.

Ludwig, C. J. H., Gilchrist, I. D., McSorley, E., & Baddeley, R. J. (2005). The temporal impulse response underlying saccadic decisions. *Journal of Neuroscience, 25*, 9907-9912.

Mansfield, R.J.W. (1976). Visual adaptation: retinal transduction, brightness and sensitivity. *Vision research, 16*, 679-690.

Meese, T.S., & Baker, D.H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a "Swiss cheese" stimulus. *Journal of Vision, 11*.

Nachmias, J., & Sansbury, R.V. (1974). Grating contrast - discrimination may be better than detection. *Vision Research, 14*, 1039-1042.

Navarro, R., Artal, P., & Williams, D.R. (1993). Modulation transfer of the human eye as a function of retinal eccentricity. *Journal of the Optical Society of America A-Optics Image Science and Vision, 10*, 201-212.

Pelli, D.G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A-Optics Image Science and Vision, 2*, 1508-1532.

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

Rieke, F., & Rudd, M.E. (2009). The challenges natural images pose for visual adaptation. *Neuron, 64*, 605-616.

Schnapf, J.L., Nunn, B.J., Meister, M., & Baylor, D.A. (1990). Visual transduction in cones of the monkey macaca fascicularis. *Journal of Physiology-London, 427*, 681-713.

Schneeweis, D.M., & Schnapf, J.L. (1999). The photovoltage of macaque cone photoreceptors:

Adaptation, noise, and kinetics. *Journal of Neuroscience, 19*, 1203-1216.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. In: N. Osborne, & G. Chader (Eds.), Progress in retinal research, 3. Pergamon Press, Oxford, pp. 263-346.

Solomon, J. (2009). The history of dipper functions. *Attention, Perception, and Psychophysics, 71*, 435-443.

Trong, P.K., & Rieke, F. (2008). Origin of correlated activity between parasol retinal ganglion cells. *Nature Neuroscience, 11*, 1343-1351.

Uzzell, V.J., & Chichilnisky, E.J. (2004). Precision of spike trains in primate retinal ganglion cells. *Journal of Neurophysiology, 92*, 780-789.

Vassilev, A., Mihaylova, M., & Bonnet, C. (2002). On the delay in processing high spatial frequency visual information: Reaction time and VEP latency study of the effect of local intensity of stimulation. *Vision Research, 42*, 851-864.

Weber, A.J., & Harman, C.D. (2005). Structure-function relations of parasol cells in the normal and glaucomatous primate retina. *Investigative Ophthalmology and Visual Science, 46*, 3197-3207.

Westheimer, G., & Campbell, F.W. (1962). Light distribution in the image formed by the living human eye. *Journal of the Optical Society of America, 52*, 1040-1044.

Winn, B., Whitaker, D., Elliott, D.B., & Phillips, N.J. (1994). Factors affecting light-adapted pupil size in normal human-subjects. *Investigative Ophthalmology and Visual Science, 35*, 1132-1137.

Wohrer, A., & Kornprobst, P. (2009). Virtual retina: A biological retina model and simulator, with contrast gain control. *Journal of Computational Neuroscience, 26*, 219-249.

von Wiegand, T.E., Hood, D.C., & Graham, N. (1995). Testing a computational model of light-adaptation dynamics. *Vision Research, 35*, 3037-3051.