Supporting information

Rod intrusion

We excluded TCSF data from the model development that was likely to be influenced by rodmediated detection of the flicker. Except for the TCSF experiments of Stockman *et al.* [\[1\]](#page-8-0) shown in Fig 5, in which rod-mediated detection was precluded, little attempt was made to exclude rod involvement at lower light levels in most TCSF studies. Some rod intrusion could be expected at the lower levels in the experiments of De Lange [\[2\]](#page-8-1), Kelly [\[3\]](#page-8-2), Roufs [\[4\]](#page-8-3) and Rovamo *et al.* [\[5\]](#page-8-4), who used circular 2°, 68°, 1°, and 1.66° centrally-viewed white flickering targets, respectively. With good fixation control, the small centrally-viewed targets used by all but Kelly might be expected to fall mainly in the rod-free fovea and thus produce mainly cone-detected flicker. However, the need for Roufs to dark-adapt his subjects for up to 45 minutes at low light levels (see p. 263 of [4]) and De Lange for 15 minutes ["with the (room) light switched off" at lower levels (see [p. 49](#page-8-5) of [6]) suggests that rod detection was important and central fixation was not necessarily well maintained. Rod intrusion should be less for Swanson *et al.* [\[7\]](#page-8-6), who used a light that was metameric with 600 nm, and least for von Wiegand *et al.* [\[8\]](#page-8-7), who used a 635-nm target light, but in both cases some rod intrusion is possible at low temporal frequencies and low luminances (e.g. [\[9,](#page-8-8) [10\]](#page-8-9)).

Since there is likely to be rod intrusion in the data at low levels in Figs 2-4 and 6, we were faced with the problem of determining which levels should be excluded from our analysis so that we could be confident that we were modelling cone-mediated vision.

Broadly speaking, rod saturation begins at about 100 scotopic trolands [\[11-13\]](#page-8-10), which corresponds, for example, to about 40 to 70 *photopic* trolands for standard whites such as Illuminants D65 and A [\(Table 1\(2.4.4\) of \[14\]](#page-9-0)). Thus, a conservative cut-off might be about 40 photopic trolands. However, based on the form of the data and the fact that the experimental conditions for De Lange and Roufs did not favour rod-detection, we have chosen a lower cut-off near 10 photopic trolands. Additionally, this cut-off coincides with the level below which the parameter *k* in the model*,* when allowed to vary*,* began to decrease and the TCSF to change in shape from bandpass to low-pass. Consequently, a plausible simplification, and an assumption that we adopted in our modelling, is that *k* is constant for cone-mediated vision.

Time-domain representation of the model

Going from the frequency domain to the time domain strictly requires knowledge of the phase response as well as amplitude response (Equation (2)) of the system. However, we note that the phase responses of LP-stages are determined by their temporal response (which is an exponential decay over time), and our model can be represented in the complex Fourier domain as,

$$
H(f) = g \frac{\left(if + (1-k)f_c \right)^2}{\left(if + f_c \right)^6 \left(if + f_{c1} \right)^2} , \tag{A}
$$

where $i = \sqrt{-1}$. The triphasic temporal response, $R(t)$, of the model is found by taking the inverse Fourier transform of Equation (A), which gives in terms of time constants, i.e., $\tau = \frac{1}{2}$ $\tau = \frac{1}{2\pi f_c}$ and

$$
\tau_{\iota} = \frac{1}{2\pi f_{cl}}:
$$

$$
y(t) = g\left(Ae^{-\frac{t}{\tau}} + Be^{-\frac{t}{\tau_{L}}}\right),
$$
 (B)

where:

$$
A(t) = \sum_{u=0}^{5} \frac{\left(\frac{k}{\tau}\right)^{2} \frac{(6-u)}{u!} (xt)^{u}}{x^{7}} + \sum_{v=0}^{4} \frac{2 \frac{k}{\tau} \frac{(5-v)}{v!} (xt)^{v}}{x^{6}} + \sum_{w=0}^{3} \frac{(4-w)}{w!} (xt)^{w}
$$

\n
$$
B(t) = \sum_{z=0}^{1} \left\{ \frac{\left(\frac{k}{\tau}\right)^{2} (xt-6)}{x^{7}} + \frac{2 \frac{k}{\tau} (xt-5)}{x^{6}} + \frac{(xt-4)}{x^{5}} \right\}
$$

\nand,
\n
$$
x = \frac{1}{\tau} - \frac{1}{\tau_{i}}
$$

Equation (B) has been plotted in the lower right orange panel of Fig 8 with *k*=0.80, *fc*=15 and *fcL*=30; *g* has been chosen to normalise the peak to 1.

The amplitude response of a cascade of leaky integrators is asymptotically a power law, but exponential for a range of visually significant frequencies

The amplitude response, *A(f)*, of a single stage of leaky integration (low-pass filtering), or an "LP-stage" for short, is given by,

$$
A(f) = \frac{c}{\sqrt{f^2 + f_c^2}} \text{ or } \log_{10}(A(f)) = C - \frac{1}{2} \log_{10}(f^2 + f_c^2),
$$
 (C)

where *f* is frequency in Hz, $C = \log_{10}(c)$ is the logarithm of the overall gain, $f_c = 1/2\pi\tau$ is the corner (or cut-off) frequency of the low-pass filter where *τ* (seconds) is the time constant (or integration time) of the leaky integrator. This logarithmic form of Equation (C) has been plotted against frequency in Panels [A] and [B] of Fig 12 as the solid cream lines. The high frequency asymptote, when f is large compared to f_c , is given by,

$$
A(f) \approx cf^{-1} \text{ or } \log_{10}(A(f)) \approx C - \log_{10}(f), \tag{D}
$$

which is a power law function of *f*, with an exponent of minus one. Plotted on double logarithmic coordinates (see the dashed red line in Panel [B] of Fig 12) this becomes a straight line with a slope of minus one. However, this approximation is valid only when the frequency is high, say above 3 times the corner frequency, as shown by the red shaded region in Fig 12. At lower frequencies the amplitude response is much better approximated by an exponential function, which is a straight line when log amplitude is plotted on a linear frequency scale (Panel [A]). To understand why, consider the Taylor series expansion of the logarithmic response function (Equation (C)) around *f*=*fc*. (We choose to expand around f_c as it is an inflection point of $log_{10}(A)$, i.e., the second derivative is equal to zero at this point and there is no second-order term in the Taylor expansion and so log₁₀(A) will be closest to a straight line at this point.)

$$
\log_{10}(A(f)) = C - \frac{1}{2} \log_{10}(f^2 + f_c^2)
$$

= $C - \frac{1}{2 \log_e(10)} \left[\log_e(2f_c^2) - \frac{1}{f_c}(f - f_c) + \frac{1}{6f_c^3}(f - f_c)^3 - \frac{1}{8f_c^4}(f - f_c)^4 + O(f - f_c)^5 \right]$ (E)

For values of *f* close to *fc* the expansion will be dominated by the constant and linear terms, i.e.,

$$
\log_{10}\left(A(f)\right) \approx D - \frac{1}{2\log_e(10)f_c}(f - f_c),\tag{F}
$$

where D = C $-\frac{\log_e\left(2f_c^2\right)}{2\log_e\left(10\right)}$ = $\log_{10}\left(A(f_c)\right)$ *c e f* $D = C - \frac{C}{\sigma L} \frac{C}{\sigma L} = \frac{C}{\sigma L} \left(A(f_c) \right)$. This function has been plotted in both panels of Fig 12 as

the dashed black line and can be seen to be a good approximation to the filter shape between about 0.36*fc* and 1.92*fc* (i.e., the error is between ±3.5% in the green shaded region), over which sensitivity declines by about 0.31 log_{10} units (i.e., by about 51%).

For a cascade of LP-stages in a linear system the convolution theorem states that the total response of the cascade is equal to the product of the responses of the individual stages. Or, equivalently, the logarithm of the cascade response is the sum of the logarithms of the individual responses. At high frequencies *n* LP-stages will tend toward a power law with an exponent of minus *n* (a straight line of slope minus *n* when plotted on double logarithmic coordinates). This result has been noted many times before [\[15,](#page-9-1) [16\]](#page-9-2) and has occasionally been used as justification for rejecting a leaky integrator model altogether [\[17\]](#page-9-3) because the *measurements* at high frequencies don't assume this form. However, between about 0.36*fc* and 1.92*fc* a cascade of *n* identical LP-stages has an approximately exponential frequency response with an exponent of $-n/\sqrt{2\log_e(10)f_c}$, and sensitivity will decline by about 0.31*n* log₁₀ units. A cascade of six such stages will have an approximately exponential response over 1.86 log_{10} units, close to the maximum flicker sensitivity range seen in TCSFs. This range crucially depends on the LP-stages having the same corner frequencies, if the corner frequencies differ the frequency ranges over which they each approximate an exponential function will only partially overlap and the good exponential approximation to the TCSF will cover a smaller region. The value of the exponent of the approximation in this region, i.e., the slope of the linear region when log sensitivity is plotted against linear frequency, will be

$$
-\frac{1}{2\log_e(10)}\sum_{j=1}^n\frac{1}{f_{cj}}\text{ where }f_{cj}\text{ denotes the corner frequency of the }j^{th}\text{ LP-stage.}
$$

For a cascade of leaky integrators all the corner frequencies (or equivalently, all the time constants) should take similar values.

In addition to producing the largest range of frequencies over with the TCSF is approximately exponential, it is also the case that having identical corner frequencies produces maximal temporal contrast sensitivity at all frequencies. Consider the simplest case with two stages with corner frequencies C_1 and C_2 . The amplitude response, $A(f)$, is then given by:

$$
A(f) = \frac{C_1 C_2}{\sqrt{f^2 + C_1^2} \sqrt{f^2 + C_2^2}}\,,\tag{G}
$$

where *f* is temporal frequency in Hz. Note that the numerator is the product of the corner frequencies in order to set the low frequency asymptote to 1. The high-frequency asymptote on a log-log plot is then a straight line with slope -2 which passes through the point (C_1C_2 , 1). The way we have defined our amplitude response means the low-frequency limb is fixed at *A*=1 for all *C1* and *C2* and the location of the high-frequency limb depends solely on the product of *C1* and *C2*, so the family of curves that have the same low- and high-frequency limbs can be parametrised as $C_1 = C/X$ and $C_2 = CX$, for some *C* (which is the geometric mean of C_1 and C_2). C and *X* are both greater than 0. In the case when the corner frequencies are identical, i.e., *X*=1, Equation (G) reduces to:

$$
A_{s}(f) = \frac{C^{2}}{(f^{2} + C^{2})}, \qquad (H)
$$

while if $X \neq 1$:

$$
A_d(f) = \frac{C^2}{\sqrt{f^2 + \left(\frac{C}{X}\right)^2} \sqrt{f^2 + \left(CX\right)^2}} ,
$$
 (1)

where the subscripts *s* and *d* denote corner frequencies that are the *same* or *different*. The question is which of these two amplitude responses is larger for any given frequency. We note that *C*, *X* and *f* are all strictly positive and so both A_s and A_d are positive, which therefore means that $A_s > A_d$ if, and only if, $A_s^2 > A_d^2$. We have:

$$
A_s^2 = \frac{C^4}{\left(f^4 + 2C^2 + C^4\right)} \text{ and } A_d^2 = \frac{C^4}{f^4 + \left(\frac{1}{X^2} + X^2\right) + C^2 + C^4},
$$
 (J)

so that (by inspection of the denominators in Equations (J)), $A_s^2 > A_d^2$ if, and only if, $2 < \frac{1}{\sqrt{2}} + X^2$ *X* $\langle \frac{1}{2}+X^2 \text{ and}$ this is clearly true for all $X \neq 1$ and is independent of *f*. So, we see that having stages with identical corner frequencies will always lead to optimal sensitivity at all frequencies compared to two stages

with different corner frequencies. The above reasoning can be easily extended to the case of more than two filters.

A more intuitive explanation of why having identical corner frequencies is optimal is to consider a serial cascade of two filters (leaky integrators), one of which is slower (has a lower corner frequency) than the other. If the slower filter comes first, then the second filter will be capable of passing higher frequencies than it receives as input and its sensitivity would be wasted. Similarly, if the slower filter comes second then the first filter passes high frequencies which the second attenuates, which again would be wasteful. If the filters have the same corner frequency, then they are optimally tuned to pass (or attenuate) the same set of frequencies. While it is highly unlikely that all the significant LP-stages of light adaptation in the visual system have identical time constants, any process which seeks to optimise the efficiency of the system might be expected to make them as near to identical as possible.

Subtractive inhibition maintains the exponential fall-off but reduces its slope and shifts it to higher frequencies

While a cascade of LP-stages can produce an exponential decline in sensitivity over the visible range of flicker frequencies, it does not explain the loss of sensitivity at low frequencies. We model the loss of low-frequency sensitivity as the result of inhibition. A standard high-pass filter can be constructed by passing a signal through a leaky integrator with unity DC gain and subtracting the result from the original signal. If the DC gain is less than unity, the filter will only partially cancel signals at low frequencies. These kinds of filter can be thought of as partial high-pass filters and are often used in electric engineering (where they are referred to as "lead-compensators") in order to increase stability in control circuits and sharpen the temporal response. The amplitude response of such a filter is given by:

$$
A(f) = \frac{\sqrt{f^2 + ((1-k)f_c)^2}}{\sqrt{f^2 + f_c^2}} ,
$$
 (K)

where *k* is the gain of the filter whose response is to be subtracted and can be thought of as the strength of inhibition. Note that when *k*=1 there is complete inhibition at low frequencies; i.e., a standard high-pass filter, while for *k*=0 there is no inhibition and the numerator and denominator in Equation (K) cancel, *i.e.*, an all-pass filter. Our model is made up of 6 LP-stages and two leadcompensators with the same DC gain. Based on previous work we assume that two central LP-stages do not light adapt while the other four LP-stages do. This is also consistent with the idea that a major limiting factor of visual sensitivity is the dynamic range of post-receptoral spiking neurons, so the processes of light adaptation should be most pronounced early in the visual pathway. For simplicity, and to account for the large frequency range of exponential approximation, we assume that the variable LP-stages change together and the filters in the inhibition adapt in the same way. The equation for this model is given by,

$$
A(f) = g \frac{f^2 + ((1-k)f_c)^2}{(f^2 + f_c^2)^3 (f^2 + f_{c1}^2)} \text{ or}
$$

\n
$$
\log_{10}(A(f)) = G + \log_{10}(f^2 + ((1-k)f_c)^2) - 3\log_{10}(f^2 + f_c^2) - \log_{10}(f^2 + f_{c1}^2),
$$
 (L)

where f_c and f_{cL} are the corner frequencies of the variable and fixed stages, respectively. This function, where $G = log_{10}(g)$, was fitted to the TCSF data by varying G , f_c , f_{cl} and k using a nonlinear least squares procedure in SigmaPlot or MATLAB. Note that f_{cl} was fixed across observers and light levels, *k* was fixed separately for each observer (except at low levels), and *G* and *fc* varied both with observer and with light level. The fits are plotted as the black or gray lines in Figs 2-6. The shape of Equation (L) depends on k , f_c and f_{c} .

The general equation for our full model does not lend itself to a simple Taylor series expansion as it did in the single leaky integrator case, so instead we use a technique from computer vision to examine the approximate straight-line portion of the log sensitivity versus linear frequency plot, namely the Hough transform [\[18\]](#page-9-4). Any straight-line can be parameterised by two quantities, one denoting a direction and one denoting a position, often given as the slope and the intercept of the yaxis. In the Hough transform these quantities are typically taken as the orientation of the line, *θ*, and its shortest distance to the centre of the image, *ρ*. In computer vision, the transform is used to map each point in an image to a segment of a sinusoidal curve through *θ*-*ρ* space. Points that are collinear will produce curves that intersect at a single point (*θ, ρ*); i.e., a straight-line in the image will appear as a peak in the summed distribution of curves in *θ*-*ρ* space. We generated an image of the TCSF curve (with values of $f_c=15$, $f_{cl}=30$ and $k=0.8$) and applied the Hough transform to extract the largest (approximately) straight line in the image. From this we can find the range of the linear fit in terms of sensitivity (~1.75 log_{10} units) and frequency (in this case, between about 13.1 and 47.1 Hz), and the slope is -19.4 Hz/decade.

High-frequency linearity and low frequency Weber's law

At high frequencies (when $f \gg f_c \& f_{cL}$) the equation for our model (J) simplifies to:

$$
A(f) \approx \frac{g}{f^6}, \tag{M}
$$

which, in the traditional double logarithmic coordinates of Bode [\[19\]](#page-9-5), is a straight-line with slope of -6 passing through *log10(g)* at 1 Hz where *-6* log10*(f) = 0*. If the high-frequency power law region of each TCSF were within the visible range, then high-frequency linearity would hold if and only if *g* were constant. Inspection of Panel [B] of Fig 9 shows this is approximately true up to about 3.5 log_{10} Tds. However, as noted above, the region where the power law approximation holds lies largely above the temporal acuity limit above which sensitivity cannot be measured and below which TCSFs are approximately exponential functions. In semi-logarithmic coordinates the exponential losses in sensitivity appear as straight-lines of different slopes, and so cannot strictly conform to "highfrequency linearity", i.e., they could only coincide at a single intersection point, not over an extended range. Plotted on a logarithmic frequency scale, however, these lines accelerate downwards and only appear to coincide. We suggest that the notion of high-frequency linearity is an inappropriate inference based on the way in which amplitude sensitivity has been plotted in the past; it is not a feature of visual sensitivity.

At low frequencies (when $f \ll f_c \& f_{cl}$) our model (J) simplifies to:

$$
A(0) = \frac{g(1-k)^2}{f_c^4 f_{c1}^2}
$$
 (N)

Now Weber's Law holds if *ΔI/I* is constant, where *ΔI=A* is the amplitude threshold. In our model *k* and f_{cl} are fixed across light levels, but *g* and f_c vary, and, from Equation (N), Weber's law will hold if:

$$
\frac{f_c^4}{lg} = W \,, \tag{O}
$$

where W is a constant that is proportional to the Weber fraction. Below about 3.4 log₁₀ Td, *g* is constant (Fig 9[B]) so Weber's law will hold if *fc* increases in proportion to the 4th root of *I*. However, the best fitting power law function relating f_c to *I* (see Equation (3), plotted as the blue curve in Fig 9[A]) has an exponent of 0.181, which corresponds to the 5.5th root of *I*. Above 3.4 log₁₀ Td, f_c is constant while *g* decreases in proportion to *I* (see Equation (4) and Fig 9[B]), so here W will be constant. Thus, according to our model, Weber's law is not strictly maintained at low frequencies at low- to mid-light levels. Note that low-frequency flicker thresholds are distinct from flash thresholds, which do indeed show good adherence to Weber's law over a large range of light levels [\[20\]](#page-9-6), but this will likely depend on the shape and size of the impulse response rather than flicker sensitivity at any specific frequency.

References

- 1. Stockman A, Langendörfer M, Smithson HE, Sharpe LT. Human cone light adaptation: from behavioral measurements to molecular mechanisms. J Vision. 2006;6(11):1194-213. doi: 10.1167/6.11.5. PubMed PMID: 17209729.
- 2. de Lange H. Research into the dynamic nature of the human fovea-cortex systems with intermittent and modulated light. I. Attenuation characteristics with white and colored light. J Opt Soc Am. 1958;48:777-84.
- 3. Kelly DH. Flicker fusion and harmonic analysis. J Opt Soc Am. 1961;51:917-8. PubMed PMID: 13752373.
- 4. Roufs JAJ. Dynamic properties of vision-I. Experimental relationships between flicker and flash thresholds. Vision Res. 1972;12:261-78.
- 5. Rovamo J, Raninen A, Donner K. The effects of temporal noise and retinal luminance on foveal flicker sensitivity. Vision Res. 1999;39:533-9.
- 6. de Lange H. Attenuation characteristics and phase-shift characteristics of the human foveacortex systems in relation to flicker-fusion phenomena. Delft: TU Delft; 1957.
- 7. Swanson WH, Ueno T, Smith VC, Pokorny J. Temporal modulation sensitivity and pulsedetection thresholds for chromatic and luminance perturbations. J Opt Soc Am A. 1987;4(10):1992-2005. PubMed PMID: 3430210.
- 8. von Wiegand TE, Graham N, Hood DC. Testing a computational model of light-adaptation dynamics. Vision Res. 1995;35(21):3037-51. PubMed PMID: 8533341.
- 9. Hecht S, Verrijp CD. The influence of intensity, color and retinal location on the fusion frequency of intermittent illumination. Proceedings of the National Academy of Sciences of the United States of America. 1933;19(5):522-35. PubMed PMID: 16587779; PubMed Central PMCID: PMC1086060.
- 10. Wald G. Human vision and the spectrum. Science. 1945;101(2635):653-8. doi: 10.1126/science.101.2635.653. PubMed PMID: 17777531.
- 11. Aguilar M, Stiles WS. Saturation of the rod mechanism of the retina at high levels of stimulation. Optica Acta. 1954;1:59-64.
- 12. Adelson EH. Saturation and adaptation in the rod system. Vision Res. 1982;22(10):1299-312. PubMed PMID: 7179751.
- 13. Sharpe LT, Fach CC, Nordby K, Stockman A. The incremental threshold of the rod visual system and Weber's law. Science. 1989;244(4902):354-6. PubMed PMID: 2711186.
- 14. Wyszecki G, Stiles WS. Color Science: concepts and methods, quantitative data and formulae. 2nd ed. New York: Wiley; 1982.
- 15. Kelly DH. Flicker. In: Jameson D, Hurvich LM, editors. Handbook of Sensory Physiology. VII/4. Berlin: Springer-Verlag; 1972. p. 273-302.
- 16. Watson AB. Temporal sensitivity. In: Boff K, Kaufman L, Thomas J, editors. Handbook of Perception and Human Performance. 1. New York: Wiley; 1986. p. 6-1-6-43.
- 17. Kelly DH. Diffusion model of linear flicker responses. J Opt Soc Am. 1969;59(12):1665-70. PubMed PMID: 5360496.
- 18. Duda RO, Hart PE. Use of the Hough transformation to detect lines and curves in pictures. Communications of the Association for Computing Machinery. 15(1):11-5.
- 19. Bode HW. Network analysis and feedback amplifier design. Princeton: van Nostrand; 1945.
- 20. Stiles WS. Mechanisms of Colour Vision. London: Academic Press; 1978.

Tables

Table **A**

Table **A** (continued)

Best-fitting falling high frequency slopes in units of Hz per log₁₀ unit and R^2 values for each fit. Levels are from low to high as listed in the keys of Figure 2-6. The high-frequency slope cannot be reliably estimated at the lowest three levels for Rovamo *et al.* (1999) because of limited data. For other details see text.

Table **B**

Table **B** (continued)

Best-fitting variable corner frequencies (fc), logarithmic gains (log₁₀(g)) and feedforward gain (k). Two LP-stages, which were fixed across observers, had corner frequencies of 30.92 ±2.23 Hz. Adjusted R^2 for simultaneous fit = 99.6% Levels are from low to high as listed in the keys of Figure 2-6. For details see text.