Manuscript submitted to *eLife*. Type of article: *Research Advance*. English language: *no proof read*.

Software availability: <https://github.com/tzvet/Data-Model-MotionSuppressionFacilitation-2018>

Title: Suppression and facilitation of motion perception in humans

Authors: Tzvetomir Tzvetanov

Affiliation: Anhui Province Key Laboratory of Affective Computing and Advanced Intelligent Machine, School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, P.R. China

Contact: tzvetan@hfut.edu.cn

Abstract

In a study that investigated the putative neuronal origin of suppression and facilitation in human motion perception Schallmo et al. (2018) used various techniques to investigate how motion perception is shaped by excitatory and inhibitory interactions within hMT+ or earlier areas. They further proposed that neuronal normalization is a sufficient account of the behavioural results, discounting accepted and precise neuronal mechanisms of excitation or inhibition. In this Research Advance, it is shown (1) that once the full computational model that predicts the psychophysical results is defined, it is not “divisive normalization” but actually excitatory and inhibitory processes that are the neuronal mechanisms shaping facilitation and suppression in the behavioural domain, then (2) that the experimental design they used allows a *quantitative* comparison and usage of such “contrast–size tuning” data.

Introduction

Changes in perception are interesting scientific tools in the search of how and where is visual perception created in our brain (Spillmann & Dresch, 1995; Spillmann & Ehrenstein, 1996; Spillmann & Werner, 1996; Eagleman, 2001; Albright & Stoner, 2002; Born & Bradley, 2005). Human motion perception is known to exhibit various astonishing effects, among which very known in the public domain are motion illusions (for some examples among many websites, see <https://michaelbach.de/ot/index.html>), with even simple motion stimuli providing sometimes counter-intuitive results.

Concerning the percept of a single motion direction, that is, of an object or a collection of objects (as rain drops) moving in a single unidirectional motion, it is considered to be created from the interaction between excitatory, pooling information across space and directions, and inhibitory, suppressing information across space and directions, interactions between motion sensitive neurons. For example, the percept of random-dot kinematograms/patterns (RDK/RDPs) was reported to be strongly shaped by the context in which the target motion was presented (Tynan & Sekuler, 1975; Levinson & Sekuler, 1976; Marshak & Sekuler, 1979; Chang & Julesz, 1984; Watamaniuk, Sekuler & Williams, 1989; Nawrot & Sekuler, 1990). Limiting the topic to the spatial spread of the interactions, two interesting studies by two Japanese scientists (Murakami & Shimojo, 1993; Murakami & Shimojo, 1996) showed that size tuning of these centre-surround contextual interactions in motion perception seems to be eccentricity invariant. That is, when rescaling the measures for different eccentricities all results naturally fell on a single curve. This somewhat recalls the size tuning characteristics of receptive fields and their magnification with visual eccentricity. The authors proposed a simple spatial centre-surround model of receptive fields for explaining their observations.

Yet in another type of psychophysical probes, D. Tadin and colleagues demonstrated that perception of simple moving grating stimuli (Tadin, Lappin, Gilroy & Blake, 2003), or a type of RDK (Tadin & Lappin, 2005), have very peculiar perceptual results with respect to size and contrast of the stimuli (see Figure 1A). Globally large and low contrast stimuli needed much less presentation time for

45 being perceived than large and high contrast stimuli, while the opposite was observed for small
 46 stimuli. That is, there is an inverted effect on perception as a function of contrast for small and large
 47 stimulus sizes (Figure 1A). These observations were also attributed to the contrast and size tuning
 48 properties of motion tuned neuronal populations.

49 This tuning has a simple form with excitatory and inhibitory components (Figure 1B,C). It has a
 50 visuospatial receptive field structure with a relatively small excitatory centre but spatially larger in-
 51 hibitory surround (panel B) and contrast responses such that the excitatory component is activated
 52 much rapidly at low contrasts (panel C). The final neuronal activity is a combination of the re-
 53 sponses of these two drives. It is their precise shape and method to combine them that gives the par-
 54 ticular property of the neurons. If one creates stimuli that somewhat match the preferred character-
 55 istics of these neurons, then it is considered that these neuronal responses are directly influencing
 56 the percept of the stimulus. This gives us a simple way to probe, non-invasively, the particular com-
 57 putational structure of the motion perception system. D. Tadin’s insights and results, including later
 58 work of collaborators and other researchers, are noteworthy because they showed how we can ac-
 59 cess the spatial and contrast characteristics of these motion tuned neurons.

60 The work of Schallmo et al. (2018) builds over the above very short overview about motion percep-
 61 tion and its origin, particularly in the spatial and contrast domains through the use of D. Tadin’s ex-
 62 perimental design. But they stated that the recent decade of surge in more application oriented stud-
 63 ies, which are based on the accumulated knowledge from the earlier works on motion perception, is
 64 not totally justified because (citing their Introduction): “Strong assumptions are often made about
 65 the neuronal processes underlying these seemingly complex interactions between size and contrast
 66 during motion perception.”. Thus, they proposed that the overall effects and observations for the
 67 single grating design of D. Tadin and collaborators (and other researchers who use it) are not neces-
 68 sarily due to excitatory and inhibitory interactions between motion sensitive neurons but instead
 69 arises from the divisive normalization of neuronal activity, that emerges from classic neuronal net-
 70 work computations, and which they, and others, have used as a “computational principle”/“compu-
 71 tational framework”.

72 In the following Research Advance manuscript, I comment on two methodological and theoretical
 73 aspects of their work, and demonstrate that the original design of D. Tadin does indeed allow to ex-
 74 tract information about the putative excitatory and inhibitory processes that shape motion percep-
 75 tion in the spatial and contrast domains. First, I discuss the psychophysics methods and analyses

employed by the authors for the measures they provide. Second, I analyse their modelling approach and show that their model and its application contains conflicting knowledge with established computational neuroscience and behavioural modelling. From these analyses, in a last step, I develop the correct model for predicting motion perception in D.Tadin's design and I show that such data can be used for *quantitative* inference of putative changes in underlying excitatory or inhibitory components, changes that partly can and partly cannot be inferred from the study of Schallmo et al. (2018).

Results

The article of Schallmo et al. (2018) used five techniques for approaching the question about how contrast and size of the stimulus affect the necessary stimulus duration time to perceive the correct direction of motion. From these five methods, the psychophysics measures and modelling are the main foundations. These two methods are strongly intertwined in the argumentation and inferences that can be made from the design in the study. In the following, first two points about psychophysics and data processing are emphasized. Then, shortcomings in the modelling are discussed and the correct model is presented, fit on previously published data, and used to discuss the results of Schallmo et al. (2018).

Psychophysics

Definition of the psychometric function in D. Tadin's design

The method used by the authors is as follows. In this design observers were presented with a single moving stimulus whose contrast, size, presentation duration and direction of motion is varied across trials. The idea is to extract the duration threshold for various stimulus sizes and contrasts. Thus, observers are generally instructed (or learn it during the training before data collection) that (1) on each trial only one of two opposite motion directions can appear (e.g. left- or rightward directions), (2) on some trials the motion may appear so weak that it is not clearly perceived and they may need to guess its direction, and (3) they had to report the direction of motion they thought the presented stimulus had on the trial. These are necessary conditions for measuring the psychometric function in this design, and additionally researchers may, or may not, add more randomness from observer's point of view by measuring thresholds for different contrasts and sizes of the stimulus within the same block. This design is a common "one-stimulus presentation two-alternatives forced choice" design, let's call it here 1stim2AFC (Morgan, Watamaniuk and McKee (2000) called it Method of Single Stimulus; Kingdom and Prins (2009) called it 1AFC), that can be used for either measuring discrimination thresholds or measuring misperception effects (e.g. motion repulsion, Tzvetanov and Womelsdorf (2008); Tzvetanov (2012) ; e.g. tilt illusion, Westheimer

Figure 2: Psychometric function definition. (A) cumulative Gaussian function used to describe psychometric performance in 1stim2AFC designs; black and red curves plot no-bias and biased cases. (B) Replotting the function from (A) by defining Percent correct as the y-variable. (C) Pooling the negative and positive intensities into a single psychometric function of percent correct defined between 50% and 100% correct responses. (D) Threshold at 84% as a function of the bias from pooled psychometric function in (C). All plots have psychometric functions with standard deviation of 1.

(1990); Kapadia, Westheimer and Gilbert (2000)). In this particular design, the experimentalist is interested in the discrimination thresholds of the observers, defined as the necessary duration time of the stimulus for discriminating at some predefined level of correctness (let's say 84%) the direction of motion of the stimulus. It is defined with respect to the midpoint of the psychometric function, as described below. The duration of the stimulus is varied between long durations (persons clearly see the motion direction) and short durations (persons have hard time deciding about the direction). The psychometric function spans the two dimensions of stimulus duration (continuous) and direction (binary). Therefore, the resulting data must be analysed as a single psychometric function spanning the full range of left-to-rightward motions (sign of the stimulus) of different durations (intensity of the stimulus). If we assign negative/positive as leftward/rightward motions, absolute intensity as the duration, and percent "rightward" responses as the y-variable one obtains a single psychometric function spanning the full range of 0 to 100% of proportion responses (Figure 2A). This definition clearly allows to define also the biases that can appear in this 1stim2AFC design, biases that can be due to response bias in insecure trials (decision to respond with always the same key when not sure) or perceptual biases (persons really see something different from "0"). Figure 2A depicts such two functions. The discrimination threshold is thus $x(p=84\%) - x(p=50\%)$.

Instead of analysing the data in this manner, Schallmo et al. (2018) defined the pooled psychometric function by computing the percent correct responses for each physical direction of motion (Figure 2B) and then pooled the data over the two directions (Figure 2C). The consequence of doing this is that, if a bias is present, this psychometric function does not define unambiguously the perceptual threshold of the observers, but also includes the bias. When one compares the extrapolated threshold in Figure 2C for the black (no bias) and red (bias of 1) curves, the change is ~52% higher. One can compute the threshold as a function of the amount of bias when no change in the real discrimination threshold is present (Figure 2D), demonstrating the strong change in thresholds with change in bias.

While one might argue that the results of the experiments still show strong and significant effects across conditions, which is undoubted, the reported thresholds have an unnecessary bias that is easily taken care in at least one of two ways: either through the data analysis procedure described above, or either at the experimental methods level by providing feedback to the observers about their response correctness on each trial. The later method was used by Tadin and Lappin (2005) while Schallmo et al. (2018) did not report to use it. This data analysis is important when one has to make model-to-data adjustment. (further details on psychometric function definition and plausible consequences on threshold estimation are presented in **Appendix A1**).

The composite variable "Size Index" involves simultaneously spatial suppression and summation

The authors reported their data by showing the mean duration thresholds and used the "Size Index" (SI) variable to indicate changes in spatial suppression and facilitation. They defined this variable as the difference of log-thresholds for the smallest and a larger size conditions at a given contrast level. In order for this composite variable to be representative of the concepts of suppression, which they defined as $SI < 0$, and facilitation, which they defined as $SI > 0$, it means that changes of thresholds between a small and larger size automatically labels the change as due to only suppression or only summation. This happens when: (1) at very low contrasts inhibition is very weak and thus the results are mainly influenced by spatial summation instantiated by excitatory mechanisms, therefore one should expect a decreasing threshold as a function of stimulus size (see Figure 1A, $c=9\%$); and (2) at very high contrast inhibition is predominant and thus the results are mainly influenced by suppression, therefore one should expect an increasing threshold as a function of size (see Figure 1A, $c=92\%$, but see the two data points for the smallest size). The overall psychophysical data of Schallmo et al. (2018) does not comply with these rules, in particular the low contrast data, which

has some “V” shape, indicating that both excitatory and inhibitory mechanisms seem to act simultaneously in setting observers perceptual thresholds for motion.

Conclusion from Psychophysics

The type of data obtained from their design must be fit with the 0-100% percent responses range psychometric function, for example a cumulative Gaussian (Figure 2A), and then extract the discrimination thresholds of the observers. Then, contrary to the reported composite “Size Index”, the threshold data should be used directly to infer something about changes of inhibition and excitation in the visual motion system of the observers. This last point will appear very clearly once the model is established and used to predict how inhibition and excitation, together, shape the thresholds.

Modelling

A note on normalization of neuronal activity

Here, before presenting the exact model for their experimental design, I want to comment on the issue of the computational principle/framework that the authors put forward. Schallmo et al. (2018) proposed “divisive normalization” as the binding and main framework for understanding facilitative and suppressive effects in motion perception, as other authors have proposed it for neurophysiology (Reynolds & Heeger, 2009; Carandini & Heeger, 2012).

How does the normalisation property appear? Let’s look at a classic neuronal network with excitatory and inhibitory nodes, respectively noted y_e and y_i . The activity of the excitatory node can be written:

$$\tau_e dy_e/dt = -y_e + F(y_e, y_i, I_e) \quad (1)$$

Here, $F()$ is some function that describes how the nodes are connected and influence each other (through some weights, “firing rate”/“transducer” function, and the inputs to each node I_e and I_i) (Ermentrout, 1998). One general result for neuronal networks is that after a short time of activity the network stabilises itself on a “steady state” (but it could be also a pure oscillatory activity, for a simple introduction to the topic, see Wilson (2005)). What this means is that the activity of a neuron has an initial strong dynamics that typically shows oscillation activity just after input onset (e.g. a burst of firing), and then rapidly stabilises on some steady state activity (for model examples, e.g. Piëch, Li, Reeke and Gilbert (2013), or Wilson (2005)). This steady state activity can (rare) or cannot (in general) be computed in a simple analytical manner, depending on the exact functional form of the above network equations (the function $F()$). This stable point is what Schallmo et al. (2018) consider as a computational principle/framework (see also Appendix of Heeger (1992)).

From the above short presentation, it comes natural that neuronal activity stabilisation, aka normalisation, sometimes appearing as “divisive normalization”, accounts for the neurophysiological results, but it does so due to the exact functional form of the equations (transducer, excitatory and inhibitory effects...) that the modeller decided to put in it for matching the model to the data. They are carefully considered based on prior knowledge, especially neurophysiological but also behavioural reports. Schallmo et al. (2018), themselves, do fix their normalised activities by making the assumption that excitatory (facilitation) effects are in the numerator and inhibitory (suppressive) effects are in the denominator (their Equation 1 reproduced here):

$$R = E/(S + \sigma) \quad (2)$$

The exact reason for having such a form of stabilised activity is not given, out of the “divisive normalization principle” (although their model description hints to a classic feedforward model). This is only one possibility of normalisation result. The following two examples of neuronal networks will give further possibilities.

215 The first example is a V1-layer 4 cell from the model proposed by Grossberg and Raizada (2000)
216 (their equations 17-18) for explaining contrast and attentional effects observed in neurophysiology
217 of V1. The activity y_{ijk} of these V1-layer 4 cells at equilibrium is:

$$218 \quad y_{ijk} = \frac{C_{ijk} + \eta^+ x_{ijk} - \sum_{pqr} W_{pqrijk}^+ m_{pqr}}{1 + C_{ijk} + \eta^+ x_{ijk} + \sum_{pqr} W_{pqrijk}^+ m_{pqr}} \quad (3)$$

219 where C_{ijk} are processed LGN inputs that are modulated by layer 6 cells activities (x_{ijk}) organised in
220 an excitatory-centre ($\eta^+ x_{ijk}$) and inhibitory-surround ($W_{pqrijk}^+ m_{pqr}$) structure, this later being transmit-
221 ted through a layer 4 inhibitory network activities m_{pqr} (see their equation 19). Leaving aside further
222 considerations about their model, this example shows how nominator and denominator include ex-
223 citatory and inhibitory effects from various origins. It should be noted that these other feedback
224 activities are themselves function of layer 4's activities and other V1 layers, thus making the exact
225 computation not analytically straightforward and necessitating numerical simulations.

226 The second example is a model of centre-surround interactions in motion processing taken from
227 Kim and Wilson (1997) (their equation 6). The time dependent activity, C_θ , of one neuron in a
228 centre population sensitive to direction of motion θ is:

$$229 \quad \tau dC_\theta/dt = -C_\theta + S \left(E_{C,\theta} - \sum_{i=\theta-120}^{\theta-45} \alpha_i C_i - \sum_{i=\theta+45}^{\theta+120} \alpha_i C_i - Nf(x) \sum_{k=\theta-60}^{\theta+60} \phi_k A_k \right) \quad (4)$$

230 that gives at equilibrium the result:

$$231 \quad C_\theta = S \left(E_{C,\theta} - \sum_{i=\theta-120}^{\theta-45} \alpha_i C_i - \sum_{i=\theta+45}^{\theta+120} \alpha_i C_i - Nf(x) \sum_{k=\theta-60}^{\theta+60} \phi_k A_k \right) \quad (5)$$

232 where $S()$ is a simple transducer ($S(x)=x$ if $x>0$, else 0), $E_{C,\theta}$ is the preprocessed pattern motion input
233 to the centre at direction θ , and A_k are the activities of spatially surround neurons that themselves
234 undergo excitation and inhibition in a similar manner (the remaining factors are network connection
235 weights, see their tables 1 and 2). In this example the stabilized activities of the neurons are a
236 simple function of input and subtraction of the inhibitory parts coming from the other cells with sta-
237 bilized activities. If one wants to see a “divisive normalisation” in this example, we can say that the
238 denominator is equal to 1, which is due to the particular network structure taken in the article. The
239 normalisation property here comes from the facts that the input $E_{C,\theta}$ is itself limited between a min-
240 imum a maximum, there are only inhibitory interactions, and the transducer does not allow for neg-
241 ative activities (see equation 4). In this case too, because of the intertwined interactions through all
242 spatial and motion direction lateral connections, the computation of the stable activities were ob-
243 tained through numerical simulations. (It is noteworthy to mention that Tadin and Lappin (2005)
244 use a simplified version of such a network, see their Table 1 and further below, section **The correct**
245 **model for D. Tadin's design**).

246 What did we see from these examples? First, that the exact relation for the neuronal activity at equi-
247 librium is not necessarily an excitatory effect divided by an inhibitory effect; second, that the rela-
248 tion can be widely different depending on the exact functional network and connectivities that are
249 assumed by the modeller. The fact is: equilibrium activities of a neuronal network can be obtained
250 from divisive normalisation; it is a simple plausible computation; the exact terms in the denomin-
251 ator and nominator depend on the network structure the researchers are interested in, based on pre-
252 liminary knowledge coming from neurophysiology or from behavioural results hinting to a particu-
253 lar structure.

254 Putting “divisive normalisation” on a pedestal of a principle/framework as such, without considera-
255 tion of the network structure of interest giving the normalisation property, some researchers may ad-
256 vocate (Reynolds & Heeger, 2009; Carandini & Heeger, 2012), but it is not the unique possibility.

257 Concerns in Schallmo et al. (2018)'s model of motion perception and its application

258 In their work, Schallmo et al. (2018) proposed that the threshold of perception, the parameter that is
259 reported, is modelled through a simple rule of *Criterion* divided by model's *Response* written as
260 (Equation 3 in their Methods):

$$261 \text{Threshold} = \text{Criterion} / \text{Response} \quad (6)$$

262 This equation is claimed to be somehow related to the DDM of decision making (Ratcliff, 1978;
263 Ratcliff & McKoon, 2008; Forstmann, Ratcliff & Wagenmakers, 2016) (see also Link (1975); Link
264 (1992); Smith (2000)), and it seems to have been taken from the two earlier studies of Tadin and
265 Lappin (2005); Betts, Sekuler and Bennett (2012). The exact model description of how this equation
266 is obtained is not fully described in Schallmo et al. (2018), neither in the two other studies using it
267 (see **Appendix A2**), but instead refers to the study of Huk and Shadlen (2005). These later two au-
268 thors investigated the detection of RDKs with behavioural and neurophysiological measures in
269 monkeys on which they applied the DDM. But does the DDM predict a similar relation as in Equ-
270 ation 6? The interest in the drift-diffusion model is that it predicts the mean psychometric functions
271 for proportion responses and decision times (response times) (e.g. Luce (1986); Palmer, Huk and
272 Shadlen (2005)). The proposed equations are (Equations 4-5 directly taken from Huk and Shadlen
273 (2005)):

$$274 \begin{cases} P_+ = 1 / (1 + \exp(-2\mu A)) \\ \bar{t}_d = (A/\mu) \tanh(A\mu) \text{ if } \mu \neq 0, \text{ else } \bar{t}_d = A^2 \end{cases} \quad (7)$$

275 where A is the height of the boundary for decision making and μ is the drift rate. This later para-
276 meter is assumed to relate in a simple manner to model's low-level response, R , for example
277 $\mu(x) = R(x) \times \text{sign}(x)$ (x is stimulus intensity, see **Psychophysics**). If we take the equation of P_+ and
278 search for the x_{84} value giving $P_+ = P_{thr} = 0.84$, one obtains:

$$279 \begin{cases} R(x_{84}) = \frac{1}{2A} \ln(P_{thr} / (1 - P_{thr})) = \frac{1}{2A} \ln(21/4) \\ x_{84} = R^{-1} \left(\frac{1}{2A} \ln(21/4) \right) \end{cases} \quad (8)$$

280 where the last equation is obtained by assuming that model's low-level response $R(x)$ is invertible.
281 In the DDM result the stimulus *Threshold* is inversely related to the *boundary height* A , through
282 some non-linear inverse function of model's response. As such, Equation 6 is conflicting with the
283 derivation obtained from the psychometric function of the DDM. The argument about using Equ-
284 ation 6 seems the fact that for a fixed *Criterion* stronger neuronal responses lead to shorter time
285 *Thresholds*, and vice versa, thus providing a simple relation between neuronal responses and per-
286 ception (see **Appendix A2**). Overall, their model relating perceptual *Threshold* and low-level model
287 neuronal responses (Equation 6) is a new construct whose origin and derivation is not clearly given.

288 An important further concern in the model application done by Schallmo et al. (2018) is the way
289 they left their model's decision *Criterion* to differ between low and high contrast conditions (their
290 Figure 2C,G and Appendix 1 – Table 1, first column of results). Because all conditions of size and
291 contrast were presented randomly within the same block of measures (their page 12, Methods –
292 Paradigm and data analysis – Psychophysics: “Staircases were run separately to determine
293 thresholds for each of the six stimulus conditions (two contrasts x three sizes, as above). Condition
294 order was randomized across trials.”) and there was no *prior* information about what will be presen-
295 ted on each trial, the decision criterion cannot differ between conditions measured within the same
296 block. Allowing for different *Criteria* for different contrasts is not realistic because it means that the
297 observers fixed their decision criterion on a trial basis *after* each stimulus was presented. Therefore,
298 this is an issue that is conflicting with common ideas in detection theory.

General background for the modelling

Now we can turn to the modelling of the behavioural results in the simple experimental layout of D. Tadin. First let's recall the factors of interest in such studies, $\{size, contrast, duration\}$ of the stimulus, and the background hypothesis, that it is the excitatory-centre inhibitory-surround receptive field organisation in area MT/hMT+, which has a typical structure such that similar-to-centre motions in the surround reduce the activity due to centre stimulus presentation (Born & Bradley, 2005), that are a direct reason of the perceptual results obtain with motion stimuli.

In predicting psychophysical performance, the models are generally split into two independent stages, the low-level neuronal activities and the high-level decision stage. When combined they must predict the psychometric functions for each combination of size and contrast (see **Psychophysics**). The low-level activities are fed to the decision stage that only does one computation – to predict the dependent variable, here percent “rightward” responses.

An important point must be clarified concerning this experimental design and its modelling: the independent variable of the psychometric function is stimulus duration (t_{stim}). Because of this, there is a very strong assumption in the modelling that is made: the neuronal activity, R , of the motion tuned neurons are a direct measure of the duration of the stimulus, exactly as for stimulus contrast. That is, there is some monotonic relation between stimulus duration and R , the low-level model of neuronal firing rate. This point is a necessary condition for being able to model the psychometric function that is assumed monotonically increasing with stimulus duration. Let's assume that it has a classic saturating behaviour defined by a hyperbolic ratio equation (Albrecht & Hamilton, 1982) in the time domain:

$$R(t_{stim}) \propto t_{stim}^n / (t_{stim}^n + t_{50}^n) \quad (9)$$

with $n, t_{50} > 0$. This equation states that for an infinitely long presentation of the stimulus the neuronal activity is finite. Furthermore, here it is assumed that this component is multiplicative of the “interaction” components.

Concerning the decision stage model, the two most common, and currently largely advertised and used, models will be considered. The first one is the classic Signal Detection Theory (SDT) (Green & Swets, 1966; Macmillan & Creelman, 2005) applied on neuronal activities (e.g. Britten, Shadlen, Newsome and Movshon (1992)). The second one is the “Drift-Diffusion Model” (DDM) (Link, 1975; Ratcliff, 1978; Luce, 1986; Link, 1992; Smith, 2000; Smith & Ratcliff, 2004; Huk & Shadlen, 2005; Palmer et al., 2005; Ratcliff & McKoon, 2008; Forstmann et al., 2016) that Schallmo et al. (2018) associate to their decision stage.

The correct model for D. Tadin's design

After this long but necessary *aparté scientifique*, it is now possible to present a correct model predicting motion perception in D. Tadin's design. We take the low-level processing stage of the model to be represented by a population of neurons coding the two possible motions with opposite directions, whose activity R 's depend on all independent variables of stimulus contrast (c), size (s), and duration (t_{stim}). Here, both low-level models of Tadin and Lappin (2005) (but modified for having positive defined activities) and Schallmo et al. (2018), and both decision models (SDT and DDM) are considered, giving as model equation:

$$\begin{cases}
 R_{Tadin}(c, s, t_{stim}) = R_0 + S(E(s, c, t_{stim}) - I(s, c, t_{stim})) \\
 R_{Schallmo}(c, s, t_{stim}) = R_0 + E(s, c, t_{stim}) / (\sigma + I(s, c, t_{stim})) \\
 P_{SDT}(+|c, s, t_{stim}) = \int_0^{+\infty} N\left(\frac{R^+ - R^-}{\sqrt{Var(R^+) + Var(R^-)}}\right) dx \\
 P_{DDM}(+|c, s, t_{stim}) = \frac{1}{1 + \exp(-2A(R^+ - R^-))}
 \end{cases} \quad (10)$$

with $N(M)$ representing a normal distribution with mean M and unit variance, and R^+ and R^- representing the activity of neurons coding right- and leftward motion directions, respectively; $E()$ and $I()$ are the excitatory and inhibitory drives; R_0 is a “spontaneous firing rate” of the neurons (no input: $c=0$, or $s=0$, or $t_{stim}=0$; but see **Discussion**); $S()$ is the simple transducer from Equations 4-5; “+” sign in the psychometric function is defined as the rightward motions (see Figure 2); because the experimental design is a 1stim2AFC task with only two possible motion directions (see **Psychophysics**), SDT’s prediction is based on the difference of activities of the two motion coding populations, and $Var(R)$ is assumed equal to R (Fano factor of one); for the same reason the drift rate μ in the DDM (Equation 7) is taken as the difference between the activities of neurons with opposite motion directions.

Important note: in the above Equation 10, activities R s of the motion tuned neurons may represent the equilibrium state values for the given stimulus input parameters, as discussed in section **A note on normalization of neuronal activity**, or more generally a mean activity across some time window of post-stimulus presentation; starting from equation 10 model presentation can be considered as following the tradition of “pattern analyzers” (Graham, 2011), where it is usual to directly use static mathematical models with predefined pattern sensitive inputs.

What remains to be defined are the excitatory and inhibitory drives $E()$ and $I()$. For the moment we keep a general formulation. Because we assumed that neuronal response to stimulus duration, $R_{dur}(t_{stim})$, is independently pooled from responses to contrast and size, and that it has the same modulation on both components, we can write:

$$\begin{cases}
 E(c, s, t_{stim}) = R_{dur}(t_{stim}) \times R_{exc}(c, s) \\
 I(c, s, t_{stim}) = R_{dur}(t_{stim}) \times R_{inh}(c, s)
 \end{cases} \quad (11)$$

By using the threshold definition, $P(t_{stim}^{thr}) = 0.84$, one finds for each low-level model considered here that (see **Appendix A3**):

$$\begin{aligned}
 \text{Tadin \& Lappin : } R_{dur}(t_{stim}^{thr}) &= \frac{C_{Decision}}{R_{exc}(c, s) - R_{inh}(c, s)} \\
 \text{Schallmo et al. : } R_{dur}(t_{stim}^{thr}) &= \frac{\sigma C_{Decision}}{R_{exc}(c, s) - C_{Decision} R_{inh}(c, s)}
 \end{aligned} \quad (12)$$

The constants $C_{Decision}$ are dependent on the decision stage models of psychometric functions as follows:

$$\begin{cases}
 C_{Decision} \equiv C_{SDT} = \frac{1 + \sqrt{1 + 8R_0}}{2} \\
 C_{Decision} \equiv C_{DDM} = \frac{1}{2A} \ln(21/4)
 \end{cases} \quad (13)$$

But these constants are NOT the criteria for giving the response, neither some “*arbitrary response value that must be reached to make the perceptual judgment*” (see **Appendix A2**). They are simply mathematical constants that are obtained from the full model assumptions and derivations, and what

is defined as threshold for percent responses on the psychometric function (here 84%). In SDT case, the observer decision criterion is: if $R^+ - R^- > 0$ then give response “+”, else response “-”. The case of the DDM was presented in the previous section and it does contain the “boundary height A ”, but in that case the decision criterion is which of the two boundaries is reached (if “ $+A$ ” then give response “+”, if “ $-A$ ” give response “-”; for further details see the DDM articles and reviews cited earlier), and the decision “height” is in the denominator.

Making the last step and computing the duration threshold of the theoretical observer, by using equations 9 and 12, one obtains:

$$\begin{aligned} \text{Tadin \& Lappin : } t_{stim}^{thr} &= t_{50} \left(\frac{C_{Decision}}{R_{exc}(c, s) - R_{inh}(c, s) - C_{Decision}} \right)^{1/n} \\ \text{Schallmo et al. : } t_{stim}^{thr} &= t_{50} \left(\frac{\sigma C_{Decision}}{R_{exc}(c, s) - C_{Decision} R_{inh}(c, s) - \sigma C_{Decision}} \right)^{1/n} \end{aligned} \quad (14)$$

Therefore, although equation 6 proposed by Schallmo et al. (2018) has some similar form with the correct equation 14, its origin and interpretation are not satisfying. The correct general equations for modelling the psychometric functions of proportion responses as a function of stimulus parameters are given in equations 10, which from considerations about experimental threshold and model parametrisation leads to equation 14.

Comparing predictions of the two low-level models

Now that the derivations of the perceptual thresholds have been laid down, we can analyse whether these equations make interesting inferences about a particular model, or point to changes in thresholds as a function of excitatory or inhibitory strengths. Because the decision stage of the models make predictions that differ only in the constant $C_{Decision}$, from now on, any application of a full model will be done with SDT. Furthermore, we fix, without loss of generality, in Schallmo et al.’s model $\sigma=1$ (amplitude rescaling of $E()$ and $I()$ in Equation 10).

The two models give the predictions of duration threshold with equations 14. Interestingly, the condition $C_{SDT}=1$, which happens when $R_0=0$, and assuming that $R_{inh}(\text{Tadin})=R_{inh}(\text{Schallmo})$ and $R_{exc}(\text{Tadin})=R_{exc}(\text{Schallmo})$, gives exactly the same final prediction from both models, independent from the particular functional form of the excitatory and inhibitory drives. That is, despite the very different initial model assumptions (Equations 10), one finds that the final prediction of threshold variation as a function of stimulus size and contrast are equal.

When $R_0 > 0$, the exact difference between the two predicted thresholds depends on all three variables of excitatory, inhibitory and spontaneous firing rates. Nevertheless, even when $R_0 > 0$, one can see that if the two models differ only in the inhibitory responses R_{inh} , when $R_{inh}(\text{Tadin})=C_{SDT}R_{inh}(\text{Schallmo})$ both models still give exactly the same prediction of thresholds through a simple scaling of the inhibitory drives. Therefore, if the functions for excitatory and inhibitory effects in the two low-level models are chosen to be of the same mathematical form, then the two models can predict exactly the same thresholds by a proper rescaling of inhibitory amplitude.

To conclude, selecting either of the low-level models in Equation 10 is irrelevant for the final predictions. Consequently, it is the mathematical forms of the excitatory and inhibitory drives that predict the perceptual thresholds, not the “divisive normalisation” of Schallmo et al., neither the simple subtraction of Tadin & Lappin.

Application of the model

Until now, without taking a particular mathematical form of the contrast and size tuning properties, it was possible to make model analyses based on general considerations about the functions. Because the model predicts exact thresholds, and the deduced final equation 14 can be fit to the data (similarly to the previous applications of the resembling equation 6, see Tadin and Lappin (2005); Betts et al. (2012)), now we have to choose specific mathematical forms for the different components. By using Tadin & Lappin's low-level model, based on prior knowledge, and writing the full equation of the model, we have:

$$\begin{cases} t_{stim}^{thr} = t_{50} \left(\frac{C_{Decision}}{R_{exc}(c, s) - R_{inh}(c, s) - C_{Decision}} \right)^{1/n} \\ R_{exc}(c, s) = a_e \frac{c^p (1 - c_{1/2,e}^p)}{c^p (1 - 2c_{1/2,e}^p) + c_{1/2,e}^p} \operatorname{erf} \left(\frac{s}{(\sigma_e / (1 + m e^{-k/c}))} \right) \\ R_{inh}(c, s) = a_i \frac{c^p (1 - c_{1/2,i}^p)}{c^p (1 - 2c_{1/2,i}^p) + c_{1/2,i}^p} \operatorname{erf} \left(\frac{s}{\sigma_i} \right) \end{cases} \quad (15)$$

where $\{a_e, a_i\}$ are the absolute response amplitudes of the excitatory and inhibitory components, the middle, complex-looking, ratio of contrasts in $R_{exc}()$ and $R_{inh}()$ is the classic hyperbolic ratio for contrast response functions (CRF) (Albrecht & Hamilton, 1982) (but rewritten such that at $c=1$ its response is 1 and at $c=c_{1/2}$ its response is 1/2), and the last part is the error function as a function of size (for the inhibitory part) and size and contrast (for the excitatory part) (Tadin & Lappin, 2005) (see also **Appendix A4** explaining why in this model excitatory-centre tuning width must vary with contrast). In the above equations it is assumed that the pure contrast responses for both excitatory and inhibitory drives have the same exponent p but different half-amplitude constants $\{c_{1/2,e}, c_{1/2,i}\}$. The size function, $\operatorname{erf}()$, assumes that both receptive field components have a circular Gaussian shape and the space information is simply pooled across the stimulated area (stimulus centred on the receptive field). The model contains a total of 12 parameters: $\{t_{50}, n, C_{Decision}, a_e, a_i, p, c_{1/2,e}, c_{1/2,i}, \sigma_e, \sigma_i, m, k\}$. Without loss of generality, n and a_e are set to 1 (parameters $\{C_{Decision}, a_i\}$ are interpreted as proportional to the excitatory amplitude a_e). This leaves a total of 10 free parameters.

Example of model fit

A demonstration of fit to the data of Tadin and Lappin (2005) is carried. There are enough data points (32) on a sufficiently large range of contrasts and sizes for constraining all 10 parameters. Fit results are: $\{t_{50}=197, C_{Decision}=0.051, a_i=0.876, p=0.75, c_{1/2,e}=0.0044, c_{1/2,i}=0.081, \sigma_e=2.5, \sigma_i=3.6, m=7.78, k=0.305\}$ (see continuous curves in Figure 1A and compare to the data). Although some parameters seem to have stronger influence on a single input dimension (e.g. σ_i on size only), it is the 3D data, *Duration Threshold* vs $\{Contrast, Size\}$, that constrain all parameters simultaneously through the final non-linear model of Equation 15. Some interesting *quantitative* points can be derived from the final parameters: (1) inhibitory amplitude is quite strong, ~90% from the excitatory amplitude, which is a nice counterpart of neurophysiological results showing that short duration motion stimuli seem to preferentially stimulate neurons with strong surround inhibition (Churan, Khawaja, Tsui & Pack, 2008); (2) the constant $C_{Decision}=0.051$ gives, if $a_e=100, R_0 \sim 10$ which is about 10% of a_e , a rather common spontaneous firing rate of neurons (but see **Discussion**); (3) the excitatory half-amplitude constant $c_{1/2,e}$ is low, an interesting counterpart of neurophysiological observations for similar centre-surround designs showing low semi-saturation constants (Tsui & Pack, 2011); (4) the power of the contrast response function is very low ($p=0.75$), which was also reported to be the case in neuronal fits of size tuning data of MT (Tsui & Pack, 2011).

In summary, the correct computational model fit to the behavioural data provides *quantitative* parameters that globally match expectations about parameter values gathered from neurophysiological

Figure 3: (A) Fit of model to the data of Schallmo et al. (2018), data from their Figure 2C. (B) Model prediction of threshold changes when only one among the three parameters (a_i , $c_{1/2,i}$, σ_i) is changed such that inhibition is increased. "Data fit" curves are from (A). Red curves are for high contrast stimuli and blue curves are for low contrast stimuli.

studies. Thus, one can apply the model on behavioural data to infer putative changes of inhibitory or excitatory mechanisms under experimental manipulations, experimental manipulations that were performed by Schallmo et al. (2018) (although the data may contain some biases, see **Psychophysics**, here it is assumed that these biases correspond to a constant factor across all thresholds, thus simply scaling them all; this is done in the model by changing the global time constant, t_{50}).

The psychophysics measures of Schallmo et al. (2018)

In their study, the authors reported duration thresholds for a total of six different experimental conditions, 3 sizes and 2 contrasts. Therefore, the data does not allow any direct and exact model fit. Despite this, given the above prior knowledge and tests, we can make the assumption, based on the hypothesis of same origin of effects, that parameters should be globally similar between the different studies; especially here the excitatory part. Thus, we fix: $c_{1/2,e}=0.005$; $p=0.75$; $C_{Decision}=0.05$; $\sigma_e=2.5$; $m=7.8$ and $k=0.3$; parameter t_{50} is left to vary because it simply scales the overall data (on a log-scale it shifts the curves up or down). Thus, the interesting parameters about inhibition are left free: amplitude (a_i); half-amplitude constant ($c_{1/2,i}$), and inhibition tuning width (σ_i). A fit gave: $t_{50}=421$, $a_i=0.74$, $\sigma_i=7.7$, $c_{1/2,i}=0.072$; the results are depicted on Figure 3A. It shows that the data are globally explained, although some discrepancies at medium size stimuli and low contrast are present (this discrepancy hints to some model shortcomings, see **Appendix A5**).

Manipulating individual observer's level of inhibition

What should we expect as an outcome in this experimental design if the overall level of inhibition within a person is manipulated? In the model the effects of the inhibitory component are instantiated through three model functions, one of amplitude, one of contrast, and one of size, with their respective parameters $\{a_i, c_{1/2,i}, \sigma_i\}$. The effect of enhanced inhibition can affect all three functions. Thus, there are three simple ways to obtain stronger inhibition: an increase in amplitude (a_i), a decrease in contrast half-amplitude constant ($c_{1/2,i}$) such that inhibition is activated at lower contrasts, and a decrease in tuning width (σ_i) such that inhibition is activated earlier in spatial domain. These three simple examples are contrasted in Figure 3B to the results of the previous data fit (Figure 3A). As one can see, amplitude increase affects mainly thresholds at large stimulus sizes, half-amplitude decrease mainly affects low contrast responses and thus changes thresholds at low contrast, and inhibitory size decrease affects thresholds for small to medium size stimuli. One important result is that all three changes *increase* duration thresholds. This can easily be understood by referring to equation 15, where one can see that increasing inhibition ($R_{inh}(c,s)$) decreases the denominator which in turn increases the overall bracket. That is, while the exact threshold variation is dependent on the combined changes of inhibitory parameters, increase in inhibition makes overall thresholds higher.

The experiment of Schallmo et al. (2018) about "**Pharmacologically enhanced inhibition**" indeed show such a behaviour. That is, duration thresholds of the drug-intake condition increased when

compared to the placebo condition. Thus, contrary to their conclusion from the results (p. 7: “This result is not consistent with the idea that spatial suppression is directly mediated by neuronal inhibition.”), their results are consistent with the original hypothesis of drug-based inhibition enhancement of spatial suppression (if one assumes all other components are unchanged, see **Discussion**).

Comparing inhibition parameters between observers

In a last interesting experiment, the authors compared duration thresholds and non-invasive measures of GABA+ concentrations in visual cortices. They found that mean individual thresholds anti-correlated with the concentration of GABA+, obtained from Magnetic Resonance Spectroscopy, of the observer. This, as they stated it, is rather interesting because, as it was shown just above, stronger inhibition increases duration thresholds. Nevertheless, the exact contribution of the inhibitory component to the mean individual thresholds is unknown because the individual thresholds are dependent on all factors of excitatory and inhibitory drives, plus the individual’s half-amplitude time constant (t_{50} , see Equations 9 and 15) and observers’ spontaneous neuronal activity. In order to know each component of each observer, a full model approach is necessary with the corresponding measures.

Conclusion from modelling

From the above overall description few conclusions are drawn. First, the computational model predicting the duration thresholds in D. Tadin’s design is simple and easy to obtain. From careful considerations about its basic assumptions and functional form of the neuronal activities, it was shown that two “competing” models of neuronal activities, i.e. “divisive normalization” or simple subtraction of excitatory and inhibitory drives, cannot be dissociated in this design. Indeed, by an amplitude change of inhibitory component between the two models they allow to predict exactly the same thresholds. Thus, it is the exact mathematical form of the excitatory and inhibitory drives that defines the shape of suppressive and facilitative effects on duration thresholds as a function of size and contrast. Second, the computational model was shown to provide quantitative estimates of parameters that globally matched known neurophysiological counterparts. Third, the application of the model to the experimental design of Schallmo et al. (2018) showed (1) that in the experiment with pharmacologically manipulated inhibition the behavioural results of increased thresholds are consistent with stronger inhibition, and (2) that to know exactly the amount of inhibitory change under a given manipulation or between different observers it is necessary to have a full model approach.

Discussion

This Research Advance work analysed the spatial summation and suppression effects that raise in human perception of visual motion, specifically in the very simple stimulus design of D. Tadin (Tadin et al., 2003). It was prompted from the report of Schallmo et al. (2018) that (1) “divisive normalization” explains this paradigm’s results instead of more classic considerations about excitatory and inhibitory components, and (2) that their results seemed to show that suppression is not driven by inhibitory mechanisms. Therefore, this work presented the model for predicting the perceptual thresholds in the experimental design of D. Tadin and made multiple inferences from it. The conclusion from the model were presented in the relevant section. Here are discussed some points about psychophysics and modelling that were not developed further in the **Results** section.

When testing computational models of perception, we match the model prediction to the behavioural measures obtained in a particular experimental design through the use of the psychometric function (PF). Its exact form, and thus definition, depends on the experimental design. In the “1-stimulus-2-Alternatives-Forced-Choice” (1stim2AFC) design analysed here, it was argued that the PF should be defined as a function of motion duration and direction, thus representing a typical PF for “discrimination” (discriminating between two possible motion directions) that can be defined in

the full range of percent responses (0% to 100%). Because in this paradigm PF thresholds (inverse of slope) vary in a large range, researchers customary report results in pooling opposite motion directions that allows to plot the results on log-scale of stimulus intensity. It provides a simple visualisation of the threshold change that is seen as a shift of the PF on the log-scale data (this representation is similar to the classic “detection” results representation, e.g. typical contrast “detection” in 2stim2AFC designs). The current work allowed to model the psychometric function by considering the inputs and outputs that the model should incorporate and predict. It provided the general model equation (Equation 10), where the low-level model responses are presented together with the decision stage method, that led to predict the perceptual thresholds.

The model approach in this work is based on “static model”, or “pattern analyzers” (Graham, 2011). It approximates the responses of the neuronal populations to the inputs in specific ways by assuming that all neuronal interactions lead to simple “static” functional forms.

First, the major independent variable stimulus duration, t_{stim} , was modelled as activating neurons independently from the two other variables of stimulus contrast and size. The results show that it is a sufficient assumption. Furthermore, it was hypothesized that the exact way this variable affects neuronal responses is to modulate independently both excitatory and inhibitory drives. On the contrary, one can hypothesize that the effect is to modulate the combined response of excitatory and inhibitory components. While the “subtractive inhibition” model of Tadin and Lappin (2005) is not affected by this difference, the “divisive normalization” model of Schallmo et al. (2018) gives different predictions (see **Appendix A6**).

Second, the model incorporates the “spontaneous firing rate” of the neurons, R_0 , as the minimum possible activity. Nevertheless, spontaneous activity of the neurons are known to be lower for surround-suppressed cells in comparison to non-suppressed cells (Churan et al., 2008). This hints to the possibility that, when a moving stimulus is presented to them, the activity of the neurons sensitive to opposite motion directions is in fact lower than R_0 , an effect already reported (Snowden, Treue, Erickson & Andersen, 1991; Britten, Shadlen, Newsome & Movshon, 1993) and successfully modelled (Simoncelli & Heeger, 1998), and that it may vary with stimulus size and contrast. The model might be improved by including such an effect.

Third, it is known that the neuronal responses have contrast and size tuning such that at least one of the drives, excitatory or inhibitory, has intertwined contrast and size response functions. In the current work only the size tuning component of the excitatory drive was assumed to be changed with contrast variations. It is possible that instead of using size tuning to change with contrast, one can describe behavioural results by using the contrast tuning components to change with size variations. Such a possibility was used by Tsui and Pack (2011) for modelling their neurophysiological results and one might successfully apply such a model also to the behavioural results.

Last, if one manipulates inhibition, as performed by (Schallmo et al., 2018), the question naturally arises of how this experimental manipulation should be incorporated in the model. Here, it was argued only through the change of the inhibitory drive model. Since the model approach is based on “static” models, one cannot discard the argument that all the components of the model of “steady state” activity are modified. This is so because in changing connection weights of one type (excitatory or inhibitory) in the system affects the global equilibrium, and this later part is modelled, not the real weights. Therefore, because the experimental design of D. Tadin simultaneously measures the spatial and contrast domain characteristics of motion perception, the model necessarily must incorporate all these factors in its excitatory and inhibitory drives, which naturally leads to high multidimensional parameter space (here 10 parameters). The consequence is that, unless there is prior knowledge, the experimental measures must be carefully designed and carried in order to have data that can correctly constrain all parameters.

This work showed that we are very successful in providing simple and elegant modelling of behavioural results, results that demonstrate interesting perceptual phenomena that we associate to neuro-

physiological counterparts. But our modelling must be carefully considered and weighted with respect to prior knowledge. When neurophysiological, computational and behavioural results are combined in appropriate conditions for comparisons, we do gain important insights and knowledge about a given neurophysiological system and how it affects perception. In the case of suppression and facilitation of motion perception in humans, our understanding of how the percept is created and modulated has strong grounds from which we can make interesting inferences and, importantly, allow us to go toward more application oriented studies.

References

- Albrecht, D. G. and Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function, *Journal of Neurophysiology* 48 : 217-236.
- Albright, T. D. and Stoner, G. R. (2002). Contextual influences on visual processing, *Annual Review of Neuroscience* 25 : 339-379.
- Betts, L. R., Sekuler, A. B. and Bennett, P. J. (2009). Spatial characteristics of center-surround antagonism in younger and older adults., *J Vis* 9 : 1-15.
- Betts, L. R., Sekuler, A. B. and Bennett, P. J. (2012). Spatial characteristics of motion-sensitive mechanisms change with age and stimulus spatial frequency, *Vision Research* 53 : 1-14.
- Betts, L. R., Taylor, C. P., Sekuler, A. B. and Bennett, P. J. (2005). Aging reduces center-surround antagonism in visual motion processing., *Neuron* 45 : 361-366.
- Born, R. and Bradley, D. (2005). Structure and function of visual area MT, *Annual Review of Neuroscience* 28 : 157-189.
- Britten, K. H., Shadlen, M. N., Newsome, W. T. and Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance, *The Journal of Neuroscience* 12 : 4745-4765.
- Britten, K. H., Shadlen, M. N., Newsome, W. T. and Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals, *Visual Neuroscience* 10 : 1157-1169.
- Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation, *Nature Reviews Neuroscience* 13 : 51-62.
- Cavanaugh, J. R., Bair, W. and Movshon, J. A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons., *Journal of Neurophysiology* 88 : 2530-2546.
- Chang, J. J. and Julesz, B. (1984). Cooperative phenomena in apparent movement perception of random-dot cinematograms, *Vision Research* 24 : 1781-1788.
- Churan, J., Khawaja, F. A., Tsui, J. M. and Pack, C. (2008). Brief motion stimuli preferentially activate surround-suppressed neurons in macaque visual area MT, *Current Biology* 18 : R1051-R1052.
- Eagleman, D. M. (2001). Visual illusions and neurobiology, *Nature Reviews Neuroscience* 2 : 920-926.
- Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems, *Reports on Progress in Physics* 61 : 353-430.
- Forstmann, B., Ratcliff, R. and Wagenmakers, E.-J. (2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions, *Annual Review of Psychology* 67 : 641-666.

- 622 Garcia-Perez, M. A. and Alcala-Quintana, R. (2013). Shifts of the psychometric function:
623 distinguishing bias from perceptual effects, *The Quarterly Journal of Experimental Psychology* 66 :
624 319-337.
- 625 Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1
626 simple cells): Useful additions of the last 25 years, *Vision Research* 51 : 1397-1430.
- 627 Green, D. M. and Swets, J. A., 1966. Signal detection theory and psychophysics. John Wiley &
628 Sons, Inc., .
- 629 Grossberg, S. and Raizada, R. (2000). Contrast-sensitive perceptual grouping and object-based
630 attention in the laminar circuits of the primary visual cortex, *Vision Research* 40 : 1413-1432.
- 631 Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex., *Visual Neuroscience* 9 :
632 181-197.
- 633 Huk, A. C. and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal
634 integration of visual motion signals during perceptual decision making., *Journal of Neuroscience*
635 the Official Journal of the Society for Neuroscience 25 : 10420.
- 636 Kapadia, M. K., Westheimer, G. and Gilbert, C. D. (2000). Spatial distribution of contextual
637 interactions in primary visual cortex and in visual perception, *Journal of Neurophysiology* 84 :
638 2048-2062.
- 639 Kim, J. and Wilson, H. (1997). Motion integration over space: interaction of the center and surround
640 motion, *Vision Research* 37 : 991-1005.
- 641 Kingdom, F. A. A. and Prins, N., 2009. Psychophysics: a practical introduction. Academic Press, .
- 642 Levinson, E. and Sekuler, R. (1976). Adaptation alters perceived direction of motion, *Vision*
643 *Research* 16 : 779-781.
- 644 Link, S. W. (1975). The relative judgment theory of two choice response time, *Journal of*
645 *Mathematical Psychology* 12 : 114-135.
- 646 Link, S. W., 1992. The wave theory of difference and similarity. Lawrence Erlbaum Associates,
647 Hillsdale, New Jersey, .
- 648 Luce, R. D., 1986. Response Times. Oxford University Press, .
- 649 Macmillan, N. A. and Creelman, C. D., 2005. Detection theory: a user's guide. Lawrence Erlbaum
650 Associates, .
- 651 Marshak, W. and Sekuler, R. (1979). Mutual repulsion between moving visual targets, *Science* 205 :
652 1399-1401.
- 653 Morgan, M., Watamaniuk, S. N. and McKee, S. P. (2000). The use of an implicit standard for
654 measuring discrimination thresholds, *Vision Research* 40 : 2341-2349.
- 655 Murakami, I. and Shimojo, S. (1993). Motion capture changes to induced motion at higher
656 luminance contrasts, smaller eccentricities, and larger inducer sizes, *Vision Research* 33 : 2091-
657 2107.
- 658 Murakami, I. and Shimojo, S. (1996). Assimilation-type and contrast-type bias of motion induced
659 by the surround in a random-dot display: evidence for center-surround antagonism, *Vision Research*
660 36 : 3629-3639.
- 661 Nawrot, M. and Sekuler, R. (1990). Assimilation and contrast in motion perception: explorations in
662 cooperativity, *Vision Research* 30 : 1439-1451.

- 663 Palmer, J., Huk, A. C. and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and
664 accuracy of a perceptual decision, *Journal of Vision* 5 : 376-404.
- 665 Piëch, V., Li, W., Reeke, G. N. and Gilbert, C. D. (2013). Network model of top-down influences on
666 local gain and contextual interactions in visual cortex, *Proceedings of the National Academy of*
667 *Science USA* 110 : E4108-E4117.
- 668 Ratcliff, R. (1978). A theory of memory retrieval, *Psychological review* 85 : 59-108.
- 669 Ratcliff, R. and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice
670 decision tasks, *Neural Computation* 20 : 873-922.
- 671 Reynolds, J. H. and Heeger, D. J. (2009). The Normalization Model of Attention, *Neuron* 61 : 168-
672 185.
- 673 Sceniak, M. P., Ringach, D. L., Hawken, M. J. and Shapley, R. (1999). Contrast's effect on spatial
674 summation by macaque V1 neurons, *Nature Neuroscience* 2 : 733.
- 675 Schallmo, M.-P., Kale, A. M., Millin, R., Flevaris, A. V., Brkanac, Z., Edden, R. A., Bernier, R. A.
676 and Murray, S. O. (2018). Suppression and facilitation of human neural responses, *eLife* 7 : e30334.
- 677 Simoncelli, EP and Heeger, DJ (1998). A model of neuronal responses in visual area MT., *Vision*
678 *Research* 38 : 743-761.
- 679 Smith, P. L. (2000). Stochastic Dynamic Models of Response Time and Accuracy: A Foundational
680 Primer, *Journal of Mathematical Psychology* 44 : 408.
- 681 Smith, P. L. and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions, *Trends in*
682 *Neurosciences* 27 : 161-168.
- 683 Snowden, R., Treue, S., Erickson, R. and Andersen, R. (1991). The response of area MT and V1
684 neurons to transparent motion, *The journal of Neuroscience* 11 : 2768-2785.
- 685 Spillmann, L. and Dresp, B. (1995). Phenomena of illusory form: can we bridge the gap between
686 levels of explanation?, *Perception* 24 : 1333-1364.
- 687 Spillmann, L. and Ehrenstein, W. (1996). 43, From neuron to gestalt: Mechanisms of visual
688 perception. In: R. Greger, U. W. (Ed.), *Comprehensive Human Physiology*, Springer-Verlag Berlin
689 Heidelberg.
- 690 Spillmann, L. and Werner, J. (1996). Long-range interactions in visual perception, *Trends in*
691 *Neuroscience* 19 : 428-434.
- 692 Tadin, D. and Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast,
693 *Vision Research* 45 : 2059-2064.
- 694 Tadin, D., Lappin, J. S., Gilroy, L. A. and Blake, R. (2003). Perceptual consequences of center-
695 surround antagonism in visual motion processing, *Nature* 424 : 312-315.
- 696 Tsui, J. M. and Pack, C. C. (2011). Contrast sensitivity of MT receptive field centers and surrounds,
697 *Journal of Neurophysiology* 106 : 1888.
- 698 Tynan, P. and Sekuler, R. (1975). Simultaneous motion contrast: velocity, sensitivity and depth
699 response, *Vision Research* 15 : 1231-1238.
- 700 Tzvetanov, T. (2012). A single theoretical framework for circular features processing in humans:
701 orientation and direction of motion compared, *Frontiers in Computational Neuroscience* 6.
- 702 Tzvetanov, T. and Womelsdorf, T. (2008). Predicting human perceptual decisions by decoding
703 neuronal information profiles, *Biological Cybernetics* 98 : 397-411.

Watamaniuk, S., Sekuler, R. and Williams, D. (1989). Direction perception in complex dynamic displays: the integration of direction information, *Vision Research* 29 : 47-59.

Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea, *Vision Research* 30 : 1913-1921.

Wilson, H. R., 2005. Spikes, decisions, and actions: The dynamical foundations of neuroscience. Oxford University Press, .

Appendix

Appendix A1

Psychometric function definition, fitting and threshold estimation

As presented in the main text, Schallmo et al. (2018) defined the psychometric function between 50% and 100% correct responses on stimulus intensity range $-\infty$ to $+\infty$. They used the common ad-hoc Weibull function of the form:

$$P(|x|, a, b, \lambda = 0.04) = \frac{1}{2} [1 + (1 - \lambda) (1 - \exp(-(|x|/a)^b))] \quad (\text{A1-1})$$

with the two parameters $\{a, b\}$ defining the threshold (a) and slope (b), λ is observer's lapse rate, and x is the stimulus intensity (negative values are "leftward" motions).

A first issue is that the above definition states that the psychometric function might have a sigmoidal form (if b is much bigger than 1, from ~ 2 and above) on a linear scale of x that starts from zero. What it means is that below some absolute intensity of the stimulus, $|x_{low}|$, observer's response is near chance level. This value x_{low} depends on the "slope" parameter b of the Weibull function as well as the threshold a . The natural question coming from this definition is: If this effect exists, what is its origin? Because modelling of the psychometric function includes two levels, low- (stimulus processing) and high-level model (decision stage), there are two possibilities to consider. Either (1) this effect is based on some low-level sensory processing that must be carefully considered and incorporated in the model, either (2) this effect is a high-level decision stage process that should also somehow be incorporated.

It turns out that case (2) was already proposed and discussed for the experimental design of 1stim2AFC. Garcia-Perez and Alcalá-Quintana (2013) proposed that for very weak stimuli observers may have an *indecision interval* (they quote Fechner as the first to mention this "interval of uncertainty") where observers are not sure about having seen something relevant for deciding, and thus decide to guess. This allows to explain typical observations in the 1stim2AFC design of possible psychometric function shifts when observers always use the same response to respond in "undecided" trials. Because of the data processing method in Schallmo et al. (2018) (see main text) the above indecision model will shift the "Weibull" psychometric function toward higher stimulus intensities and steeper functions (higher slopes). Even in symmetric use of the two possible answers, when the interval of indecision is not zero it should create sigmoidal psychometric functions, with steeper slopes, than predicted by the simpler model of low-level + classic decision stage. Thus, it is possible that the ad-hoc Weibull function fits with high slopes b of the 1stim2AFC design do correspond to the model predicted data. Interestingly, because of their experimental methods (all stimuli presented within the same block), the above effect should be the same across all psychometric functions that were measured. Thus, one should expect that it corresponds to a scaling of the thresholds toward higher values, a scaling that is a simple shift on log-scale.

Overall, psychometric function fitting should be carried as discussed in section **Psychophysics** in order to avoid the above issues.

Appendix A2

Relation between the model of Tadin and Lappin (2005); Betts et al. (2012); Schallmo et al. (2018) and the drift-diffusion model

The proposed model for predicting duration Thresholds (Equation 6) seems to be related in some loose way to the drift-diffusion model without clear description of the way it was obtained, out of the fact that this rule allows to relate inversely duration Thresholds to neuronal activity of the model (R). Here are quotations from their articles where the authors describe and discuss their model and its putative origin.

In Tadin and Lappin (2005), the authors state (page 2063):

“Threshold (T) was taken to equal the number of times a response (R) needed to be repeated to reach a certain *Criterion*—essentially modeling “how long” the response needed to be maintained in order to generate sufficient evidence about the motion direction of the stimulus [emphasis by author TT]. This computation models the accumulation of evidence that likely underlies duration threshold measurements (see Footnote 1).”

And their Footnote 1 states:

“Use of duration thresholds was based on the assumption that if the neuronal response to a stimulus is weak and/or noisy, then longer stimulus exposure will be required for correct perception. More specifically, deciding whether an object is moving in one of two possible directions can be conceptualized as a process involving accumulation of sensory evidence over time (Gold & Shadlen, 2000; Roitman & Shadlen, 2002). When neuronal responses are noisy or attenuated, as with a highly suppressed motion stimulus, sensory evidence accumulates more slowly and a correct decision thus may require longer exposure duration (Roitman & Shadlen, 2002).”

Betts et al. (2012) state (page 2):

“The use of the stimulus duration threshold as a psychophysical measure of behavior is based on the idea that a perceptual decision regarding the direction of motion occurs only after a sufficient amount of information has been presented (i.e., after some response criterion has been surpassed) (Huk & Shadlen, 2005; Roitman & Shadlen, 2002). In this way, *weak and/or noisy signals require longer stimulus presentations to surpass the criterion for the perceptual response* [emphasis by author TT]. In the context of the model, the threshold, T , is the duration for which the response of strength R must be presented in order to surpass the Criterion (Tadin & Lappin, 2005). The *Criterion parameter determines the vertical placement of the model output on log-log coordinates, and inverts the independent variable such that increased responsiveness produces decreased duration thresholds* [emphasis by author TT].”

Last, Schallmo et al. (2018) state (page 13):

“To equate the response predicted by the model to motion discrimination thresholds, we assumed an inverse relationship between predicted response and discrimination time, such that:

$$\text{Threshold} = \text{Criterion} / \text{Response} \quad (3)$$

where Threshold is the amount of time needed to discriminate the direction of stimulus motion, *Criterion represents an arbitrary response value that must be reached to make*

the perceptual judgment [emphasis by author TT] (Huk and Shadlen, 2005), and Response is the predicted response rate (from Equation 1). This framework is consistent both with previous modeling efforts (Betts et al., 2012; Tadin and Lappin, 2005)[...]"

From the above snippets of their publications, it appears that (1) the decision stage model the authors advocate in their work has some loose relation to the DDM model and (2) this decision stage seems more a practical construct for predicting the dependent variable by the model.

Thus, despite tentatives to connect them, the comparison of DDM and their model that is carried in the main text shows that the two models are different.

Appendix A3

Derivation of Equation 12

From Equation 10 we can write:

$$R(c, s, t_{stim}) = R_0 + F(c, s, t_{stim}) \quad (A3-1)$$

where $F \equiv F(c, s, t_{stim})$ is the general non-specific function of the model. Then, using the SDT model, the definition of the threshold, and defining $F^+ \equiv F + R_0$ and $F^- \equiv R_0$, we have:

$$F(c, s, t_{stim}^{thr}) + R_0 - R_0 = \sqrt{F(c, s, t_{stim}^{thr}) + 2R_0} \quad (A3-2)$$

Which after simplification gives the second order equation and its solution:

$$\begin{aligned} F^2 - F - 2R_0 &= 0 \\ \Rightarrow F &= \frac{1 \pm \sqrt{1 + 8R_0}}{2} \equiv C_{SDT} \end{aligned} \quad (A3-3)$$

In this last equation the negative solution is not biologically meaningful and thus discarded. A similar to the last equation is obtained in the case of the DDM. From here, one derives the main text equations 12 and 14.

Appendix A4

Here is analysed the question of what kind of excitatory and inhibitory drives can, or can not, predict the behavioural results. One noteworthy point from Equations 12 and 14 is that they allow simple and interesting inferences, inferences that, to the best knowledge of the author, for the particular design of D. Tadin, were missing until now.

Independence of contrast and size effects can not predict behavioural results: simple case

If we make the assumption of independence of stimulus contrast and size effects on each drive, $R_{exc}(c, s) = R_{c, exc}(c)R_{s, exc}(s)$ (similar equation for inhibition), and that inhibitory contrast response function, $R_{c, inh}(c)$, is a scaled version of the excitatory one, that is, $R_{c, inh}(c) = kR_{c, exc}(c) \equiv kR_c(c)$, for all models one obtains:

$$R_{dur}(t_{stim}^{thr}) \propto 1/R_c(c) \quad (A4-1)$$

That is, because of the monotonic relation for each function on both sides of the above equation (both $R_{dur}(t_{stim})$ and $R(c)$ are monotonically increasing with their variable), duration thresholds are inversely related to contrast for any size of the stimulus. In fact, the behavioural results for small stimulus sizes follow this relation (see Figure 1), but not for large stimulus sizes. Thus, independent effects of stimulus contrast and size cannot predict the psychophysical observations of inverted relation between duration thresholds and contrast at large stimulus sizes. We can discard this possibility as incompatible with behavioural evidence. This result is a nice counterpart of the neurophysiolo-

gical findings in low-level visual processing systems, that stimulus contrast and size are somehow strongly intertwined in neuronal responses of areas V1 and MT (Sceniak, Ringach, Hawken & Shapley, 1999; Cavanaugh, Bair & Movshon, 2002; Tsui & Pack, 2011).

Independent contrast and size effects in the model: different, non-scaled, contrast tuning of excitatory and inhibitory components

Here is further analysed the condition of independent effect of contrast and size. Now it is assumed that $R_{c,inh}(c) \neq kR_{c,exc}(c)$, for example a simple shift of the inhibitory contrast response toward higher contrasts, and we assume that the components describing contrast responses are normalised such that their responses at $c=1$ is equal to one.

Then we can predict the inverted effect of contrast on duration thresholds as a function of stimulus size (e.g. taking the ratio of high-to-low contrasts of equation 14 one finds a simple relation between size and contrast functions and can deduce the relation at which the inversion happens; not shown). But the behavioural data show a very strong variation with contrast of the “minimum” (or dip) of the duration threshold as a function of stimulus size. In the data of Tadin and Lappin (2005) this minimum changes from ~2-3 degrees at $c=9\%$ down to ~0.5-0.6 degrees at contrasts 42% and 92% (see Figure 1A), a factor change of ~3-6. To demonstrate that this change is hardly predicted with independent contrast and size effects, we take for the size functions the error functions:

$$R(c, s) = R_{c,exc}(c)a_e\text{erf}(s/\sigma_e) - R_{c,inh}(c)a_i\text{erf}(s/\sigma_i) \quad (\text{A4-2})$$

Because the dip of the effect happens when the response $R(c,s)$ reaches a maximum on the size dimension (in Tadin & Lappin’s model; but in Scaglino et al.’s there is also the decision constant), we can compute the stimulus size at which the dip appears, s_{dip} , which gives:

$$s_{dip}(c) = \left(\frac{\sigma_e^2 \sigma_i^2}{\sigma_i^2 - \sigma_e^2} \ln \left(\frac{R_{c,exc}(c) a_e \sigma_i}{R_{c,inh}(c) a_i \sigma_e} \right) \right)^{1/2} \quad (\text{A4-3})$$

From this result, it is possible to find the necessary ratio of excitatory to inhibitory low-contrast responses for observing a given dip change as a function of the ratio of high-contrast responses. By using the ratio between size dips at low and high contrasts, $r_{dip} = s_{dip}(lc)/s_{dip}(hc)$ (lc – low contrast; hc – high contrast), it is found that:

$$\frac{R_{c,exc}(lc)}{R_{c,inh}(lc)} = \left(\frac{a_e \sigma_i}{a_i \sigma_e} \right)^{r_{dip}^2 - 1} \left(\frac{R_{c,exc}(hc)}{R_{c,inh}(hc)} \right)^{r_{dip}^2} \quad (\text{A4-4})$$

If we take the best case of $R_{c,inh}(hc) = R_{c,exc}(hc)$, e.g. at contrast of one the normalised contrast coding components are equal, and $a_e = a_i$ (maximum surround inhibition), then we have:

$$\frac{R_{c,exc}(lc)}{R_{c,inh}(lc)} = \left(\frac{\sigma_i}{\sigma_e} \right)^{r_{dip}^2 - 1} \quad (\text{A4-5})$$

That is, the ratio of excitatory-to-inhibitory low-contrast responses is a function of the ratio inhibitory-to-excitatory size tunings to the power of squared r_{dip} ! This model can predict factor of dip changes of ~3–6 only if the inhibitory responses at low contrasts are essentially zero ($R_{c,inh}(c) = (\sigma_i/\sigma_e)^{-35} R_{c,exc}(c) - (\sigma_i/\sigma_e)^{-8} R_{c,exc}(c)$). This would correspond to a strongly expansive inhibitory contrast response function at intermediate to high contrasts, and thus explains the observation that this model does not correctly predict the behavioural data (Tadin & Lappin, 2005).

To conclude, the model of low-level responses cannot have independent contrast and size effects. Some form of contrast dependent size tuning widths is necessary to predict the behavioural results, as argued by Tadin and Lappin (2005) and reported neurophysiologically (Sceniak et al., 1999; Cavanaugh et al., 2002).

868 Necessity of centre excitatory tuning width to vary with contrast

869 One can further analyse which single component, excitatory or inhibitory, influences the most the
870 position of the dip. We can again use Equation A4-3 and rewrite it as:

$$871 \quad s_{dip}(c) = \sigma_e \left(\frac{k_s^2}{k_s^2 - 1} \ln \left(\frac{R_{c,exc}(c) a_e k_s}{R_{c,inh}(c) a_i} \right) \right)^{1/2} \quad (A4-6)$$

872 where $k_s = \sigma_i / \sigma_e$ is the ratio of inhibitory-to-excitatory size tuning. From this equation it is much
873 easier to see that, at fixed ratio k_s , the dip position is linearly related to σ_e , while at fixed σ_e it is
874 much slowly changing with σ_i . Thus, the behavioural data of duration thresholds showing large
875 changes of the dip position can only be explained if excitatory centre tuning width is contrast de-
876 pendent. The other factors of “pure” contrast responses and inhibitory tuning widths depending on
877 contrast do not allow for large changes in dip positions.

878 Appendix A5

879 Measures of spatial suppression and summation with sine gratings do not conform to the 880 model

881 The data from the RDKs of Tadin and Lappin (2005) presented in Figure 1 are well fit by the
882 model. In an attempt to explain also other data, it was found that the model has difficulties in fitting
883 them. These data were obtained with moving sine grating stimuli, that were seen through Gaussian
884 envelopes (Gabor patches). The fitting was unsuccessful in matching the model to the data. When
885 the 10 parameters were left free the final best fit gave biologically not plausible values; even with
886 these best parameters, the fit was still difficult to reconcile with the data. The reason of this discrep-
887 ancy can be seen in Figure A1. Panels A-C replots the data from three studies that used moving sine
888 gratings at different contrasts and sizes. As it can be seen, the grating data has a particular beha-
889 viour: at low contrasts (~3%) the thresholds have the typical decrease with increasing stimulus size;
890 at slightly higher contrasts (4.2-5.5%) the data show a U-shape with increasing stimulus size; from
891 contrasts above 10% the thresholds exhibit mainly monotonic increasing behaviour with possible
892 plateau effect at very high contrasts (panel C). Panel D presents the model predictions. It can be
893 seen that the model shows a typical U-shape of thresholds vs. stimulus size for all but the lowest
894 contrasts of 2.8-5.5%, where it is very weak. One can use Equation A3-6 to see that at very high
895 contrasts one has approximately:

$$896 \quad s_{dip}(c) = K \sigma_e(c) \quad (A5-1)$$

with K approximately a constant (k_s increases with contrast and the ratio of contrast functions slowly decreases toward 1, which makes the bracket in Equation A4-6 to increase very slowly with increasing contrast, in opposite direction of s_{dip} variation). Thus, one has the choices: (i) to change the function relating excitatory centre size and contrast such that at high contrasts it drops much rapidly as in the data, or instead (ii) suppose that the valid fit in Figure 1 hints to model inadequacy for explaining results with grating type of stimuli.

The data at low contrast of Schallmo et al. (2018)

The above short section showed that there are difficulties in matching the model to grating type of stimuli. In analysing Schallmo et al. (2018)'s data, it is apparent that their results for low contrast condition exhibited a different behaviour, a "U-shape" for thresholds vs. stimulus size; this U-shape is very pronounced in their Figures 4 and 5. This seems to match the U-shape results for slightly higher contrasts reported in the section above. Given that the stimuli of Schallmo et al. (2018) differ in the window type used (circular window with blurred edges), it is difficult to make comparison between the studies using moving gratings. Despite this, it should be noted that leaving the full model (10 parameters) to fit the 6 data points of Figure 2C in Schallmo et al. (2018) also gives excellent fit but with biologically not meaningful values (results not shown); a similar outcome as when trying to apply the model to Gabor patches stimuli.

Thus, it seems sine grating type of stimuli have peculiar results that are not explainable with the model derived in the main text.

Appendix A6

Schallmo et al. (2018)'s model with global modulation by $R_{dur}(t_{stim})$

In Equations 10-11, it was assumed that $R_{dur}(t_{stim})$ modulates independently each of the two drives. Instead, one can assume that $R_{dur}(t_{stim})$ modulates the full response to the combined drives. In this case, Schallmo et al. (2018)'s model provides slightly different equations, which are presented here.

The low-level response of the model becomes:

$$R(c, s, t_{stim}) = R_0 + R_{dur}(t_{stim}) \times R_{exc}(c, s) / (\sigma + R_{inh}(c, s)) \quad (A6-1)$$

Remarque: one may think to re-interpret the above equation as: "only the excitatory drive is modulated by stimulus duration"; if one recalls that this is a "static" model of the final "equilibrium/steady state" activities, the interpretation should not be mistaken between the two possible implementations.

From equation A6-1 and the decision stage model, at threshold one obtains:

$$R_{dur}(t_{stim}^{thr}) = C_{Decision}(\sigma + R_{inh}(c, s)) / R_{exc}(c, s) \quad (A6-2)$$

Which, by using Equation 9, gives:

$$t_{stim}^{thr} = t_{50} \left(\frac{C_{Decision}(\sigma + R_{inh}(c, s))}{R_{exc}(c, s) - C_{Decision}R_{inh}(c, s) - \sigma C_{Decision}} \right)^{1/n} \quad (A6-3)$$

This last equation can be compared to the prediction of the main text (equation 14). This equation has a main effect to create inhibition effects on duration threshold much stronger than the main text model. This equation also provides a fit of the data of Tadin and Lappin (2005) (not shown). As for the main models (see **Appendix A5**) this model also succeeds to fit the Gabor patches results, but with biologically unrealistic parameters (not shown). No further analyses are provided here, neither detailed comparisons with the main text model.