# Visual illusions via neural dynamics: Wilson-Cowan-type models and the efficient representation principle

Marcelo Bertalmio, Luca Calatroni, Valentina Franceschi, Benedetta Franceschiello, Alexander Gomez-Villa, Dario Prandi

# Visual illusions via neural dynamics: Wilson-Cowan-type models and the efficient representation principle

Marcelo Bertalmío
DTIC, Universitat Pompeu Fabra, Barcelona, Spain
marcelo.bertalmio@upf.edu

Luca Calatroni
CMAP, CNRS, École Polytechnique,
Institut Polytechnique de Paris, Palaiseau, France
luca.calatroni@polytechnique.edu

Valentina Franceschi
IMO, Université Paris-Sud, Orsay, France
valentina.franceschi@u-psud.fr

Benedetta Franceschiello
FAA, LINE, Radiology, CHUV, Lausanne, Switzerland
benedetta.franceschiello@fa2.ch

Alexander Gomez-Villa
DTIC, Universitat Pompeu Fabra, Barcelona, Spain
alexander.gomez@upf.edu

Dario Prandi
CNRS, L2S, CentraleSupelec, Gif-sur-Yvette, France
dario.prandi@l2s.centralesupelec.fr

**Abstract**

In this work we have aimed to reproduce supra-threshold perception phenomena, specifically visual illusions, with Wilson-Cowan-type models of neuronal dynamics. We have found that it is indeed possible to do so, but that the ability to replicate visual illusions is related to how well the neural activity equations comply with the efficient representation principle. Our first contribution is to show that the Wilson-Cowan equations can reproduce a number of brightness and orientation-dependent illusions, and that the latter type of illusions require that the neuronal dynamics equations consider explicitly the orientation, as expected. Then, we formally prove that there can't be an energy functional that the Wilson-Cowan equations are minimizing, but that a slight modification makes them variational and yields a model that is consistent with the efficient representation principle. Finally, we show that this new model provides a better reproduction of visual illusions than the original Wilson-Cowan formulation.

**Keywords**: Wilson-Cowan equations; Brightness perception; Efficient representation principle; Variational modelling

## 1 Introduction

The goal of this work is to point out the intimate connections existing between three popular approaches in vision science: the Wilson-Cowan equations, the study of visual illusions, and the efficient representation theory.

As other articles in this special issue make abundantly clear, Wilson-Cowan equations have a long and successful story of modelling cortical low-level dynamics [Cowan et al., 2016]. Nonetheless, not many works have pursued the study of psychophysics with Wilson-Cowan equations (e.g. [Adini et al., 1997, Herzog et al., 2003, Bertalmío and Cowan, 2009, Ernst et al., 2016, Bertalmío et al., 2017]), and we are not aware of publications dealing with suprathreshold brightness perception through the Wilson-Cowan formulation.

The study of visual illusions has always been key in the vision science community, as the mismatches between reality and perception provide insights that can be very useful to develop new models of visual perception [Kingdom, 2011] or of neural activity [Eagleman, 1959, Murray and Herrmann., 2013], and also to validate the existing ones. It is commonly accepted that visual illusions arise due to neurobiological constraints [Purves et al., 2008] that modify the underpinned mechanisms of the visual system.

The efficient representation principle, introduced by Attneave [Attneave, 1954] and Barlow [Barlow et al., 1961], states that neural responses aim to overcome these neurobiological constraints and to optimize the limited biological resources by being tailored to the statistics of the images that the individual typically encounters, so that visual information can be encoded in the most efficient way. This principle is a general strategy observed across mammalian, amphibian and insect species [Smirnakis et al., 1997] and is embodied by neural processing according to abundant experimental evidence [Fairhall et al., 2001, Mante et al., 2005, Benucci et al., 2013].

Our work aims at pulling together the three approaches just mentioned, providing a more unified framework to understand vision mechanisms. First, we show that the Wilson-Cowan equations are able to reproduce a number of visual illusions. Secondly, we formally prove that Wilson-Cowan equations (with constant input) are not variational, in the sense that they are not minimizing any energy functional. Next, we detail how a simple modification turning the Wilson-Cowan equations variational yields a local histogram equalisation method that is consistent with the efficient representation principle. We finally show how this new formulation provides a better reproduction of visual illusions than the Wilson-Cowan model.

# 2 Materials and methods

## 2.1 Visual illusions

Visual illusions have always been considered as a window between reality and perception, enabling neuroscientists to disentangle the complicated process of vision [Eagleman, 1959, Murray and Herrmann., 2013]. Computational models able to reproduce perceptual phenomena represent very effective methods to test new hypotheses and generate new insights, both for neuroscience and applied disciplines such as image processing. These illusions can be classified according to the main feature detection mechanisms involved during the visual process [Shapiro and Todorovic, 2016]. In this contribution we considered two main groups of visual illusions to assess the efficacy of our model in reconstructing the perceptual process: *brightness illusions* and *orientation-dependent illusions*.

### 2.1.1 Brightness illusions

Brightness illusions are a class of phenomena where image regions with the same gray level are perceived as having different brightness, depending on the shapes, arrangement and gray level of the surrounding elements. Fig. 1 shows the nine brightness illusions we have chosen to perform tests on in this paper. They are all very popular and at the same time they represent a diverse set, as we can see from the following descriptions.

**White's illusion:** the left gray rectangle appears darker than the right one, while both are identical [White, 1979] (Fig. 1(a)).

**Simultaneous brightness contrast:** the left gray square appears lighter than the right one, while both are identical [Brucke, 1865] (Fig. 1(b)).

**Checkerboard illusion:** the mid-gray square in the fifth column appears darker than the one in the seventh column, while both are identical [DeValois and DeValois, 1990] (Fig. 1(c)).

**Chevreul illusion:** a pattern of homogeneous bands of increasing intensity from left to right is presented. However, the bands in the image are perceived as inhomogeneous, i.e. darker and brighter lines appear at the borders between adjacent bands [Ratliff, 1965] (Fig. 1(d)).

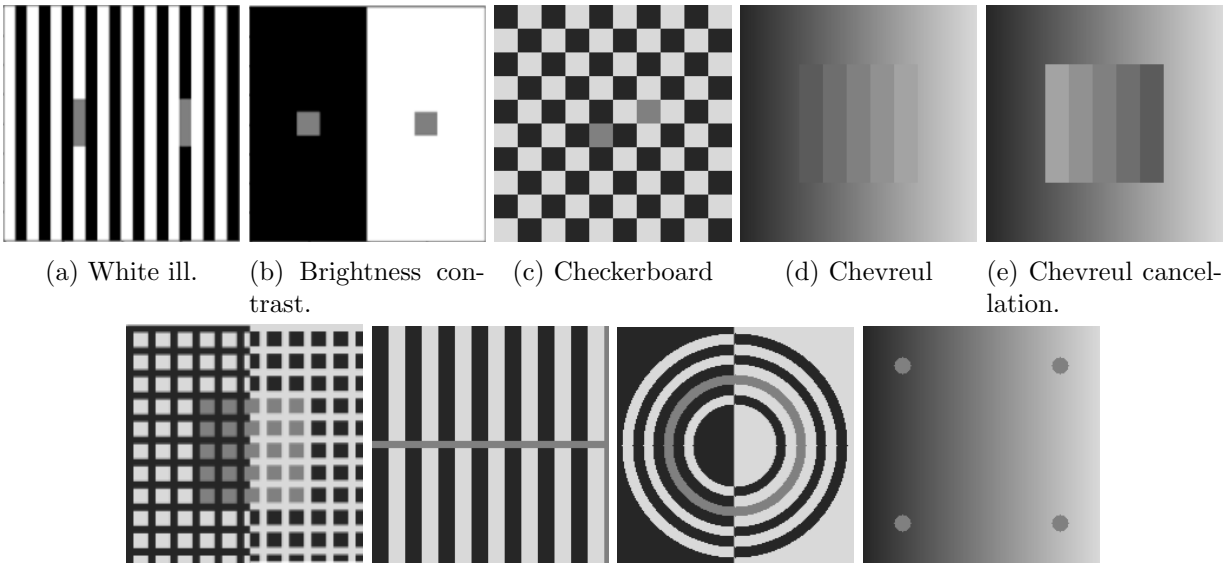
**Chevreul cancellation:** when the order of the bands is reversed, now decreasing in intensity from left to right, the effect is cancelled [Geier and Hudák, 2011] (Fig. 1(e)).

**Dungeon illusion:** two gray rectangles are perceived as darker or lighter depending on the gray intensities of both the background and the grid, see [Bressan, 2001]. The left rectangle is perceived as darker than the one on the right (Fig. 1(f)).

**Grating induction:** the background grating (which can be tuned to different orientations) induces the appearance of a counter-phase grating in the homogeneous gray horizontal bar [McCourt, 1982] (Fig. 1(g)).

**Hong-Shevell illusion:** the mid-gray half-ring on the left appears darker than the one on the right, while both are identical [Hong and Shevell, 2004] (Fig. 1(h)).

**Luminance illusion:** four identical dots over a background where intensity increases from left to right, and the dots on the left are perceived being lighter than the ones on the right [Kitaoka, 2006] (Fig. 1(i)).



(a) White ill. (b) Brightness contrast. (c) Checkerboard (d) Chevreul (e) Chevreul cancellation.

(f) Dungeon ill. (g) Grating induction. (h) Hong-Shevell ill. (i) Luminance ill.

Figure 1: From left to right, top to bottom: White's illusion, Brightness contrast, the checkerboard illusion, the Chevreul illusion, Chevreul cancellation, the Dungeon illusion, the Grating induction, the Hong-Shevell illusion and the Luminance illusion.

### 2.1.2 Orientation-dependent illusions

We also consider orientation-dependent illusions, where the perceptual phenomena (e.g. in terms of brightness or contrast) is affected by the orientation of the image elements.

**Poggendorff illusion.** The Poggendorff illusion, presented in the modified version considered in this work in Fig. 2(a), is a very well known geometrical optical illusion in which the presence of a central surface induces a misalignment of the background lines. This illusion depends both on the orientation of the background lines and the width of the central surface [Weintraub and Krantz, 1971], as the more the angle is close to $\pi/2$ the less is the bias, but in this example the perceived bias is also dependent on the brightness contrast between central surface and background lines.

**Tilt illusion.** The Tilt illusion is a phenomenon where the perceived orientation of a test line or grating is altered by the presence of surrounding lines or grating with a different orientation. In our case we consider the effect that the orientation of a surround grating pattern has on the perceived contrast of a grating pattern in the center: the inner circles in Figs. 2(b) and 2(c) are identical but the latter is perceived as having more contrast than the former.

(a) Poggendorff illusion. (b) Tilt illusion, same $\theta$. (c) Tilt illusion, different $\theta$.

Figure 2: From left to right: a modified version of the Poggendorff illusion based on Grating Induction, a modified Tilt illusion with concentric circles having the same orientation and a modified Tilt illusion with concentric circles having different orientations.

## 2.2 Wilson-Cowan-type models for contrast perception

In this section we introduce four different evolution equations derived from the Wilson-Cowan formulation, that will be studied in this paper. We recall that, denoting by $a(x,t)$ the state of a population of neurons with spatial coordinates $x \in \mathbb{R}^2$ at time $t > 0$, the Wilson-Cowan equations proposed in [Wilson and Cowan, 1972] can be written[1] as

$$\frac{\partial}{\partial t}a(x,t) = -\beta a(x,t) + \nu \int_{\mathbb{R}^2} \omega(x\|y)\sigma(a(y,t))\,dy + h(x), \tag{2.1}$$

where $\beta > 0$ and $\nu \in \mathbb{R}$ are fixed parameters, $\sigma : \mathbb{R} \to \mathbb{R}$ is a non-linear sigmoid saturation function, the kernel $\omega(x\|y)$ models interactions at two different spatial locations $x$ and $y$ (we will assume that the integral of $\omega$ is normalised to 1) and $h$ is the input signal.

### 2.2.1 Wilson-Cowan equations do not fulfill any variational principle

Over the last thirty years, the use of variational methods in imaging has become increasingly popular as a regularisation strategy for solving general ill-posed imaging problems in the form

$$\text{find } u \quad \text{s.t.} \quad f = \mathcal{T}(u). \tag{2.2}$$

[1] In [Wilson and Cowan, 1972] the sigmoid function is applied outside of the integral term and not only on the activity $a(y,t)$ as in (2.1). This corresponds to an "activity-based" model of neuron activation, while (2.1) corresponds to a "voltage-based" one. See [Faugeras, 2009], where the two models are shown to be equivalent.

Here, $f$ represents a given degraded image and $\mathcal{T}$ a (possibly non-linear) operator describing the degradation (e.g. noise, blur, under-sampling, etc.)

Due to the lack of fundamental properties such as existence, uniqueness and stability of the solution of the problem (2.2), the idea of regularisation consists of incorporating *a priori* information on the desired image $u_\star$ and on its closeness to the data $f$ by means of suitable variational terms. This gives rise, in particular, to variational methods where one looks for an approximation $u_\star$ of the real solution $u$ by solving

$$u_\star = \arg\min \mathcal{E}(u), \tag{2.3}$$

where $\mathcal{E}$ is the energy functional combining regularisation and data fit, depending also on the given image $f$. A popular way to solve the variational problem consists in finding $u_\star$ as the steady-state solution of the evolution equation given by the gradient descent of the energy functional

$$\frac{\partial}{\partial t}u = -\nabla\mathcal{E}(u), \qquad u|_{t=0} = f, \tag{2.4}$$

under appropriate conditions on the boundary of the image domain.

In the context of vision science, evolution equations have been originally used as a tool to describe the physical transmission, diffusion and interaction phenomena of stimuli in the visual cortex [Beurle and Matthews, 1956, Wilson and Cowan, 1972, Wilson and Cowan, 1973]. Variational methods are the main tool of ecological approaches, that pose the efficient coding problem [Olshausen and Field, 2000] as an optimisation problem to be solved with evolution equations that minimise an energy functional [Atick, 1992] involving natural image statistics and biological constraints. The resulting solution is optimal because it has minimal redundancy.

However, we must remark that, while considering the gradient descent of an energy functional gives always an evolution equation, the reverse is not true: not every evolution equation is minimising an energy functional. In fact, this is the case for the Wilson-Cowan equations, which do not fulfil any variational principle, as we prove in Appendix A. As a consequence, they are sub-optimal in reducing the redundancy.

### 2.2.2 A modification of the Wilson-Cowan equations complying with efficient representation

Remarkably, the efficient representation principle has correctly predicted a number of neural processing aspects and phenomena like the photoreceptor response performing histogram equalisation, the dominant features of the receptive fields of retinal ganglion cells (lateral inhibition, the switch from bandpass to lowpass filtering when the illumination decreases, and, remarkably, colour opponency, with $L$, $M$ and $S$ signals being highly correlated but $L+M$, $L-M$ and $S-(L+M)$ having quite low correlation), or the receptive fields of cortical cells having a Gabor function form [Atick, 1992, Daugman, 1985, Olshausen and Field, 2000]. Efficient representation is the only framework able to predict the functional properties of neurons from a simple principle, and given how simple the assumptions are it's really surprising that this approach works so well [Meister and Berry, 1999].

In [Bertalmío and Cowan, 2009] Bertalmío and Cowan showed how a slight modification of the Wilson-Cowan formulation leads to a variational model, as we now present. Assuming that the activity signal $a$ is in the range $[0, 1]$, we can re-write equation 2.1 in terms of a sigmoid $\hat{\sigma}$ shifted by $\frac{1}{2}$ (which we take as the average signal value) and inverted in signs, thus getting:

$$\frac{\partial}{\partial t}a(x,t) = -\beta a(x,t) - \nu\int_{\mathbb{R}^2}\omega(x\|y)\hat{\sigma}\left(a(y,t)-\frac{1}{2}\right)dy + h(x). \tag{2.5}$$

Note that this is just a re-writing of equation (2.1), so it is still not associated to any variational method. However, if we now assume $\hat{\sigma}$ to be odd and replace the $\frac{1}{2}$ term by $a(x,t)$, we obtain

$$\frac{\partial}{\partial t}a(x,t) = -\beta a(x,t) + \nu\int_{\mathbb{R}^2}\omega(x\|y)\hat{\sigma}(a(x,t)-a(y,t))\,dy + h(x), \tag{2.6}$$

and this equation is now a gradient descent equation, as it does fulfil a variational principle.

Furthermore, under the proper choice of parameters $\beta, \nu$ and input signal $h$, this evolution equation performs local histogram equalisation (LHE) [Bertalmío et al., 2007]. This is key for our purposes, since, as Atick points out [Atick, 1992], one of the main types of redundancy or inefficiency in an information system like the visual system happens when some neural response levels are used more frequently than others, and for this type of redundancy the optimal code is the one that performs histogram equalisation.

It is therefore expected that the modification of the Wilson-Cowan equations in (2.6), which better complies with the efficient representation principle, should be more effective in reducing redundancy than the original Wilson-Cowan model of equation (2.1).

### 2.2.3 Accounting for orientation

Models (2.1) and (2.6) ignore orientation and as such they are not well-suited to explain a number of visual phenomena. For this reason, following [Bertalmío et al., 2019], we extend them to a third dimension, representing local image orientation, as follows. We let $La : Q \times [0, \pi) \to \mathbb{R}$ be the cortical activation in V1 associated with the signal $a$, so that $La(x, \theta)$ encodes the response of the neuron with spatial preference $x$ and orientation preference $\theta$ to $a$. Mathematically, such activation is obtained via a suitable convolution with the receptive profiles of V1 neurons, as explained in Appendix B, see also [Duits and Franken, 2010, Petitot, 2017, Prandi and Gauthier, 2017, Citti and Sarti, 2006, Sarti and Citti, 2015]. Then, denoting for any $t > 0$ by $A(x, \theta, t)$ the cortical response at time $t$, the natural extension of equations (2.1) and (2.6) to the orientation dependent case is given by the two models:

$$\frac{\partial}{\partial t}A(x,\theta,t) = -\beta A(x,\theta,t) + \nu \int_0^\pi \int_Q \omega(x,\theta\|y,\phi)\sigma\Big(A(y,\phi,t)\Big)\, dy\, d\phi + Lh(x,\theta), \tag{2.7}$$

$$\frac{\partial}{\partial t}A(x,\theta,t) = -\beta A(x,\theta,t) + \nu \int_0^\pi \int_Q \omega(x,\theta\|y,\phi)\hat{\sigma}\big(A(x,\theta,t) - A(y,\phi,t)\big)\, dy\, d\phi + Lh(x,\theta), \tag{2.8}$$

where $Lh(x, \theta)$ denotes the cortical activation in V1 corresponding to the visual input $h$ at spatial location $x$ and orientation preference $\theta$. We remark that these models describe the dynamic behaviour of activations in the 3D space of positions and orientation. As explained in Appendix B, once a stationary solution is found, the two-dimensional perceived image can be found by simply applying the formula

$$a(x) = \frac{1}{\pi}\int_0^\pi A(x,\theta)\, d\theta. \tag{2.9}$$

### 2.2.4 Models under consideration

We summarise here the four models we are going to test in the following sections. The orientation-independent WC and LHE models are:

$$\frac{\partial}{\partial t}a(x,t) = -(1+\lambda)a(x,t) + \frac{1}{2M}\int_Q \omega(x,y)\sigma\left(a(y,t)\right)\, dy + \lambda f_0(x) + \mu(x) \tag{WC-2D}$$

$$\frac{\partial}{\partial t}a(x,t) = -(1+\lambda)a(x,t) + \frac{1}{2M}\int_Q \omega(x,y)\hat{\sigma}\left(a(x,t) - a(y,t)\right)\, dy + \lambda f_0(x) + \mu(x), \tag{LHE-2D}$$

which relate to (2.1) and (2.6) by simply choosing parameters as $\beta = 1 + \lambda$ and $\nu = 1/2M$ where $M > 0$ is a normalisation constant and input signal $h(x) = \lambda f_0(x) + \mu(x)$, where $\lambda > 0$, $f_0(x)$ is the local intensity at $x \in Q$ of given image $f_0$ and $\mu(x)$ denotes a local average of the initial stimulus $f_0$ around $x$ (a choice motivated by the averaging behaviour of M cells and already considered in similar models e.g. [Bertalmío et al., 2007, Bertalmío, 2014]).

The orientation-dependent WC and LHE models can be similarly written as:

$$\begin{aligned}\frac{\partial}{\partial t}A(x,\theta,t) =& -(1+\lambda)A(x,\theta,t) + \frac{1}{2M}\int_0^\pi \int_Q \omega(x,\theta||y,\phi)\sigma\left(A(y,\phi,t)\right)\,dy\,d\phi \\ &+ \lambda L f_0(x,\theta) + L\mu(x,\theta), \qquad \text{(WC-3D)}\\ \frac{\partial}{\partial t}A(x,\theta,t) =& -(1+\lambda)A(x,\theta,t) + \frac{1}{2M}\int_0^\pi \int_Q \omega(x,\theta||y,\phi)\hat{\sigma}\left(A(x,\theta,t) - A(y,\phi,t)\right)\,dy\,d\phi \\ &+ \lambda L f_0(x,\theta) + L\mu(x,\theta), \qquad \text{(LHE-3D)}\end{aligned}$$

which can analogously be related to (2.7) and (2.8) by choosing the very same parameters as above and by now taking as cortical activation in V1 corresponding to $h$ the quantity $Lh(x,\theta) = \lambda L f_0(x,\theta) + L\mu(x,\theta)$.

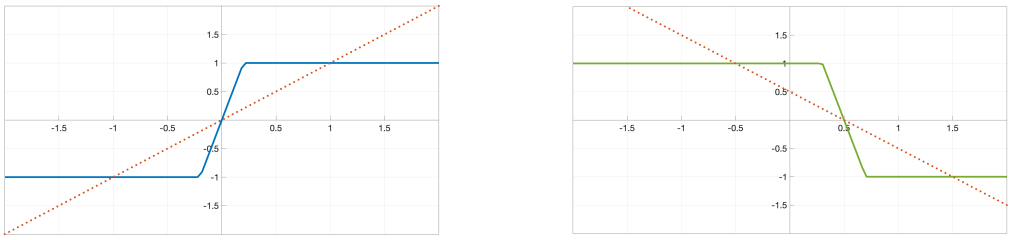
### 2.2.5 Numerical implementation

All four relevant equations (WC-2D), (LHE-2D), (WC-3D), and (LHE-3D) are numerically implemented via a forward Euler time-discretisation, as presented in [Bertalmío et al., 2007]. For a given image $a$, the cortical activation $La$ is recovered via standard wavelet transform methods, as presented in [Bertalmío et al., 2019] (see also [Duits and Franken, 2010]). The codes, written in Julia [Bezanson et al., 2017], are available at the following link: `http://www.github.com/dprn/WCvsLHE`.

All the considered images are $200 \times 200$ pixel and take values in the interval $[.15, .85]$, in order to avoid aliasing issues. We always consider $K = 30$ discretised orientations. As presented in Appendix B, the associated receptive profiles are obtained via cake wavelets [Bekkers et al., 2014], for which the frequency band `bw` is set to `bw`$= 5$. The interaction kernel is taken to be a 2D or 3D Gaussian with standard deviation $\boldsymbol{\sigma}_\omega$, the local mean average $\mu$ is obtained via Gaussian filtering with standard deviation $\boldsymbol{\sigma}_\mu$. In our experiments we used the following two piece-wise linear functions as sigmoids:

$$\hat{\sigma}(\rho) := \min\{1, \max\{\alpha\rho, -1\}\}, \qquad \sigma(\rho) := -\hat{\sigma}\left(x - \frac{1}{2}\right), \tag{2.10}$$

with $\alpha = 5$, see Figure 3. Note that $\hat{\sigma}$, which will be used for LHE models, is odd and centered in zero while $\sigma$, which will be used for WC models, is shifted in $1/2$ and shows a reversed behaviour. This in fact corresponds to a change of sign in the integral terms of LHE models w.r.t. the WC ones, as discussed in Section 2.2.2.



(a) $\hat{\sigma}$ and the line $y = x$ (b) $\sigma$ and the line $y = -x + \frac{1}{2}$

Figure 3: Sigmoid functions in the form (2.10), with $\alpha = 5$, as considered in our experiments.

Finally, the evolution equations stop when the $L^2$ relative distance between successive iterations is smaller than a tolerance $\tau = 10^{-2}$.

# 3 Results

In this section, we present the results obtained by applying the four models described above to the visual illusions described in Section 2.1. Our objective is to understand the capability of these models to *replicate* the visual illusions under consideration. That is, we are interested in whether the output produced by the models agrees with the human perception of the phenomena.

Assessing the replication of visual illusions by means of quantitative metrics is a very delicate matter. Due to the lack of a universal metric adapted to this task we will evaluate replication or lack thereof by presenting relevant line profiles, i.e., plots of brightness levels along a single row (line), of images produced by the four models in consideration. These lines are chosen as to cross a section of the image called *target*: A gray region in the image (or set of regions in the case of the Chevreul illusion), where the brightness illusion appears.

In all the results shown in this section, the original visual stimulus profile is represented as a blue dashed line. The line profiles of the output models are represented as solid red (LHE-2D), green (WC-2D), magenta (LHE-3D), and cyan (WC-3D) lines.

The parameters appearing in the models have been chosen independently for each illusion and each model, in order to obtain the best possible replication of the visual illusion. These are presented in Table 1.

## 3.1 Orientation-independent brightness illusions

| | **WC-2D** | | | | **LHE-2D** | | | | **WC-3D** | | | | **LHE-3D** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Illusion** | $\boldsymbol{\sigma_\mu}$ | $\boldsymbol{\sigma_\omega}$ | $\lambda$ | $M$ | $\boldsymbol{\sigma_\mu}$ | $\boldsymbol{\sigma_\omega}$ | $\lambda$ | $M$ | $\boldsymbol{\sigma_\mu}$ | $\boldsymbol{\sigma_\omega}$ | $\lambda$ | $M$ | $\boldsymbol{\sigma_\mu}$ | $\boldsymbol{\sigma_\omega}$ | $\lambda$ | $M$ |
| White | 10 | 20 | .7 | 1.4 | 10 | 50 | .7 | 1 | 20 | 30 | .7 | 1.4 | 2 | 50 | .7 | 1 |
| Brightness | 2 | 10 | .7 | 1.4 | 2 | 10 | .7 | 1 | 2 | 10 | .7 | 1.4 | 2 | 10 | .7 | 1 |
| Checkerboard | 10 | 70 | .7 | 1.4 | 10 | 70 | .7 | 1 | 10 | 70 | .7 | 1.4 | 10 | 70 | .7 | 1 |
| Chevreul | 2 | 5 | .7 | 1 | 2 | 10 | .7 | 1 | 2 | 40 | .5 | 1 | 5 | 7 | .7 | 1 |
| Chevreul canc. | 2 | 2 | .9 | 1 | 5 | 3 | .9 | 1 | 2 | 20 | .5 | 1.4 | 5 | 3 | .9 | 1 |
| Dungeon | 6 | 10 | .7 | 1.4 | 5 | 40 | .7 | 1 | 2 | 50 | .7 | 1.4 | 5 | 50 | .7 | 1 |
| Gratings | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 |
| Hong-Shevell | 5 | 20 | .7 | 1 | 5 | .5 | .7 | 1 | 10 | 30 | .7 | 1 | 10 | 30 | .7 | 1 |
| Luminance | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 | 2 | 6 | .7 | 1 |
| Poggendorff | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 3 | 10 | .5 | 1 |
| Tilt | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 15 | 20 | .7 | 1 |

Table 1: Parameters used in the tests.

Table 1 summarises the replication results obtained for the illusions described in Section 2.1: if the model replicates the illusion we indicate in the table the used parameters, otherwise a cross (✗) denotes no replication, i.e. the failure of the model to reproduce computational results corresponding to the visual perception of the considered illusion.

**White's illusion.** The chosen line profile for the plots in Fig. 4 corresponds to the central horizontal line of the image, which crosses both gray patches. As both plots show, all four models correctly predict the left target to be darker than the right one.

**Simultaneous brightness contrast.** The plots in Fig. 5 show the line profiles of the central horizontal line of the image, which crosses the two gray squares. We see that our four models replicate this illusion (left square lighter than the right square). In both the 2D and the 3D case, we observe that LHE methods result in an enhanced contrast effect than WC methods.

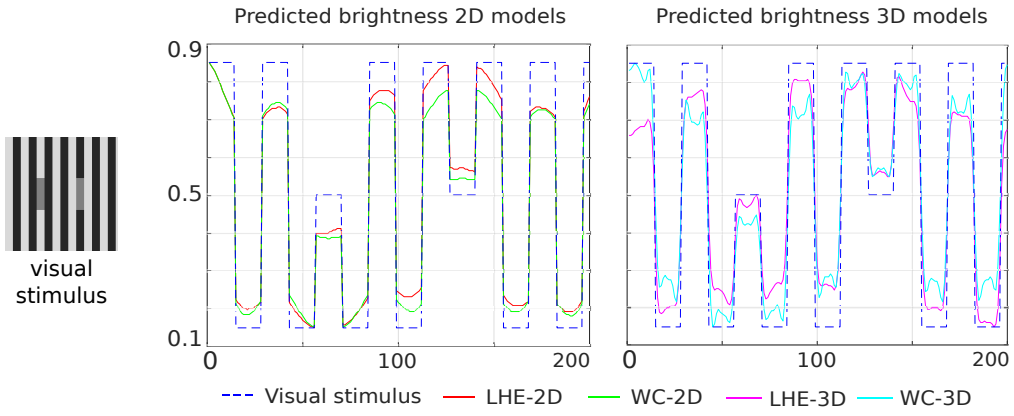



Figure 4: Predicted brightness in White illusion

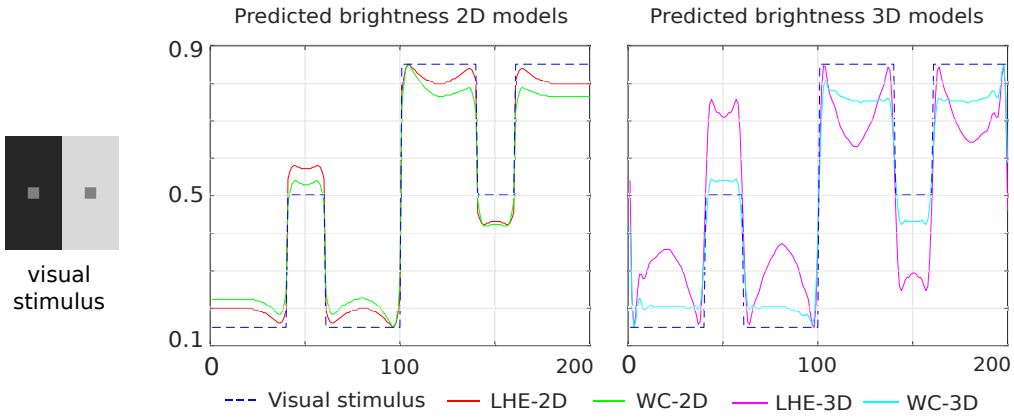



Figure 5: Predicted brightness in simultaneous brightness contrast

**Checkerboard illusion.** The chosen line profiles for this illusion are the two horizontal lines crossing, respectively, the left gray target and the right one. In Fig. 6, we chose to plot the first half of the line profile corresponding to the left target and the second half of the one corresponding to the right target. The profiles of all the four models show replication of this illusion, by which the left target is perceived darker than the right one.

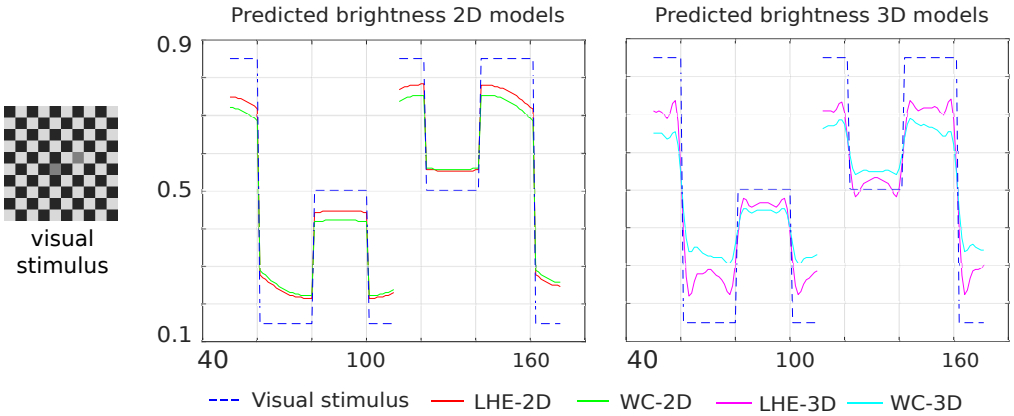



Figure 6: Predicted brightness in Checkerboard illusion

**Chevreul illusion.** Fig. 7 presents the line profiles for the central horizontal line. All four models correctly replicate the perceived changes within each band.

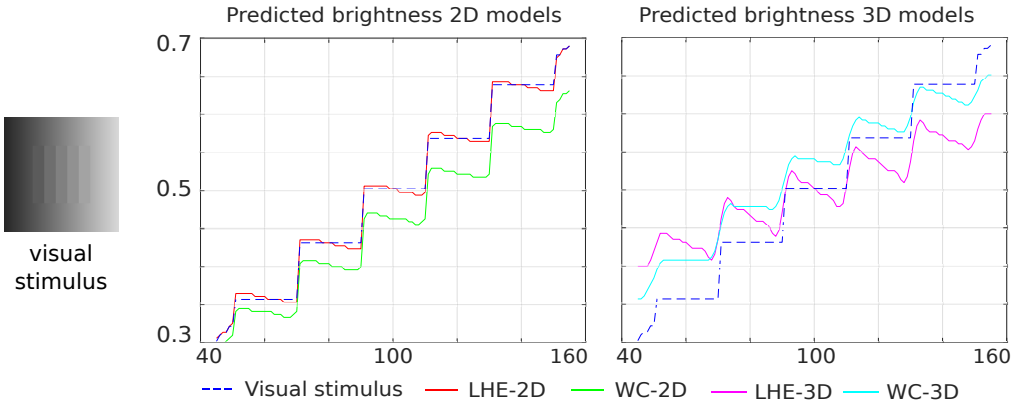



Figure 7: Predicted brightness in Chevreul illusion

.

**Chevreul cancellation.** The line profiles for the central horizontal line are presented in Fig. 8. In this case all models are able to correctly replicate the effect, although in the case of (WC-2D) and (LHE-3D) the perceptual response is not perfect, due to the presence of some oscillations. We also remark that the correct replication of this illusion is extremely sensitive to the chosen parameters.

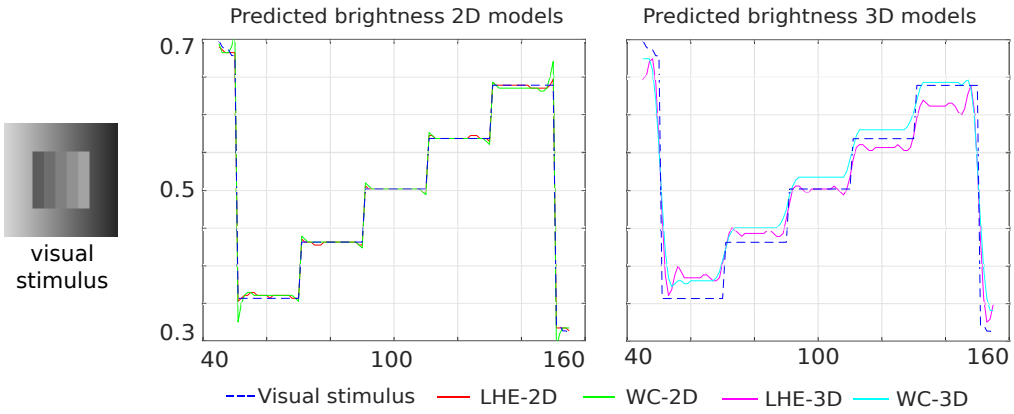



Figure 8: Predicted brightness in Chevreul cancellation

**Dungeon illusion.** Profiles of the central section (3 middle squares) of each target are shown in Fig. 9. The first part of the plot (left to right) represents the profile of the rectangle on black background. The second plot shows the target on white background. As these profiles show, our four proposed models replicate human perception (first target is predicted as darker than the second). Nevertheless, the assimilation effect (target intensity goes towards surrounding) is stronger in the 3D models.

**Grating induction.** In Fig. 10 the continuous and dashed blue lines respectively show the profile of the grating and of the central horizontal line row of the visual stimulus. Then, the line profiles of the central horizontal line of the outputs have been plotted. We observe that for both 2D and 3D models a counter-phase grating appears in the middle row, which successfully coincides with human perception. Notice that LHE methods have a higher amplitude in both cases.

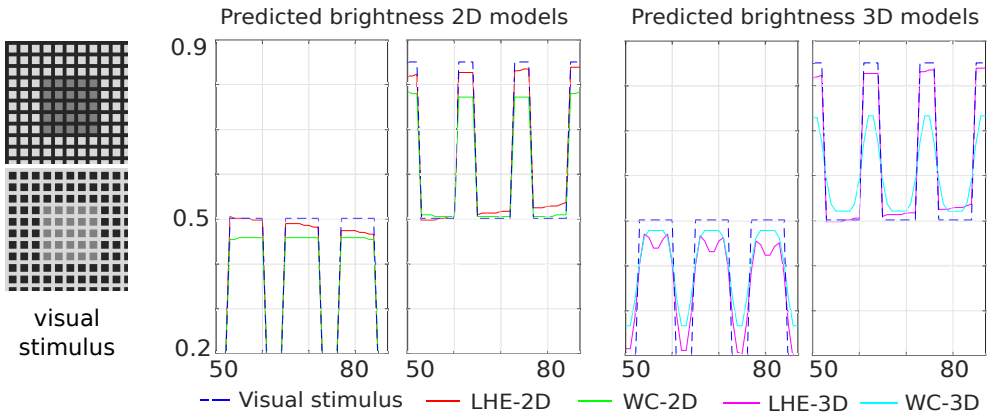



Figure 9: Predicted brightness in Dungeon illusion

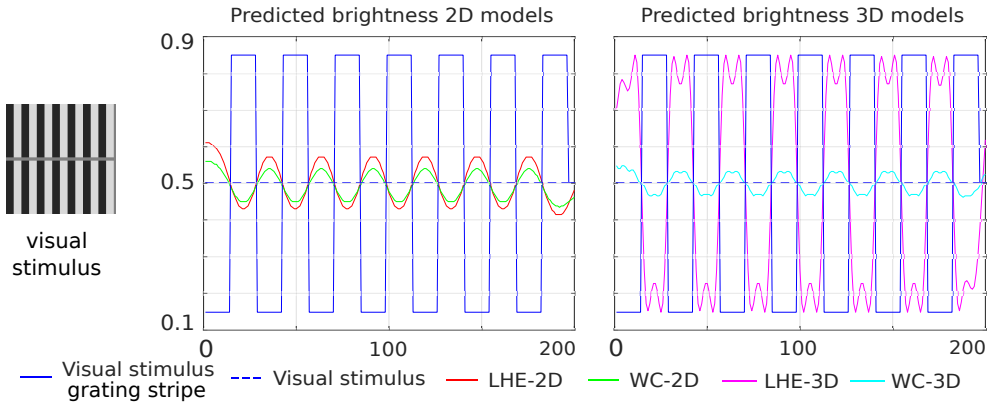



Figure 10: Predicted brightness in grating induction

**Hong-Shevell illusion.** Fig. 11 shows the line profiles of the central horizontal line around the target (gray ring) neighbourhood rings in the first half of the image. As in the case of the Dungeon illusion, we present in the first half of the plot (left to right) the output of the first stimulus (light background) and in the second half the output of the second (dark background). We see how our four proposed models replicate the assimilation effect. Hence, the gray ring in the first image is predicted as brighter than the gray ring in the second visual stimulus.

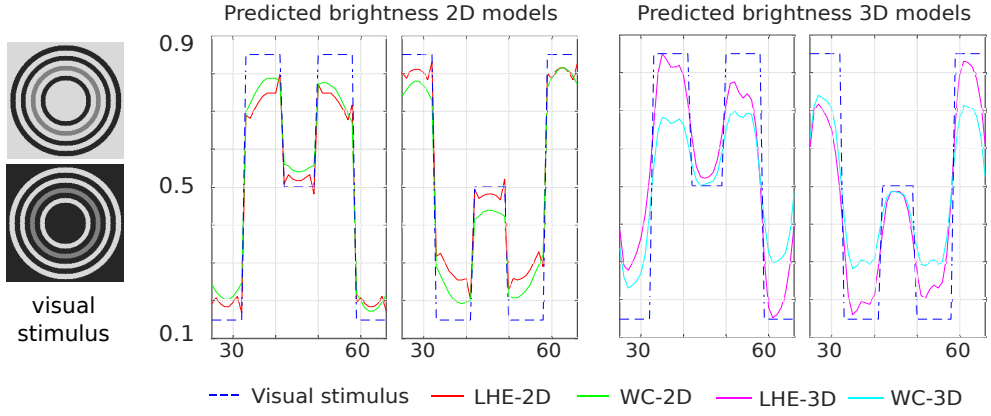



Figure 11: Predicted brightness in Hong-Shevell illusion

**Luminance illusion.** Horizontal profiles crossing top left and right targets (gray circles) are depicted in Fig. 12. For each target our four models reconstruct the left target as brighter than the right one. Hence, all models correctly predict this contrast effect. In this case, LHE presents

a higher contrast response in both responses (2D and 3D).

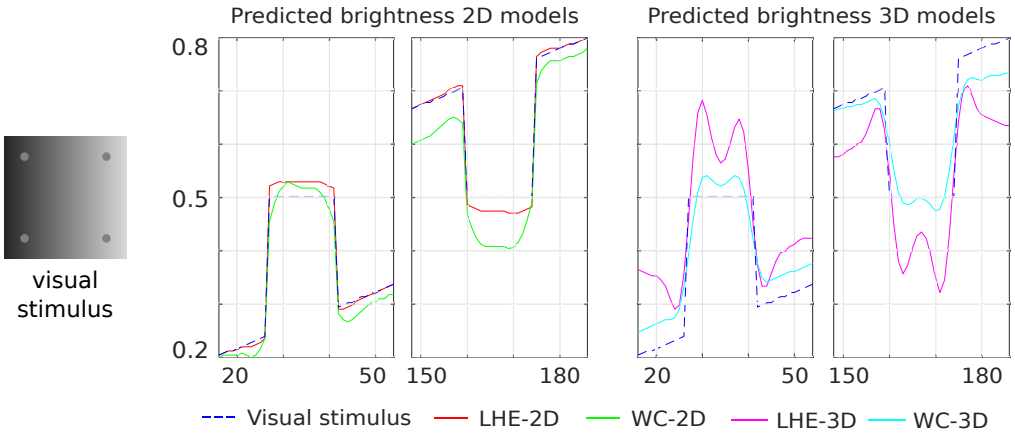



Figure 12: Predicted brightness in luminance gradient illusion

We observe that in all the considered brightness-dependent illusions the LHE-3D method presents neighbourhood-dependent oscillations which can or not happen in the WC-3D method.

## 3.2 Orientation-dependent illusions

**Poggendorff illusion.** Fig. 13 shows a profile of the middle row in the visual stimulus, while the output images and a zoom of the target gray middle area are presented in Fig. 14. In this case (WC-2D), (WC-3D), and (LHE-2D) are not able to completely replicate the illusion, since induced white lines on the gray area are not connected. On the other hand, (LHE-3D) successfully replicates the perceptual completion over the gray middle stripe.

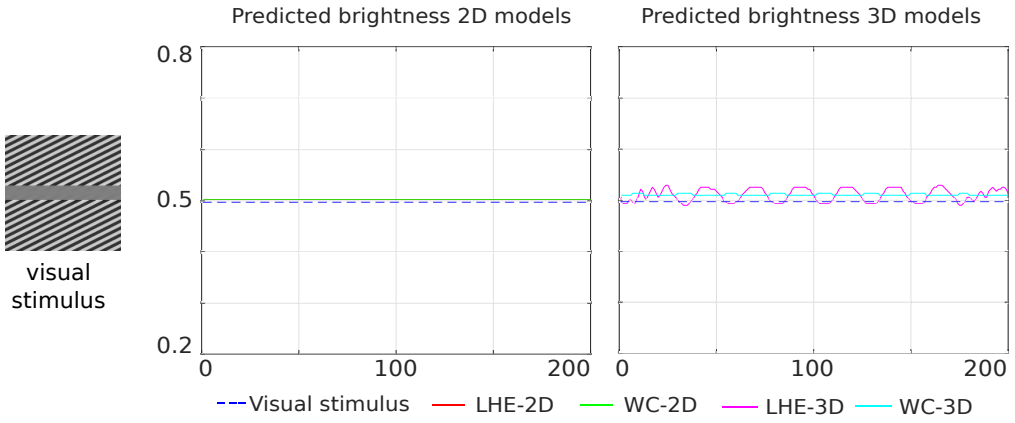



Figure 13: Predicted brightness in Poggendorff illusion

**Tilt illusion.** In Fig. 15 we present line profiles, for both visual stimuli, for a diagonal line starting at the bottom left corner of the image and ending at the top right one. In order to be able to correctly compare the two images, the line profile of the second image (from top to bottom) has been extracted after flipping the outer circle along the vertical axis, so that the responses to both stimulus have the same background. Although there is a noticeable effect, such as a reduction in contrast for the (WC-2D), the difference between the responses to the two stimuli is very mild for all models with the exception of (LHE-3D).

The fact that indeed this model is replicating the effect can be better appreciated looking at Fig. 16, which shows a composite of the inner circle for the responses to the two visual stimuli of the two orientation-dependent models. It is then evident that the (LHE-3D) model yields a stronger result than the (WC-3D) one. In fact, the former shows increased visibility (measured

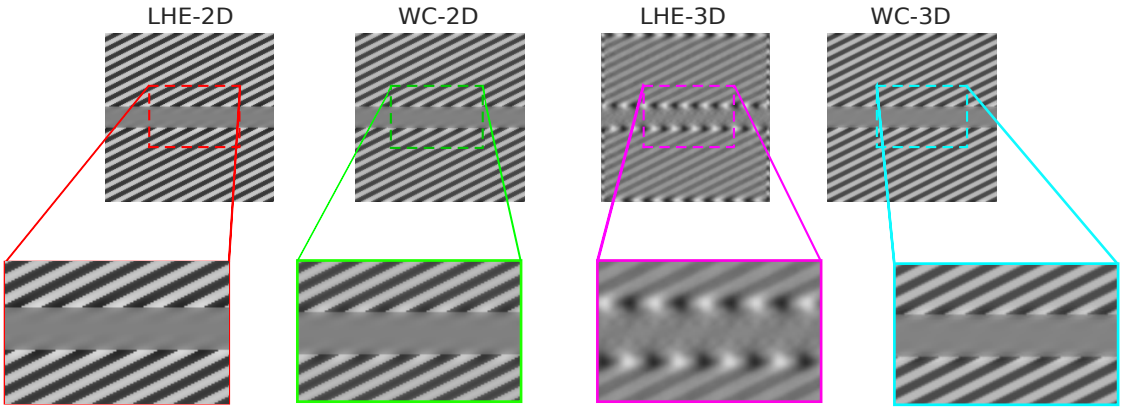



Figure 14: Zoom of the predicted completion for Poggendorff illusion

here as the contrast) for the half of the circle corresponding to the second stimulus than the one corresponding to the first stimulus. On the other hand, in the case of the (WC-3D) model (or of 2D models, not depicted here), the circle shows no difference among its two halves. This justifies our claim that the (LHE-3D) model can increase the visibility of the inner circle (replicate the illusion) based on the orientation of the outer circle.

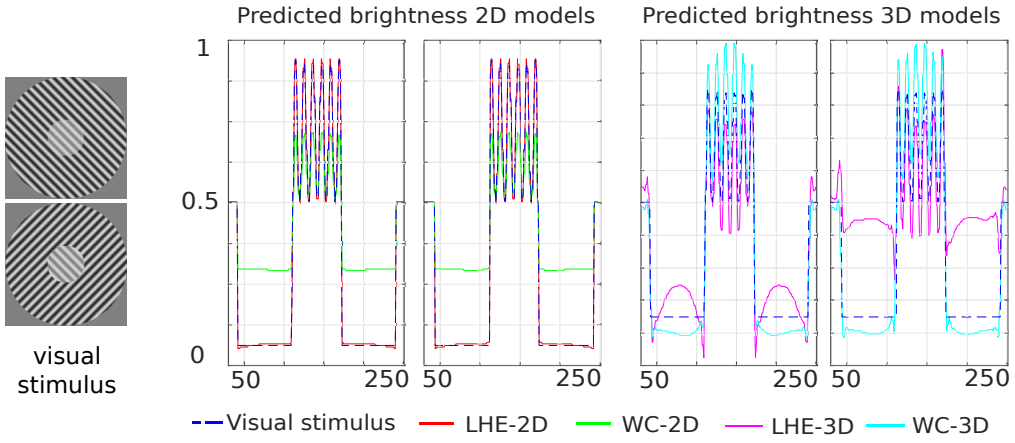



Figure 15: Predicted brightness in Tilt illusion

## 4 Discussion

The results presented in the previous section show that the four models are able to reproduce several brightness illusions. Concerning orientation-dependent illusions we observe that, as expected, 2D models cannot reproduce them, while the only 3D model that correctly reproduces the perceptual outcome is the (LHE-3D). However we stress that determining replication or lack thereof in the Tilt illusion is subtle, as the observed effects are very mild.

As already mentioned, the parameters of the presented results are chosen independently from one illusion to the other in order to qualitatively optimise the perceptual replication. Empirical observations show that these are indeed related with, e.g., the size of the target and the frequency of the background. Nevertheless, if one settles for milder replications, it would be possible to choose more uniform parameters. For instance, this happens for the (WC-3D) model in the Chevreul and Chevreul cancellation illusions, which can be reproduced simultaneously with parameters $\boldsymbol{\sigma}_\mu = 3$ and $\boldsymbol{\sigma}_\omega = 30$, although with less striking results.

Regarding the 3D models, we notice that they always give oscillatory solutions. This is especially true for the (LHE-3D) model. Such oscillations can be observed as dependent on the target surrounding.

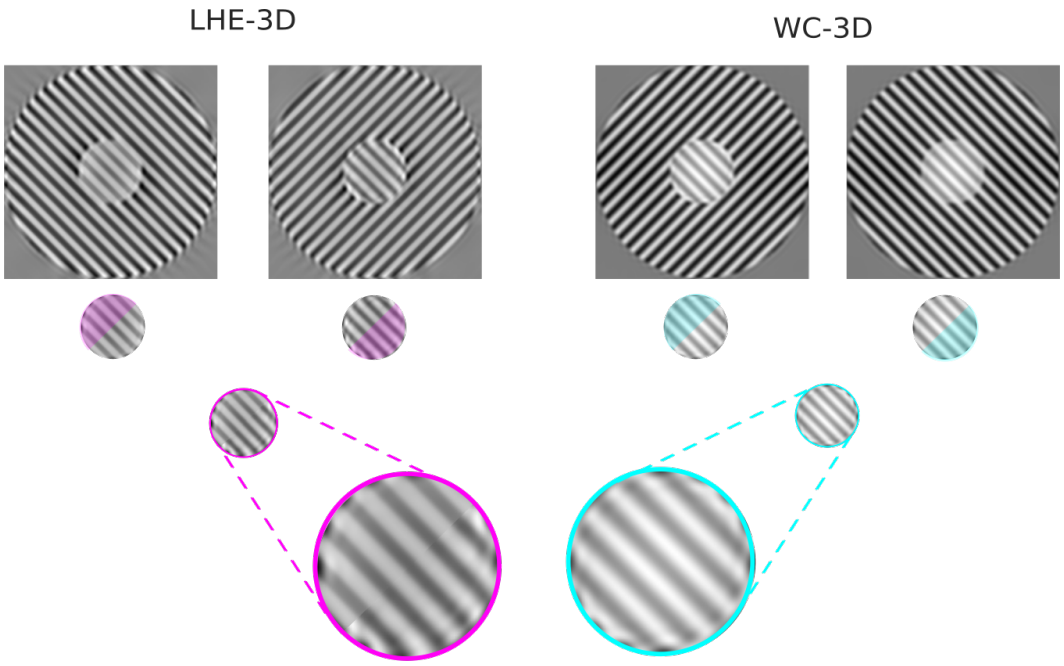



Figure 16: Detail in predicted brightness in Tilt illusion

## 5 Conclusions and future work

The analysis performed in this paper shows that, while Wilson-Cowan equations can be used to replicate brightness illusions, their variational modification, in the form of local histogram equalisation, outperforms them. Moreover, although in the case of brightness illusions we found no real advantage in considering model variants that take into account orientations, these are necessary to reproduce the two considered orientation-dependent illusions, which only the 3D variational model is able to reproduce.

These observations justify further investigation on the orientation-dependent LHE model. In particular, a more accurate modelling reflecting the actual structure of V1 should be addressed. This concerns first the lift operation, where the cake wavelet should be replaced by the more physiologically plausible Gabor filters, as well as the interaction weight $\omega$ which could be taken to be the anisotropic heat kernel of [Citti and Sarti, 2006, Sarti and Citti, 2015, Duits and Franken, 2010]. Finally, extensive numerical experiments should be performed to assess the compatibility of the model with psychophysical tests measuring the perceptual bias induced by these phenomena. This would provide insights about the robustness of the model in reproducing visual behaviour.

## Acknowledgements and Grants

M. B. would like to thank the organizers of the conference to celebrate Jack Cowan's 50 years at the University of Chicago for their kind invitation to attend that meeting, which served as inspiration for this work. L. C., V. F. and D. P. acknowledge the support of a public grant overseen by the French National Research Agency (ANR) as part of the *Investissement d'avenir program*, through the iCODE project funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02 and of the research project *LiftME* funded by INS2I, CNRS. L. C. and V. F. acknowledge the support provided by the *Fondation Mathématique Jacques Hadamard*. V. F. and D. P. also acknowledge the support of ANR-15-CE40-0018 project *SRGI - Sub-Riemannian Geometry and Interactions*. B. F. acknowledges the support of the Fondation Asile des Aveugles.

## Disclosures

All authors declare that there is no commercial relationship relevant to the subject matter of presentation.

## Author contributions

All authors equally contributed to this work.

## A Non-variational nature of Wilson-Cowan equation

In this section we show that, for non-trivial choices of weight and sigmoid functions, Wilson-Cowan equations do not admit a variational formulation.

For the sake of simplicity, we will consider only a finite dimensional variant of Wilson-Cowan equations, with constant input. Namely, for $a : \mathbb{R} \to \mathbb{R}^n$, we consider

$$\frac{d}{dt}a(t) = -\mu a(t) + W\sigma(a(t)) + h. \tag{A.1}$$

Here, $h \in \mathbb{R}^n$ is the input, $\mu > 0$ is a parameter, $\sigma \in C^1(\mathbb{R})$ is any function (we denoted $\sigma(v) = (\sigma(v_i))_i$ for $v \in \mathbb{R}^n$), and $W \in \mathbb{R}^{n\times n}$ is a symmetric interaction kernel.

Equation (A.1) admits a variational formulation if it can be written as the steepest descent associated with a functional $J : \mathbb{R}^n \to \mathbb{R}$, i.e.,

$$\frac{d}{dt}a(t) = -\nabla J(a(t)). \tag{A.2}$$

We have the following.

**Theorem.** *The Wilson-Cowan equation* (A.1) *admits the variational formulation* (A.2) *only if either $W$ is a diagonal matrix, or $\sigma$ is an affine function, i.e., $\sigma(x) = \alpha x + \beta$ for some $\alpha, \beta \in \mathbb{R}$.*

*Proof.* Writing (A.1) and (A.2) componentwise, we find the following relation for $J$:

$$\partial_i J(v) = -\mu v_i + \sum_k W_{\ell,k}\sigma(v_\ell) + h_i, \qquad v = (v_1, \dots, v_n) \in \mathbb{R}^n, \quad i = 1, \dots, n.$$

By differentiating again the above, we have

$$\partial_{ij} J(v) = -\mu\delta_{ij} + \sum_k W_{\ell,k}\sigma'(v_\ell)\delta_{j\ell} = -\mu\delta_{ij} + W_{ij}\sigma'(v_j), \qquad i, j = 1, \dots, n. \tag{A.3}$$

Namely, $\operatorname{Hess} J(v) = (-\mu\delta_{ij} + W_{ij}\sigma'(v_j))_{ij}$. Assume that $W$ is not a diagonal matrix. Then, since both the Hessian matrix and $W$ are symmetric, by choosing $i \neq j$ such that $W_{ij} \neq 0$ we get

$$\sigma'(v_i) = \sigma'(v_j) \qquad \forall v \in \mathbb{R}^n. \tag{A.4}$$

This clearly implies that $\sigma'(x)$ is constant, thus showing that $\sigma$ must be an affine function. □

We observe that the above reasoning does not apply to the LHE algorithm. Indeed, the discrete form of the latter is

$$\frac{d}{dt}a(t) = -\mu a(t) + \sum_\ell W_{i\ell}\sigma\big(a_i(t) - a_\ell(t)\big) + h. \tag{LHE}$$

Then, the corresponding variational equation (for $\mu = 0$ and $h = 0$) is

$$\partial_i J(v) = \sum_{\ell\neq i} W_{i\ell}\sigma(v_i - v_\ell), \qquad v \in \mathbb{R}^n. \tag{A.5}$$

This yields

$$\partial_{ji} J(v) = -W_{ij}\sigma'(v_i - v_j), \qquad \text{for } v \in \mathbb{R}^n, \quad i \neq j. \tag{A.6}$$

This does not contradict the symmetry of the Hessian, as $\sigma$ was chosen to be odd an thus $\sigma'$ is even. Indeed, we know by [Bertalmío and Cowan, 2009] that we can let

$$J(v) := \sum_{k,\ell} W_{k\ell}\Sigma(v_k - v_\ell), \tag{A.7}$$

where $\Sigma$ is such that $\Sigma' = \sigma$.

# B Encoding orientation-dependence via cortical-inspired models

Orientation dependence of the visual stimulus is encoded via cortical inspired techniques, following e.g., [Citti and Sarti, 2006, Duits and Franken, 2010, Petitot, 2017, Prandi and Gauthier, 2017, Bohi et al., 2017]. The main idea at the base of these works goes back to the 1959 paper [Hubel and Wiesel, 1959] by Hubel and Wiesel (Nobel prize in 1981) who discovered the so-called *hypercolumn functional architecture* of the visual cortex V1. Following [Hubel and Wiesel, 1959], each neuron $\xi$ in V1 detects couples $(x, \theta)$ where $x \in \mathbb{R}^2$ is a retinal position and $\theta$ is a direction at $x$. Orientation preferences $\theta$ are then organised in hypercolumns over the retinal position $x$, see [Petitot, 2017, Section 2].

Let $Q \subset \mathbb{R}^2$ be the visual plane. To a visual stimulus $f : Q \to [0, 1]$ is associated a cortical activation $Lf : Q \times [0, \pi) \to \mathbb{R}$ such that $Lf(\xi)$ encodes the response of the neuron $\xi = (x, \theta)$. Letting $\psi_\xi \in L^2(\mathbb{R}^2)$ be the receptive profile (RP) of the neuron $\xi$, such response is assumed to be given by

$$Lf(\xi) = \langle \psi_\xi, f \rangle_{L^2(\mathbb{R}^2)} = \int_Q \overline{\psi_\xi(x)} f(x)\, dx. \tag{B.1}$$

Motivated by neuro-phyisiological evidence, we assume that RPs of different neurons are "deducible" one from the other via a linear transformation. As detailed in [Duits and Franken, 2010, Prandi and Gauthier, 2017], see also [Bertalmío et al., 2019, Section 3.1], this amounts to the fact that the linear operator $L : L^2(Q) \to L^2(Q \times [0, \pi))$ is a continuous wavelet transform (also called *invertible orientation score transform*). That is, there exists a *mother wavelet* $\Psi \in L^2(\mathbb{R}^2)$ such that $Lf(x, \theta) = \big[f * (\Psi^* \circ R_{-\theta})\big](x)$. Here, $f * g$ denotes the standard convolution on $L^2(\mathbb{R}^2)$ and $R_{-\theta}$ is the counter-clock-wise rotation of angle $\theta$. Notice that, although images are functions of $L^2(\mathbb{R}^2)$ with values in $[0, 1]$, it is in general not true that $Lf(x, \theta) \in [0, 1]$.

Concerning the choice of the mother wavelet, we remark that neuro-physiological evidence suggests that a good fit for the RPs is given by Gabor filters, whose Fourier transform is the product of a Gaussian with an oriented plane wave [Daugman, 1985]. However, these filters are quite challenging to invert, and are parametrised on a bigger space than $\mathcal{M}$, which takes into account also the frequency of the plane wave and not only its orientation. For this reason, in this work we instead considered *cake wavelets*, introduced in [Duits et al., 2007, Bekkers et al., 2014]. These are obtained via a mother wavelet $\Psi^{\text{cake}}$ whose support in the Fourier domain is concentrated on a fixed slice, depending on the number of orientations one aims to consider in the numerical implementation. For the sake of integrability, the Fourier transform of this mother wavelet is then smoothly cut off via a low-pass filtering, see [Bekkers et al., 2014, Section 2.3] for details. Observe, however, that, since we are considering orientations on $[0, \pi)$ and not directions on $[0, 2\pi)$, we choose a non-oriented version of the mother wavelet, given by $\tilde{\psi}^{cake}(\omega) + \tilde{\psi}^{cake}(e^{i\pi}\omega)$, in the notations of [Bekkers et al., 2014].

An important feature of cake wavelets is that, in order to recover the original stimulus from its cortical activation, it suffices to simply "project" the cortical activations along hypercolumns. This yields

$$f(x) := \frac{1}{\pi} \int_0^\pi Lf(x, \theta)\, d\theta. \tag{B.2}$$

This justify the assumption, implicit in equation (2.9), that the projection of a cortical activation $F$ (not necessarily given by a visual stimulus) to the visual plane is given by

$$PF(x) = \frac{1}{\pi} \int_0^\pi F(x, \theta)\, d\theta. \tag{B.3}$$

# References


[Adini et al., 1997] Adini, Y., Sagi, D., and Tsodyks, M. (1997). Excitatory–inhibitory network in the visual cortex: Psychophysical evidence. *Proceedings of the National Academy of Sciences*, 94(19):10426–10431.

[Atick, 1992] Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems*, 3(2):213–251.

[Attneave, 1954] Attneave, F. (1954). Some informational aspects of visual perception. *Psychological review*, 61(3):183.

[Barlow et al., 1961] Barlow, H. B. et al. (1961). Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1:217–234.

[Bekkers et al., 2014] Bekkers, E., Duits, R., Berendschot, T., and ter Haar Romeny, B. (2014). A multi-orientation analysis approach to retinal vessel tracking. *JMIV*, 49(3):583–610.

[Benucci et al., 2013] Benucci, A., Saleem, A. B., and Carandini, M. (2013). Adaptation maintains population homeostasis in primary visual cortex. *Nature neuroscience*, 16(6):724.

[Bertalmío, 2014] Bertalmío, M. (2014). From image processing to computational neuroscience: a neural model based on histogram equalization. *Front. Comput. Neurosc.*, 8:71.

[Bertalmío et al., 2019] Bertalmío, M., Calatroni, L., Franceschi, V., Franceschiello, B., and Prandi, D. (2019). A cortical-inspired model for orientation-dependent contrast perception: A link with wilson-cowan equations. In Lellmann, J., Burger, M., and Modersitzki, J., editors, *Scale Space and Variational Methods in Computer Vision*, pages 472–484, Cham. Springer International Publishing.

[Bertalmío et al., 2007] Bertalmío, M., Caselles, V., Provenzi, E., and Rizzi, A. (2007). Perceptual color correction through variational techniques. *IEEE T. Image Process.*, 16(4):1058–1072.

[Bertalmío and Cowan, 2009] Bertalmío, M. and Cowan, J. D. (2009). Implementing the retinex algorithm with wilson–cowan equations. *Journal of Physiology-Paris*, 103(1-2):69–72.

[Bertalmío and Cowan, 2009] Bertalmío, M. and Cowan, J. D. (2009). Implementing the retinex algorithm with Wilson-Cowan equations. *J. Physiol. Paris*, 103(1):69 – 72.

[Bertalmío et al., 2017] Bertalmío, M., Cyriac, P., Batard, T., Martinez-Garcia, M., and Malo, J. (2017). The wilson-cowan model describes contrast response and subjective distortion. *J Vision*, 17(10):657–657.

[Beurle and Matthews, 1956] Beurle, R. L. and Matthews, B. H. C. (1956). Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 240(669):55–94.

[Bezanson et al., 2017] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.

[Bohi et al., 2017] Bohi, A., Prandi, D., Guis, V., Bouchara, F., and Gauthier, J.-P. (2017). Fourier descriptors based on the structure of the human primary visual cortex with applications to object recognition. *Journal of Mathematical Imaging and Vision*, 57(1):117–133.

[Bressan, 2001] Bressan, P. (2001). Explaining lightness illusions. *Perception*, 30(9):1031–1046.

[Brucke, 1865] Brucke, E. (1865). uber erganzungs und contrasfarben. *Wiener Sitzungsber*, 51.

[Citti and Sarti, 2006] Citti, G. and Sarti, A. (2006). A cortical based model of perceptual completion in the roto-translation space. *JMIV*, 24(3):307–326.

[Cowan et al., 2016] Cowan, J. D., Neuman, J., and van Drongelen, W. (2016). Wilson–cowan equations for neocortical dynamics. *The Journal of Mathematical Neuroscience*, 6(1):1.

[Daugman, 1985] Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169.

[DeValois and DeValois, 1990] DeValois, R. L. and DeValois, K. K. (1990). *Spatial vision*. Oxford university press.

[Duits et al., 2007] Duits, R., Felsberg, M., Granlund, G., and ter Haar Romeny, B. (2007). Image analysis and reconstruction using a wavelet transform constructed from a reducible representation of the euclidean motion group. *International Journal of Computer Vision*, 72(1):79–102.

[Duits and Franken, 2010] Duits, R. and Franken, E. (2010). Left-invariant parabolic evolutions on $SE(2)$ and contour enhancement via invertible orientation scores. Part I: linear left-invariant diffusion equations on $SE(2)$. *Quart. Appl. Math.*, 68(2):255–292.

[Eagleman, 1959] Eagleman, D. M. (1959). Visual illusions and neurobiology. *Nature reviews Neuroscience*, (2):920926 (2001).

[Ernst et al., 2016] Ernst, U. A., Schiffer, A., Persike, M., and Meinhardt, G. (2016). Contextual interactions in grating plaid configurations are explained by natural image statistics and neural modeling. *Frontiers in systems neuroscience*, 10:78.

[Fairhall et al., 2001] Fairhall, A. L., Lewen, G. D., Bialek, W., and van Steveninck, R. R. d. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787.

[Faugeras, 2009] Faugeras, O. (2009). A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience*, 3.

[Geier and Hudák, 2011] Geier, J. and Hudák, M. (2011). Changing the chevreul illusion by a background luminance ramp: lateral inhibition fails at its traditional stronghold-a psychophysical refutation. *PloS One*, 6(10):e26062.

[Herzog et al., 2003] Herzog, M. H., Ernst, U. A., Etzold, A., and Eurich, C. W. (2003). Local interactions in neural networks explain global effects in gestalt processing and masking. *Neural Computation*, 15(9):2091–2113.

[Hong and Shevell, 2004] Hong, S. W. and Shevell, S. K. (2004). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research*, 44(1):35–43.

[Hubel and Wiesel, 1959] Hubel, D. and Wiesel, T. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3):574591.

[Kingdom, 2011] Kingdom, F. A. (2011). Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy. *Vision Research*, 51(7):652–673.

[Kitaoka, 2006] Kitaoka, A. (2006). Adelsons checker-shadow illusion-like gradation lightness illusion. `http://www.psy.ritsumei.ac.jp/~akitaoka/gilchrist2006mytalke.html`. Accessed: 2018-11-03.

[Mante et al., 2005] Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature neuroscience*, 8(12):1690.

[McCourt, 1982] McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research*, 22(1):119–134.

[Meister and Berry, 1999] Meister, M. and Berry, M. J. (1999). The neural code of the retina. *Neuron*, 22(3):435–450.

[Murray and Herrmann., 2013] Murray, M. M. and Herrmann., C. (2013). Illusory contours: a window onto the neurophysiology of constructing perception. *Trends Cogn Sci.*, (17(9)):471–81.

[Olshausen and Field, 2000] Olshausen, B. A. and Field, D. J. (2000). Vision and the coding of natural images: The human brain may hold the secrets to the best image-compression algorithms. *AmSci*, 88(3):238–245.

[Petitot, 2017] Petitot, J. (2017). *Elements of Neurogeometry: Functional Architectures of Vision*. Lecture Notes in Morphogenesis. Springer International Publishing.

[Prandi and Gauthier, 2017] Prandi, D. and Gauthier, J.-P. (2017). *A semidiscrete version of the Petitot model as a plausible model for anthropomorphic image reconstruction and pattern recognition*. SpringerBriefs in Mathematics. Springer International Publishing, Cham.

[Purves et al., 2008] Purves, D., Wojtach, W. T., and Howe, C. (2008). Visual illusions: An Empirical Explanation. *Scholarpedia*, 3(6):3706. revision #89112.

[Ratliff, 1965] Ratliff, F. (1965). *Mach bands: quantitative studies on neural networks*. Holden-Day, San Francisco London Amsterdam.

[Sarti and Citti, 2015] Sarti, A. and Citti, G. (2015). The constitution of visual perceptual units in the functional architecture of V1. *J. comput. neurosc.*, 38(2):285–300.

[Shapiro and Todorovic, 2016] Shapiro, A. G. and Todorovic, D. (2016). *The Oxford compendium of visual illusions*. Oxford University Press.

[Smirnakis et al., 1997] Smirnakis, S. M., Berry, M. J., Warland, D. K., Bialek, W., and Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386(6620):69.

[Weintraub and Krantz, 1971] Weintraub, D. J. and Krantz, D. H. (1971). The Poggendorff illusion: amputations, rotations, and other perturbations. *Attent. Percept. Psycho.*, 10(4):257–264.

[White, 1979] White, M. (1979). A new effect of pattern on perceived lightness. *Perception*, 8(4):413–416.

[Wilson and Cowan, 1972] Wilson, H. R. and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *BioPhys. J.*, 12(1).

[Wilson and Cowan, 1973] Wilson, H. R. and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, 13(2):55–80.