This is the author's accepted manuscript. The final published version of this work “Reaction times in visual search can be explained by a simple model of neural synchronization”. is published by Elsevier in

Neural Networks Volume 87, March 2017, Pages 1–7

Available at: 10.1016/j.neunet.2016.12.003 (http://dx.doi.org/10.1016/j.neunet.2016.12.003)

Article history:
Received 4 August 2016
Received in revised form 1 December 2016
Accepted 2 December 2016
Available online 10 December 2016

Embargo Period – 12 month

Title page

Reaction times in visual search can be explained by a simple model of neural synchronization

Yakov Kazanovich[a] and Roman Borisyuk[a,b,*]

[a]Institute of Mathematical Problems of Biology – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, 142290, Russia

[b]School of Computing and Mathematics, Plymouth University, Plymouth, PL4 8AA, United Kingdom

[*] Corresponding author
E-mail addresses: yasha@impb.psn.ru (Yakov Kazanovich), r.borisyuk@plymouth.ac.uk (Roman Borisyuk)

Reaction times in visual search can be explained by a simple model of neural synchronization

Yakov Kazanovich[a] and Roman Borisyuk[a,b,*]

[a]Institute of Mathematical Problems of Biology – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, 142290, Russia

[b]School of Computing and Mathematics, Plymouth University, Plymouth, PL4 8AA, United Kingdom

Abstract

We present an oscillatory neural network model that can account for reaction times in visual search experiments. The model consists of a central oscillator that represents the central executive of the attention system and a number of peripheral oscillators that represent objects in the display. The oscillators are described as generalized Kuramoto type oscillators with adapted parameters. An object is considered as being included in the focus of attention if the oscillator associated with this object is in-phase with the central oscillator. The probability for an object to be included in the focus of attention is determined by its saliency that is described in formal terms as the strength of the connection from the peripheral oscillator to the central oscillator. By computer simulations it is shown that the model can reproduce reaction times in visual search tasks of various complexities. The dependence of the reaction time on the number of items in the display is represented by linear functions of different steepness which is in agreement with biological evidence.

Keywords: visual search, reaction times, oscillatory neural network, synchronization

1. Introduction

Visual search is a type of perceptual task that involves an active scan of the visual environment for a particular object (the target) surrounded by other objects (the distractors). The task of visual search can be of various complexities depending on the saliency of the target relative to the distractors. This is reflected in the duration of time that the observer spends performing the search task, and in the number of errors that are made. A vast amount of experimental evidence has been obtained to characterize the mechanisms of visual selection both at the level of information processing by neural structures and of the algorithms that are consciously or subconsciously used by subjects in visual search experiments.

In the early experiments of Treizman and Gelade (1980) it was discovered that visual search tasks can be subdivided into several categories on the basis of their difficulty. Though later experiments have shown that there are no strict barriers between the categories (Nothdurft, 1999), the following three categories are considered as basic experimental paradigms (Wolfe, Palmer, & Horowitz, 2010):

* Corresponding author
E-mail addresses: yasha@impb.psn.ru (Yakov Kazanovich), r.borisyuk@plymouth.ac.uk (Roman Borisyuk)

Feature Search is the case where the target and the distractors are maximally different, i.e. differentiated by a single property such as colour, shape, orientation, or size. In this case the reaction time (RT) is found to be short and independent of the number of objects *n* in the display (i.e. the graph RT(*n*) is nearly parallel to the *n*-axis).

Conjunction Search occurs when the target shares a property with the distractors, e.g. a red vertical bar should be found among vertical green bars and horizontal red bars. In this case the search task is more difficult and RT increases when *n* increases.

Spatial Configuration Search is even more difficult. An example of such task is the search of the number 2 among numbers 5 (2 is a reflection of 5 relative to the vertical mirror). In this case the graph RT(*n*) is more steep than in the case of Conjunction Search.

Several theoretical approaches have been put forward to explain the experimental results. Most influential are the following: *Feature Integration Theory* (Treisman & Gelade, 1980; Treisman & Sato, 1990), *Attentional Engagement Theory* (Duncan & Humphreys, 1989; Duncan & Humphreys, 1992), *Signal Detection Theory* (Eckstein, 1998; Eckstein, Thomas, Palmer, & Shimozaki., 2000; Palmer, Verghese, & Pavel, 2000; Verghese, 2001; Cameron, Tai, Eckstein, & Carrasco, 2004), *Guided Search* (Wolfe, Cave, & Franzel, 1989; Wolfe, 1994; Wolfe, 2007; Wolfe, 2010) (see also the review of Chan and Hayward (2013) and the monograph of Sternberg and Sternberg (2012)).

Over the past 30 years numerous theoretical and computational models of the visual search have been developed. Most of them are constructed along a common scheme that describes the focus of attention (FOA) formation and switching mechanisms. First, a saliency map is computed. This map is a representation of visual space in which the level of activation at a location reflects the likelihood that the location contains an object target (Koch & Ullman, 1985; Itti & Koch, 2000). Usually it is assumed that the saliency map includes both bottom-up and top-down influences. The bottom-up bias in activation results from the differences in visual features between the target and surrounding distractors. The top-down modulation of the saliency map is conditioned by the subject's knowledge of the features and their combinations that characterize targets and distractors. Often the inhibition of return is added to the models to suppress the activation in the locations previously inspected in the search for a target. Second, the saliency map is used to bias the competition between objects in the visual field for their inclusion in the FOA: an object of greater salience has higher probability of being selected in the FOA. Third, an object in the FOA is recognized as a target or as a distractor and after that a decision is taken about terminating or continuing the search.

The models differ in the details of implementation of the stages of visual search. They can be divided into two categories. The models of the first type reproduce the experimental data on the

number of errors in visual search tasks (Itti & Koch, 2000; Itti, 2005; Wilder, Mozer, & Wickens, 2011; Walles, Robins, & Knott, 2013). The performance of these models is usually evaluated and compared with experimental evidence in terms of the percentage of correct decisions or the parameters of a psychometric function that characterize the probability of correct decisions under stimulus manipulations (Palmer, Verghese, & Pavel, 2000). Also an information theoretic approach is used by some researchers (Bruce & Tsotsos, 2009; Vincent, Baddeley, Troscianko, & Gilchrist, 2009; Ma, Navalpakkam, Beck, van den Berg, & Pouget, 2011), where the selection of an object is based on the maximum-likelihood estimate of the stimulus at each given location and the performance of the model is evaluated in terms of receiver operating characteristic (ROC) curves.

The models of the second type reproduce experimental data on RT (Heinke, Humphreys, & Tweed, 2006; Herd & O'Reilly, 2006). They try to more closely follow anatomical and neurophysiological facts about the interaction of several brain areas in the bottom-up and top-down processes during FOA formation and switching. For example, a recent paper of Schwemmer et al. (2015) presents an advanced model for reproducing the distribution of RT. It is described by 20 differential equations representing the activity of four areas of the monkey brain: anterior and lateral intraparietal areas, inferior temporal cortex, and motor cortex. The behaviour of the model is controlled by 31 parameters, of which 13 parameters must be fit individually for each monkey. The model was tested on experimental data in a difficult Spatial Configuration Search with 2, 4, and 6 objects in the display. Unfortunately, this model is very complicated, adapted to a special type of experiment and has many free parameters that should be determined by training.

Recently Guided Search Theory received support from a model that is able to reproduce not only average RTs in visual search tasks but also time distributions (Moran, Zehetleitner, Müller, & Usher, 2013). The model does not suggest a neural mechanism for its implementation. It only assumes that the selection process is based on Luce's choice axiom (Luce, 1959) according to which each object in the image is selected with the probability $p_i = w_i / \sum_{i=1}^{n} w_i \quad (i = 1,...n)$, where n is the number of objects in the display and w_i are the weights assigned to objects according to their saliencies. Thus a more salient object (the target) has a higher probability to be selected first. The process of selection consists of several selection rounds which continue until a target is selected or the conditions for search termination are fulfilled. The model was applied to tasks which differed by the weights assigned to the target and distractors. In Feature Search due to high saliency of the target the weight of the target is radically greater than the weights of distractors. Purely mathematically this leads to a flat slope of the RT as a function of n. In Spatial Configuration Search the weight of the target is only slightly higher than the weight of distractors, which leads to a

steep slope of RT*(n)*. Configuration Search occupies an intermediate position between these two extreme cases.

Our model follows the same logic as the model of Moran and coauthors (2013) but aims to suggest a simple neural mechanism that may underlie the probabilistic selection of objects. In contrast to connectionists models based on the winner-take-all principle, our model operates with oscillatory networks and their synchronization. Such networks have already demonstrated their usefulness as a model of attention (Wang, 1999; Borisyuk & Kazanovich, 2004; Kazanovich & Borisyuk, 2006; Borisyuk, Kazanovich, Chik, Tikhanoff, & Cangelosi, 2009; Quiles, Wang, Zhao, Romero, & Huang, 2011; Qu, Wang, & Du, 2013). Our model is designed as a so-called network with a central element. It has a radial architecture of connections, such that there is a special central unit, called the central oscillator (CO), which has feedforward and feedback connections with peripheral units, called peripheral oscillators (PO).

This choice of architecture is inspired by the idea of a central executive network (Cowan, 1988; Baddeley, 1996) which is active when the brain is engaged in a task requiring attention (Goulden et al., 2014). The central executive is assumed to be a complex network in the forebrain that does not necessarily have direct connections to the visual cortex. Our model operates independently of the specific details of visual processing, therefore the central executive is represented in a radically simplified form of a single oscillator, the CO. Each visual object, which is represented in reality by an assembly of synchronous oscillators in the visual cortex (Gray, 1999; Singer, 1999), is reduced to a single oscillator, a PO.

Recently Ray and Mounsell (2010) call in question that synchronization hypothesis can be a universal mechanism of feature binding (at least not in the area V1). They have shown that relatively large objects whose contrast varies in space generate gamma rhythms at different frequencies in different neuronal assemblies. We think that these results do not undermine our assumption that an object is coded by a single oscillator since only small objects of a fixed contrast are used in visual search experiments that we aim to model. Besides, our model does not restrict to V1 the region where the central executive interacts with neural representations of visual objects.

Interactions in the network are organized in such a way that POs compete for synchronization with the CO. It is postulated that those POs which are synchronized by the CO are included in the FOA. This assumption is in line with the temporal correlation hypothesis (Gray, 1999; Singer, 1999). Experimental evidence shows that spiking coherence plays a major role in the control of selective attention (Fries, Reynolds, Rorie, & Desimone, 2001; Fries, Schroeder, Roelfsema, Singer, & Engel, 2002; Fell, Fernandez, Klaver, Elger, & Fries, 2003; Doesburg, Roggeveen, Kitajo, & Ward, 2008, Bosman & Womelsdorf, 2009). In the paper (Gregoriou, Gotts, Zhou, & Desimone, 2009) it has been shown that during selective attention a region in the frontal

cortex (the frontal eye field) is a source of increased synchronization of neural activity in the extrastriate area V4 at the frequency of the gamma rhythm (about 50 Hz).

The saliency of objects is represented in our model by connection strengths with which POs exert influence on the CO. One could equivalently think about the amplitudes of oscillations of POs instead of connection strengths. This would not introduce any change in the dynamics of the model. The feedback influence of the CO on POs is desynchronizing and identical in strength for all POs. We show that a proper choice of the relations between these interaction strengths leads to probabilities of target and distractor selection that correspond to the search tasks of various complexities.

We assume that our model works under the Guided Search paradigm which states that several selection attempts may be necessary before a target is found. The number of such attempts, and hence the RT, is determined by the relation between the probabilities of selection of the targets and the distractors. The Guided Search as well as its implementation in our model combine both parallel and sequential mechanisms, the first one for selecting an assessed item and the latter one to proceed from one item in the display to another. This is in agreements with the results of Bichot and coauthors (Bichot, Rossi, Desimone, 2005) who have found serial and parallel mechanisms of response enhancement in the macaque area V4 during visual search. They also have shown that neurons synchronized their activity in the gamma range whenever a preferred stimulus in their receptive field matched the target. Also Wamelsdorf and coauthors (2006) confirmed that enhanced gamma-band synchronization in the area V4 shortens RTs.

Model simulations show that the RT demonstrates an approximately linear increase with the number of items in the display, which agrees with what is observed in visual search experiments (Proulx & Egeth, 2006; Wolfe, 2007; Hulleman, 2010; Palmer, Horowitz, Torralba, & Wolfe, 2011).

The present model represents a slight revision and simplification of the model of attention which was developed in our previous publications (Borisyuk & Kazanovich, 2004; Borisyuk, Kazanovich, Chik, Tikhanoff, & Cangelosi, 2009) for consecutive selection of objects in the visual scene. In these papers no separation between targets and distractors was assumed. Now we aim to reproduce the experimental results on the RTs in visual search tasks where a target should be found among distractors. This research is encouraged by the recent paper (Moran, Zehetleitner, Müller, & Usher, 2013) which supported the concept of the Guided Search, showing how RTs in visual search can be explained purely statistically based on the probabilities of selection of targets and distractors in the attention focus. Our contribution to this field is in suggesting a neurobiological mechanism, that describes how such probabilities may appear from the dynamics of interaction between the

central executive of attention system and the regions of the cortex that represent individual object in the visual scene.

2. The model

The model has a radial structure of connections (Fig. 1). This means that there is a central oscillator (CO) and a set of n peripheral oscillators (PO). The POs represent the objects that are simultaneously present in a display. One of the oscillators corresponds to the target (we adopt the convention that this is the first oscillator) and $n-1$ oscillators correspond to distractors. The CO represents the central executive of the attention system. It interacts with POs through feedforward and feedback connections. There are no direct connections between POs.

It is postulated that an object is included in the FOA if the corresponding PO is coherent with the CO. The interaction between POs and the CO is organized so that POs compete for synchronization with the CO and in a typical case only one PO wins the competition. It is assumed that the influence of a PO on the CO depends on whether this PO represents a target or a distractor. A "target" PO has a stronger influence on the CO than a "distractor" PO. This unbalance in the interaction strength increases the chances of a "target" PO to be synchronized with the CO while other POs work incoherently with the CO.

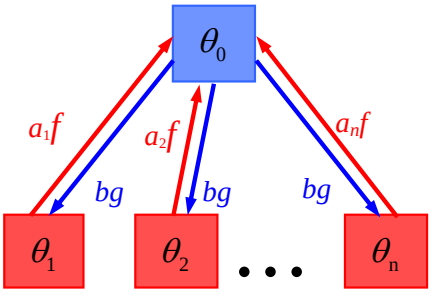


Fig. 1. Architecture of the model. The phases of oscillators are denoted as $\theta_0, \theta_1, ..., \theta_n$, f and g are interaction functions.

The dynamics of the network are described by the following equations

$$\frac{d\theta_0}{dt} = \omega_0 + \frac{1}{n}\sum_{i=1}^{n} a_i f(\theta_i - \theta_0), \tag{1}$$

$$\frac{d\theta_i}{dt} = \omega_i + bg(\theta_0 - \theta_i), \quad i = 1,...n, \tag{2}$$

$$\frac{d\omega_0}{dt} = \alpha \frac{1}{n} \sum_{i=1}^{n} a_i f(\theta_i - \theta_0),$$

(3)

$$\frac{da_i}{dt} = \beta(-a_i + c + \gamma h(\theta_0 - \theta_i)),\ i = 1,...n.$$

(4)

In these equations θ_0 is the phase of the CO, ω_0 is the natural frequency of the CO, θ_i are the phases of POs, ω_i are their natural frequencies, a_i, b are the strengths of connections $(i = 1,...,n)$. All the phases are assumed to be in the range $(-\pi, \pi]$. The functions $f(x)$ and $g(x)$ control the interaction between POs and the CO. The function $h(x)$ and the positive parameters c and γ control the adaptation of the connection strengths from POs to the CO. The parameter β determines the speed of connection strengths adaptation.

Equations (1)-(4) can be considered as a generalization of standard Kuramoto equations for phase oscillators by introducing in addition to phase equations (1)-(2) equation (3) for the adaptation of the natural frequency of the CO and equations (4) for the adaptation of connection strengths from POs to the CO.

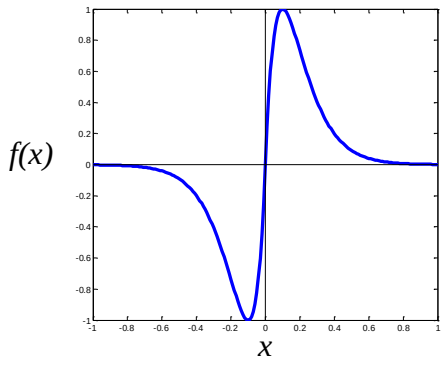
The interaction function f (Fig. 2a) is described as:

$$f(x) = \begin{cases} \lambda x e^{-\lambda x+1}, & for\ 0 \le x \le \pi, \\ \lambda x e^{\lambda x+1}, & for\ -\pi < x < 0, \end{cases} \quad \lambda > 0.$$

(5)

The function $f(x)$ is odd and has a steep rise near 0. It reaches its extremes ± 1 at the points $\pm 1/\lambda$ and then decreases to 0 as $|x|$ increases.

a



b

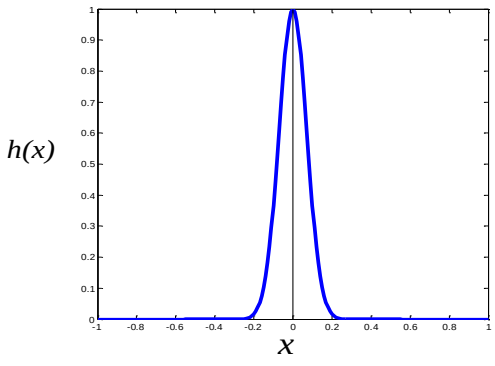


Fig. 2. Functions of dynamical equations: (a) interaction function $f(x)$, $\lambda = 10$; (b) resonance controlling function $h(x)$, $m = 100$.

The interaction function $g(x)$ is

$$g(x) = sin(x).$$

(6)

The function $h(x)$ (Fig. 2b) is described by the formula

$$h(x) = \begin{cases} (1-(x)^2)^m, & |x| < 1, \\ 0, & 1 \le |x| \le \pi. \end{cases}$$

(7)

The power m is a relatively big positive integer. In the interval $(-\pi, \pi)$ the function $h(x)$ has its maximum $h(0) = 1$ and rapidly decays to 0 as $|x|$ increases. As a result, the connection strength a_i tends to its maximum value $c + \gamma h(0)$ if the ith PO works in-phase with the CO. This is interpreted as the resonant influence of the CO on the ith PO. If the ith PO and the CO are incoherent the connection strength a_i decreases to a relatively low value c.

While the natural frequencies of the POs are constant, the natural frequency of the CO is adapted according to equation (3). The meaning of equation (3) becomes clear if it is rewritten as

$$\frac{d\omega_0}{dt} = -\alpha(\omega_0 - \frac{d\theta_0}{dt}).$$

(8)

According to (8) the natural frequency of the CO is adapted in the direction of the current frequency $\frac{d\theta_0}{dt}$. The parameter α controls the speed of adaptation.

The dynamics conditioned by equations (1)-(3) can be informally described in the following way. Due to equation (1) each PO influences on the CO trying to synchronize its activity with the activity of this PO. Evidently, the chance to succeed is greater for the PO whose influence on the CO is stronger at least at the initial moment of the run. In other words, a more salient object (the target) has a greater chance to be included in the attention focus than less salient objects (distractors). Equation (2) is responsible for organizing the competition between POs for the synchronization with the CO since the desynchronizing influence from the CO on POs does not allow several POs to simultaneously synchronize its activity with the CO. Thus only one object can be included in the focus of attention at a given run of the model. Equation (3) helps the CO to reach fast and reliable synchronization with one of the POs which results in the rapid increase of the influence of this PO on the CO. This is ensured by equation (4). At the same time equation (4) inhibits the influence on the CO of those POs which lost the competition for the synchronization

with the CO. The types of the functions in equations (1)-(3) have been chosen to ensure the proper functioning of the model and to avoid stochastic dynamics.

3. Simulation results

Two sources of stochasticity are included in the model. First, for each run initial conditions of the phases θ_i are chosen randomly from the interval $(0,\ 0.1\pi)$. Second, the natural frequencies of POs are randomly distributed in the interval (4.9, 5.1). We suppose that a formal time unit in the model is equal to $\frac{1}{20\pi}$ s. Then an oscillator that oscillates with the frequency $\omega = 5$ in formal time units will make $2\pi 50$ oscillations in 1 s. This means that POs work within the gamma range of frequencies. The initial phase of the CO is 0 and its initial natural frequency is 5. During a run of the simulation the natural frequency of the CO is adapted according to equation (3). The speed of adaptation is $\alpha = 1$. The parameters of the functions *f(x)* and *h(x)* are $\lambda = 10$ and $m = 100$, respectively.

The initial values of connection strengths from POs to the CO are set according to the assumptions of saliency of objects for the attention system. In simulations we assign to $a_1(0)$ (the connection strength that corresponds to a target) an integer value between 3 and 12. The strengths $a_i(0)$ (i = 2,…,n) (the connection strengths that correspond to distractors) has one of the values 1 or 2 and are the same for all POs representing distractors. This reflects our assumption that all distractors have the same saliency and hence have the same chance to be selected in the FOA. During a run of the simulation the connection strengths are adapted according to equation (4). The values of parameters in (4) are $c = 2, \gamma = 10, \beta = 0.05$. These parameters determine the lower and higher asymptotic boundaries for a_i. They are $c = 2$ for a PO which is incoherent with the CO and $c + \gamma = 12$ for a PO which is coherent with the CO.

The connection strength from the CO to the POs is a small negative number. In simulations we set $b = -1$. This is necessary to avoid the situation when several POs are simultaneously synchronized with the CO. When *b* is negative, the CO tends to desynchronize POs. This desynchronization can be overcome by a PO if its synchronizing influence on the CO is high enough.

The difference in asymptotic dynamics of POs can be determined by checking whether the trajectory of a PO's connection strength approaches the lower or higher boundary. Before checking this, some time should pass before the transient dynamics of the model is over. This time is set to be

$T_1 = 80$. The following 20 time units are used to determine the type of dynamics which is demonstrated by the model. Thus the whole run of the model takes $T_2 = 100$ time units.

Two thresholds $H_{high} = 10$ and $H_{low} = 3$ are introduced. If in the time interval (T_1, T_2) the trajectory of $a_i(t)$ lies above H_{high}, then the *i*th object is supposed to be included in the FOA.

All parameter values of the model are summarized in Table 1.

Massive simulations of the model have shown that the following types of dynamics can be observed under the chosen parameter values:

A. $a_1(t) > H_{high}$ for $T_1 < t < T_2$ and the "target" PO is the only oscillator for which this inequality is valid. This means that the selection of the target has been made (Fig. 3a). In this case the process of selection is over.

Table 1

Parameter values

Parameters	Values
Initial value of the phase of the CO	0
Initial values of the phases of POs	$(0,\ 0.1\pi)$
The initial value of the natural frequency of the CO	0
Range of distribution of the natural frequencies of POs	(4.9, 5.1)
Parameter λ of function f	10
Parameter m of function h	100
Initial value of connection strength $a_1(0)$	Integers between 3 and 12
Initial values of connection strengths $a_i(0)$ ($i = 2,\ldots,n$)	1 or 2
Parameter b of equation (2)	-1
Parameter α of equation (3)	1
Parameter c of equation (4)	2
Parameter γ of equation (4)	10
Parameter β of equation (4)	0.05
Time interval (T_1, T_2) for dynamics type determination	(80, 100)
Upper threshold H_{high} for connection strengths	10
Lower threshold H_{low} for connection strengths	3

B. $a_i(t) > H_{high}$ for $T_1 < t < T_2$ for a single index $i > 1$. This means that a distractor has been selected in the FOA (Fig. 3b). In this case the process of selection should be repeated in other runs of the model until the correct selection of the target will be reached.

C. $a_i(t) < H_{low}$ for all i during the time interval (T_1, T_2) (Fig. 3c). This means that the attention system failed to form the FOA during the time T_2. Also in this case the process of selection should be repeated in other runs.

Fig. 3d shows an example of the histogram of events A, B, C for 1000 runs of the model. The computations show that the number of events C is always low, usually not more than 3-5%.

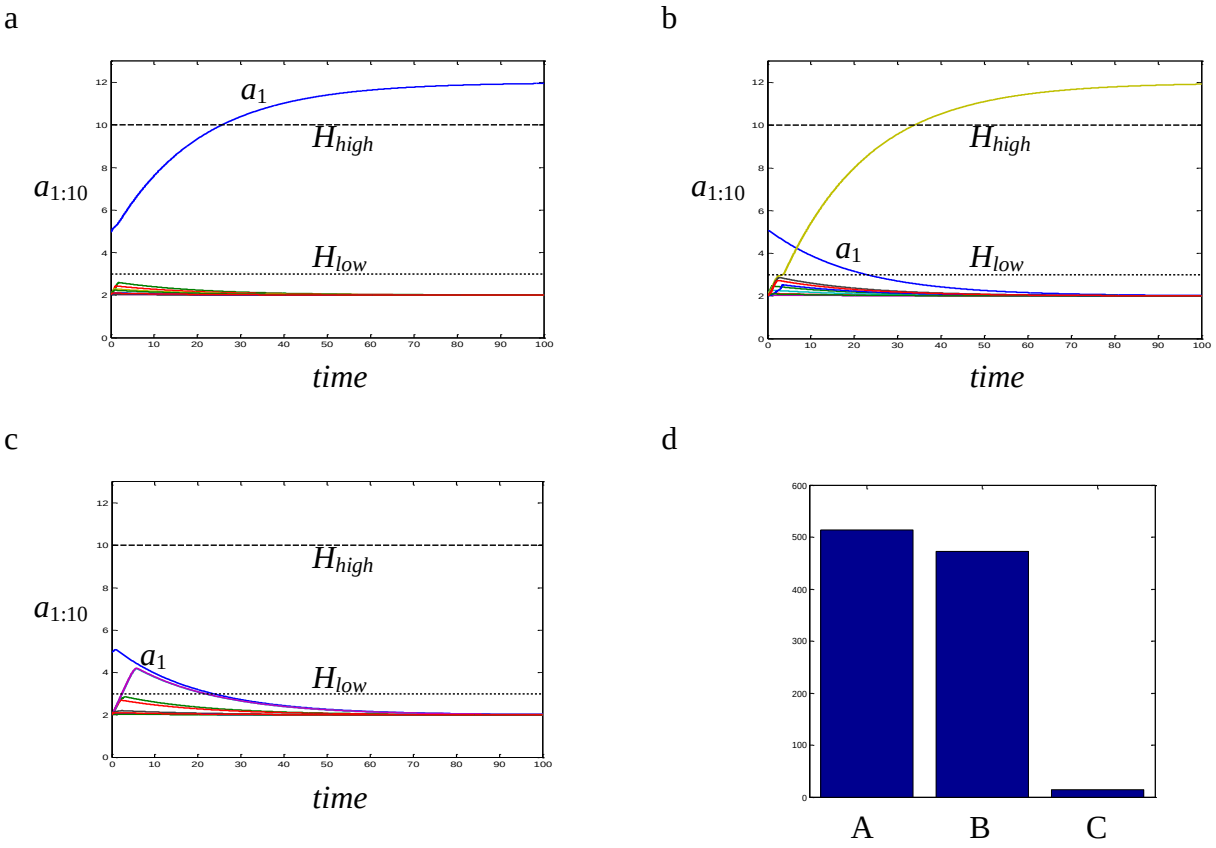
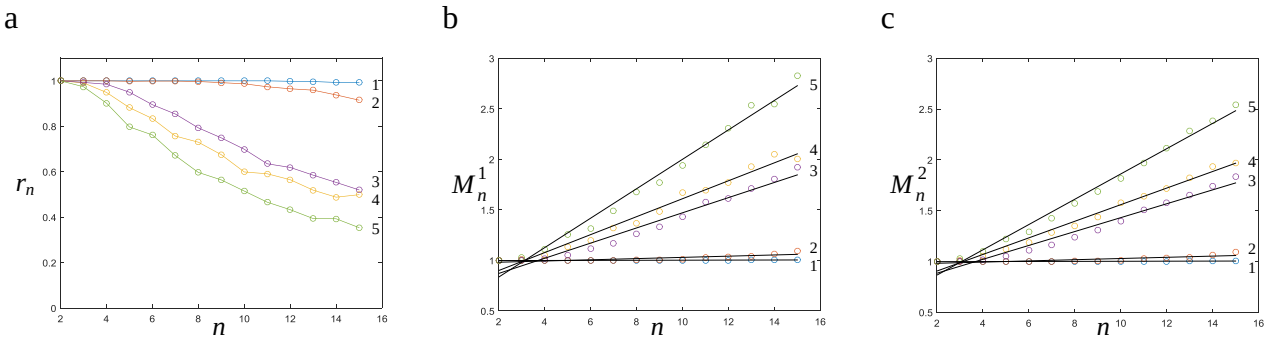


Fig. 3. The types of dynamics: (a) target selection; (b) distractor selection; (c) FOA is not formed; (d) histogram of dynamical modes. Parameter values: n = 10, $a_1(0)$ = 5, $a_i(0)$= 2 (i = 2,…,10).

Figs. 4a, d show the probabilities r_n that the target has been selected to the FOA. To find these probabilities we run 1000 simulations for a given number of objects n in the display. Different lines represent the results for different pairs of $a_1(0)$, $a_2(0)$ ($a_2(0)$=…=$a_n(0)$. The graphs show that r_n decreases as n increases. The values of r_n are also lower if the difference between the initial values of connection strengths for the target and the distractors is lower, that is if it is more difficult to separate a target from distractors.



d e f

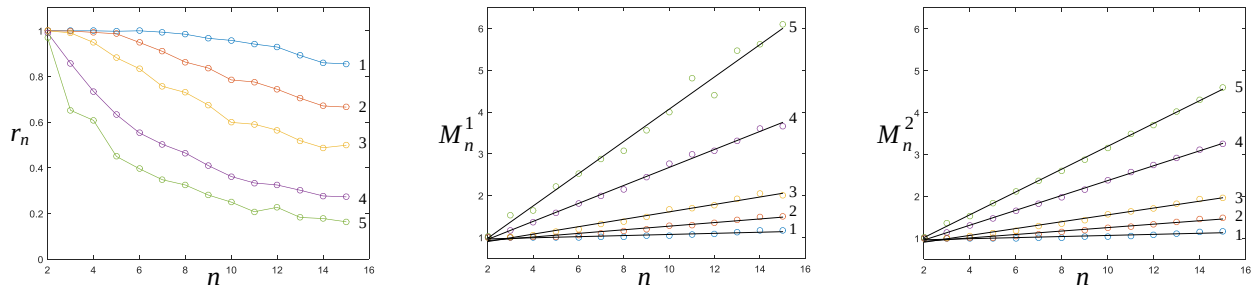


Fig. 4. Simulation results: (a), (d) the probability r_n as a function of n; (b), (e) average number of attempts M_n^1 needed to select a target with return; (c), (f). average number of attempts M_n^2 needed to select a target without return. In (a), (b), (c) $a_i(0)=1\,(i=2,...,n)$. In (d), (e), (f) $a_i(0)=2\,(i=2,...,n)$. In all panels the numeration of lines corresponds to the following values of $a_1(0)$: (1) $a_1(0)=12$, (2) $a_1(0)=8$, (3) $a_1(0)=6$. (4) $a_1(0)=4$, (5) $a_1(0)=3$. Straight lines in (b), (c) and (e), (f) are linear regression lines.

Being a hidden parameter, r_n cannot be observed experimentally. However, according to the Guided Search Theory, r_n can be used to compute the average number of attempts M needed to select a target. Since M is proportional to the average time of the target selection, the results of computations of M can be compared with experimental data.

There is no definite opinion among researchers whether repeated selections are made with or without inhibition of return. We consider both possibilities. Denote the average number of selection attempts in the cases when the return is or is not allowed as M_n^1 and M_n^2, respectively. If we don't put any restrictions on the number of attempts for selections with return, then M_n^1 satisfies the formula

$$M_n^1=\sum_{i=1}^{\infty} i r_n (1-r_n)^{i-1}=\frac{1}{r_n}. \qquad (9)$$

In the case when inhibition of return is assumed the process of selection is restricted by n attempts. The average number of attempts is

$$M_n^2=r_n+\sum_{i=2}^{n-1}\left[i r_{n+1-i}\prod_{j=n+2-i}^{n}(1-r_j)\right]+n\prod_{j=2}^{n}(1-r_j). \qquad (10)$$

The values of M_n^1 and M_n^2 have been computed according to (9)-(10). Their graphs as functions of n are shown in Fig. 4b, c, e, f by the circles together with linear regression lines (shown in black).

The simulation results of Figs. 4b, c and Figs. 4e, f properly correspond to experimental evidence on RTs in visual search tasks of different complexity (see Fig. 2 in (Wolfe., Palmer, & Horowitz, 2010)). The lines labelled 1 in Figs. 4b, c, e, f and the lines labelled 2 in Figs. 4b, c are nearly horizontal, therefore one can think that they correspond to Feature Search. Lines 3-4 in Figs. 4b, c, e, f and line 2 in Figs. 4e, f correspond to cases of Conjunction Search (). Note the difference in behaviour of the line 2 in Figs. 4a, c and Figs. 4e, f. In Figs. 4e, f the saliency of distractors is higher than that of Figs. 4b, c, therefore the search for a target in the situations of Figs. 4e, f is more difficult. Comparing the behaviour of lines 3 and 4 in Figs. 4b, c and Figs. 4e, f one can see that the change in saliency of distractors may significantly change the relation between the difficulties of the tasks. In Figs. 4b, c the difference in difficulty of tasks 3 and 4 is much lower that in the case of Figs. 4e, f. The lines labelled 5 correspond to the most difficult type of tasks: Spatial Configuration Search. Note that the search with inhibition of return (Figs, 4c, f) is only slightly faster than the search without such inhibition. This explains why it is so difficult to experimentally distinguish between these two types of search.

A question may arise how the average number of selection attempts presented in Fig. 4 relate to the experimental data on reaction times. Let us suppose that the average time to select and identify an item is a constant t_{id} that does not depend on how many selection attempts have been made. This assumption has already been used in the model (Moran, Zehetleitner, Müller, & Usher, 2013) showing a good agreement with experimental data. Under this assumption the average of the whole time for target selection is

$$\tau = t_{id} * M + t_{res}, \tag{11}$$

where M is the average number of attempts, t_{res} is the average residual time that includes all reaction time episodes which are not explicitly accounted for by the search processes. This residual time incorporates image encoding and post-decisional processes such as response planning and execution. Formula (11) means that there is a simple linear dependency between τ and M.

The simulation results show that the dependences M_n^1 and M_n^2 from n are well approximated by linear functions. As we already mentioned, this is in agreement with the experimental results of (Proulx & Egeth, 2006; Wolfe, 2007; Hulleman, 2010; Palmer, Horowitz, Torralba, & Wolfe, 2011). The linearity of all 20 models is statistically significant at conventional level of 0.01 (ANOVA p-values are less than 0.01). All models are a good fit to the data (R-square ranges from 0.56 to 0.99). Due to (11) the same is true for $\tau(n)$.

4. Discussion

The evidence on the relation between attention and synchronization of neural activity (Steinmetz et al., 2000; Fries, Reynolds, Rorie, & Desimone, 2001; Fries, Schroeder, Roelfsema, Singer, & Engel, 2002; Fell, Fernandez, Klaver, Elger, & Fries, 2003; Womelsdorf & Fries. 2007; Doesburg, Roggeveen, Kitajo, & Ward, 2008; Bosman & Womelsdorf, 2009; Deco, Buehlmann, Masquelier, & Hugues, 2011) represents a challenge for mathematical modelling. Is it possible to explain some phenomena of attention in terms of oscillatory neural networks? We demonstrate that this is the case by showing that an oscillatory network can be used to create competition between objects for inclusion in the FOA. Based on the idea of a central executive of the attention system we developed a model that describes attention in terms of partial synchronization (Kazanovich & Borisyuk, 1994; Kazanovich & Borisyuk, 1999) between the central executive in the forebrain and neural representation of objects in the visual cortex. In contrast to traditional winner-take-all models which must operate through massive interconnections of the all-to-all type, the number of connections in an oscillatory network with a central element can be of the same order as the number of elements in the network. In our model the number of connections is $2n$.

Generalized phase oscillators used in our modelling have shown to be a powerful instrument for the systems whose functioning is based on phase relations. Their advantage in comparison to standard Kuramoto oscillators is the possibility to modify natural frequencies of oscillators and the strengths of connections, which may be necessary to obtain the required dynamical behaviour. Though there is a possibility to organize a competition for synchronization with the central oscillator among Kuramoto type POs (Kazanovich, Burylko, & Borisyuk, 2013), it would be impossible to avoid chaotic behaviour in the system for most initial conditions.

The linear dependence of the reaction time from the number of items in the display is an important fact since, as we already mentioned, such dependence is observed in many visual search experiments. This linearity was not planned in advance for our model and appeared as an unexpected result in model simulations. By choosing a proper relation between initial connection strengths from "target" and "distractor"' POs it is possible to reproduce RT(n) with a slope corresponding to a search task of any given complexity. Thus our model provides additional support to the Guided Search Theory.

Another important fact that should be emphasized is that each graph in Fig. 4 was obtained under fixed initial values of connection strengths which should be interpreted as the assumption that the saliency of items does not depend on the set size. This is in agreement with the assumptions of the model of Moran et al. (2013) and can be considered as additional support for this model. The progress achieved by our model is that it suggests a neurobiological mechanism of visual search implementation, while the previous model does not go beyond purely probabilistic considerations.

Recently the Guided Search Theory has been critically reviewed in the paper (Hulleman & Olivers, 2015). The authors state that "human eye movement patterns during visual search are better explained by a model that fixates areas of the screen that will maximize information about target-presence, than by a model that fixates the item most likely to be the target". We think that our model can be expanded to enclose this approach by a more intricate assignment of connection strengths from "target" and "distractor" POs that would take into account mutual space arrangement of items in the display.

Acknowledgements

We acknowledge financial support from EPSRC (Grant EP/D036364) and BBSRC (Grant BB/L000814/1). The authors are grateful to Dr Robert Merrison-Hort for useful comments and correction of English. The authors would like to thank Dr Galina Borisyuk for advice on statistical testing.

References

Baddeley, A. (1996). Exploring the central executive. *Q. J. Exp. Psychol.*, 49A, 5-28.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, 308, 529-534.

Borisyuk, R., & Kazanovich, Y. (2004). Oscillatory model of attention-guided object selection and novelty detection. *Neural Networks*, 17, 899-915.

Borisyuk, R., Kazanovich, Y., Chik, D., Tikhanoff, V., & Cangelosi, A. (2009). A neural model of selective attention and object segmentation in the visual scene: An approach based on partial synchronization and star-like architecture of connections. *Neural Networks*, 22, 707-719.

Bosman, C., & Womelsdorf, T. (2009). Neural signatures of selective attention – synchronization and gain modulation as mechanisms of selective sensory information processing. In F. Aboitiz & D. Cosmelli (Eds.), *From Attention to Goal-Directed Behavior* (pp. 3-28). Springer-Verlag, Berlin, Heidelberg.

Bruce, N.D.B., & Tsotsos, J.K. (2009). Saliency, attention, and visual search: An information theoretic approach. *J. Vision*, 9(3):5, 1–24.

Cameron, E.L., Tai, J.C., Eckstein, M.P., & Carrasco, M. (2004). Signal detection theory applied to three visual search tasks - identification, yes/no detection and localization. *Spatial Vision*, 17, 295–325.

Chan, L.K.H., & Hayward, W.G. (2013). Visual search. *WIREs Cogn. Sci*., 4, 415–429.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163–191.

Deco, G., Buehlmann, A., Masquelier, T., & Hugues, E. (2011). The role of rhythmic neural synchronization in rest and task conditions. *Front. Human Neurosci*., 5, Article 4, 1-6.

Doesburg, S.M., Roggeveen, A.B., Kitajo, K., & Ward, L.M. (2008). Large-scale gamma-band phase synchronization and selective attention. *Cerebr. Cortex*, 18, 386-396.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. *Psychol. Rev.*, 96, 433–458.

Duncan, J., & Humphreys, G.W. (1992). Beyond the search surface: Visual search and attentional engagement. *J. Exp. Psychol. Hum. Percept. Perform.*, 18, 578-588.

Eckstein, M.P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychol. Sci*., 9, 111-118.

Eckstein, M.P., Thomas, J.P., Palmer, J., & Shimozaki, S.S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Percep. Psychophys.*, 62, 425-451.

Fell, J., Fernandez, G., Klaver, P., Elger, C.E., & Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? *Brain Res. Rev.*, 42, 265-272.

Fries, P., Reynolds, J., Rorie, A., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, 291, 1560-1563.

Fries, P., Schroeder, J.-H., Roelfsema, P. R., Singer, W., & Engel, A.K. (2002). Oscillatory neural synchronization in primary visual cortex as a correlate of stimulus selection. *J. Neurosci*., 22, 3739-3754.

Goulden, N., Khusnulina, A., Davis N.J., Bracewell, R.M., Bokde, A.L., McNulty, J.P., & Mullins, P.G.. (2014). The Salience Network is responsible for switching between the Default Mode Network and the Central Executive Network: Replication from DCM. *Neuroimage*, 99, 180-190.

Gray, C.M. (1999). The temporal correlation hypothesis is still alive and well. *Neuron*, 24, 31-47.

Gregoriou, G.G., Gotts, S.J., Zhou, H., & Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, 324, 1207-1210.

Heinke, D., Humphreys, G.W., & Tweed, C.L. *(*2006*).* Top-down guidance of visual search: A computational account. *Vis. Cogn.*, 14, 985-1005.

Herd, S.A., & O'Reilly, R.C. (2006). Serial visual search from a parallel model. *Vision Res.*, 49, 2987-2992.

Hulleman, J. (2010). Inhibitory tagging in visual search: Only in difficult search are items tagged individually. *Vision Res.*, 50, 2069-2079.

Hulleman, J., & Olivers, C.N.L. (2015). The impending demise of the item in visual search. *Behav. Brain Sci.* (In press. Available on CJO2015 doi:10.1017/S0140525X15002794).

Itti, L. (2005). Models of bottom-up attention and saliency. In L. Itti, G. Rees, & J.K. Tsotsos (Eds.), *Neurobiology of Attention* (pp. 576-582). Elsevier Acad. Press, Amsterdam.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Res.*, 40, 1489-1506.

Kazanovich, Y.B., & Borisyuk, R.M. (1994). Synchronization in a neural network of phase oscillators with the central element, *Biol. Cybern.*, 71, 177-185.

Kazanovich, Y.B., & Borisyuk, R.M. (1999). Dynamics of neural networks with a central element. *Neural Netw.*, 12, 441-454.

Kazanovich, Y.B., & Borisyuk, R.M. (2006). An oscillatory neural model of multiple object tracking. *Neural Comput.*, 18, 1413-1440.

Kazanovich, Y., Burylko, O., & Borisyuk, R. (2013). Competition for synchronization in a phase oscillator system. *Physica D*, 261, 114–124.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.*, 4, 219-227.

Luce, R.D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

Ma, W.J., Navalpakkam, V., Beck, J.M., van den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nat. Neurosci*., 14, 783-790.

Moran, R., Zehetleitner, M., Müller, H.J., & Usher, M. (2013). Competitive guided search: Meeting the challenge of benchmark RT distributions. *J. Vision*, 13(8):24, 1–31.

Nothdurft, H.C. (1999). Focal attention in visual search. *Vision Res.*, 39, 2305–2310.

Palmer, E.M., Horowitz, T.S., Torralba, A., & Wolfe J.M. (2011). What Are the Shapes of Response Time Distributions in Visual Search? *J. Exp. Psychol. Hum. Percept. Perform.*, 37, 58-71.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Res.*, 40, 1227-1268.

Proulx, M.J., & Egeth H.E. (2006). Target–nontarget similarity modulates stimulus-driven control in visual search. *Psychon. Bull. Rev.*, 13, 524-529.

Qu, J., Wang, R., Du, Y. (2013). An improved selective attention model considering orientation preferences. *Neural Comp. & Applic*., 22, 303-311.

Quiles, M.G., Wang, D.-L., Zhao, L., Romero, R.A.F., Huang, D.-S. (2011). Selecting salient objects in real scenes: An oscillatory correlation model. Neural Networks, 24, 54-64.

Ray, S., & Maunsell, J.H.R. (2010). Differences in gamma frequencies across visual cortex restrict their possible use in computation. Neuron, 67, 885-896.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron* **24**, 49-65.

Steinmetz, P.N., Roy, A, Fitzgerald, P.J., Hsiao, S.S., Johnson, K.O., & Niebur, E.. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature*, 404. 187-190.

Sternberg, R.J., & Sternberg, K. (2012). *Cognitive Psychology*. 6th Edition. Australia; Belmont, CA: Wadsworth/Cengage Learning.

Schwemmer, M.A., Feng,S.F., Holmes, P.J., Gottlieb, J., & Cohen, J.D. (2015). A multi-area stochastic model for a covert visual search task. *PLoS ONE*, 10(8): e0136097.

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychol.* **12**, 97–136.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *J. Exp. Psychol. Hum. Percept. Perform.*, 16, 459-478.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31, 523-535.

Vincent, B.T., Baddeley, R.J., Troscianko, T., & Gilchrist, I.D. (2009). Optimal feature integration in visual search. *J. Vision*, 9(5):15, 1–11.

Walles, H., Robins, A., & Knott, A. (2013). A neural network model of visual attention and group classification, and its performance in a visual search task. *Advances in Artificial Intelligence* AI-2013, 98-103. *Lecture Notes in Computer Sciences*, 8272.

Wang, D.L. (1999). Object selection based on oscillatory correlation. *Neural Networks*, 12, 579–592.

Wilder, M.H., Mozer, M.C., & Wickens, C.D. (2011). An integrative, experience-based theory of attentional control. J. *Vision*, 11(2):8, 1–30.
Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1, 202–238.
Wolfe, J.M. (2007). Guided Search 4.0: Current progress with a model of visual search. In Gray, W. (Ed.) *Integrated Models of Cognitive Systems* (pp. 99-119). University Press, New York: Oxford.
Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: an alternative to the feature integration model for visualsearch. *J. Exp. Psychol. Hum. Percept. Perform*., 15, 419–433.
Wolfe, J.M., Palmer, E.M., & Horowitz, T.S. (2010). Reaction time distributions constrain models of visual search. *Vision Res.*, 50, 1304–1311.
Womelsdorf, T., & Fries, P. (2007). The role of neuronal synchronization in selective attention. *Curr. Opin. Neurobiol*., 17, 1-7.
Womelsdorf, T., Fries, P., Mitra, P.P., & Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature*, 439, 733-736.