

A neuronal network model linking subjective reports and objective physiological data during conscious perception

Stanislas Dehaene^{*†}, Claire Sergent^{*}, and Jean-Pierre Changeux[†]

^{*}Institut National de la Santé et de la Recherche Médicale, Unité 562, Cognitive Neuroimaging, Institut Fédératif de Recherche 49, Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique/DSV, 4 Place du Général Leclerc, 91401 Orsay Cedex, France; and [†]Centre National de la Recherche Scientifique Unité de Recherche Associée 2182, Récepteurs et Cognition, Institut Pasteur, 25 Rue du Dr Roux, 75015 Paris, France

Contributed by Jean-Pierre Changeux, April 30, 2003

The subjective experience of perceiving visual stimuli is accompanied by objective neuronal activity patterns such as sustained activity in primary visual area (V1), amplification of perceptual processing, correlation across distant regions, joint parietal, frontal, and cingulate activation, γ -band oscillations, and P300 waveform. We describe a neuronal network model that aims at explaining how those physiological parameters may cohere with conscious reports. The model proposes that the step of conscious perception, referred to as access awareness, is related to the entry of processed visual stimuli into a global brain state that links distant areas including the prefrontal cortex through reciprocal connections, and thus makes perceptual information reportable by multiple means. We use the model to simulate a classical psychological paradigm: the attentional blink. In addition to reproducing the main objective and subjective features of this paradigm, the model predicts a unique property of nonlinear transition from nonconscious processing to subjective perception. This all-or-none dynamics of conscious perception was verified behaviorally in human subjects.

Studies of consciousness yield two types of data: subjective reports of perceived stimuli and objective measurements of neural activity. In recent years, a diversity of putative objective correlates of consciousness have been described. Candidates include sustained activity in primary visual area V1 (1–3), amplification of perceptual processing (4, 5), correlation across distant regions (6, 7), joint parietal, frontal, and cingulate activation (4, 8, 9), cortico-cortical or thalamo-cortical γ -band oscillations (10, 11), and P300 waveform (12). Although these diverse measurements are often viewed as incompatible findings about the nature of conscious states, an increasingly attractive alternative is that they constitute different reflections of a single underlying phenomenon: the settling of brain activity into a temporary metastable state of global activity. This article introduces a simple neuronal model that illustrates how this view can account simultaneously for psychological and neurophysiological data on consciousness.

Our model focuses on the issue of conscious access: the fact that a piece of information, once conscious, becomes selectively available for multiple processes of attention, intention, memory, evaluation, and verbal report. In agreement with psychological models of a bottleneck stage in perception (13), we postulate that this availability results from the entry of sensory stimuli processed by the posterior visual areas into a global neuronal workspace (14–16), which mobilizes excitatory neurons with long-distance axons, capable of interconnecting sensory and high level areas into global brain-scale states of activity (Fig. 1A). The neurons that are temporarily mobilized inhibit other surrounding workspace neurons, which thus become unavailable for processing other stimuli. As a consequence of this selection process, when a piece of information such as the identity of a stimulus accesses a sufficient subset of workspace neurons, their activity becomes self-sustained and can be broadcasted via

long-distance connections to a vast set of defined areas, thus creating a global and exclusive availability for a given stimulus, which is then subjectively experienced as conscious.

In the present article, we use the model to simulate a classical perceptual phenomenon: the attentional blink (17). In a typical paradigm, participants are asked to process two successive targets, 1 (T1) and 2 (T2). When T2 is presented between 100 and 500 ms after T1, the ability to report it drops, as if the participants' attention had "blinked." Objective physiological data indicate that during this blink, T2 fails to evoke a P300 potential but still elicits event-related potentials associated with visual and semantic processing (P1, N1, and N400) (12). Our simulations aim at clarifying why some patterns of brain activity are selectively associated with subjective experience. In short, during the blink, bottom-up activity, presumably generating the P1, N1, and N400 waveforms, would propagate without necessarily creating a global reverberant state. However, a characteristic neural signature of long-lasting distributed activity and γ -band emission, presumably generating the P300 waveform, would be associated with global access.

The elementary components of our simulation are single-compartment spiking neurons that, when depolarized beyond a certain threshold by neuromodulatory inputs, exhibit spontaneous activity through intrinsic membrane oscillations in the γ range (18). They are organized in schematic cortical columns with distinct infragranular, granular, and supragranular layers, reciprocally connected to a thalamic network that receives external inputs (Fig. 1). In agreement with anatomical (19) and physiological (2) models of cortical organization, we distinguish a nearest-neighbor, bottom-up network that propagates sensory stimulation across the hierarchy of areas, and a long-distance, top-down network that sends amplification signals back to all levels below it. A possible implementation, amongst others (20), proposes that bottom-up propagation is supported by fast glutamatergic α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, whereas top-down amplification is supported by slower glutamatergic *N*-methyl-D-aspartate (NMDA) receptors (21, 22). This gives top-down projections a mainly modulatory role, because the efficacy of voltage-gated NMDA receptors is enhanced if there is a concomitant depolarization by sensory inputs. Thus, global activity is more likely to be achieved if there is "resonance" (10) between bottom-up sensory information and top-down signals.

It is currently out of reach to simulate the entire postulated architecture at the single-neuron level. We therefore present a minimal network, which models only the cell assemblies evoked by stimuli T1 and T2 through four hierarchical stages of processing (Fig. 1B). T1 and T2 are initially processed by separate

Abbreviations: T1, target 1; T2, target 2; V1, primary visual area; SRA, spike-rate adaptation; GABA, γ -aminobutyric acid; AMPA, α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid; NMDA, *N*-methyl-D-aspartate.

[†]To whom correspondence should be addressed. E-mail: dehaene@shfj.cea.fr.

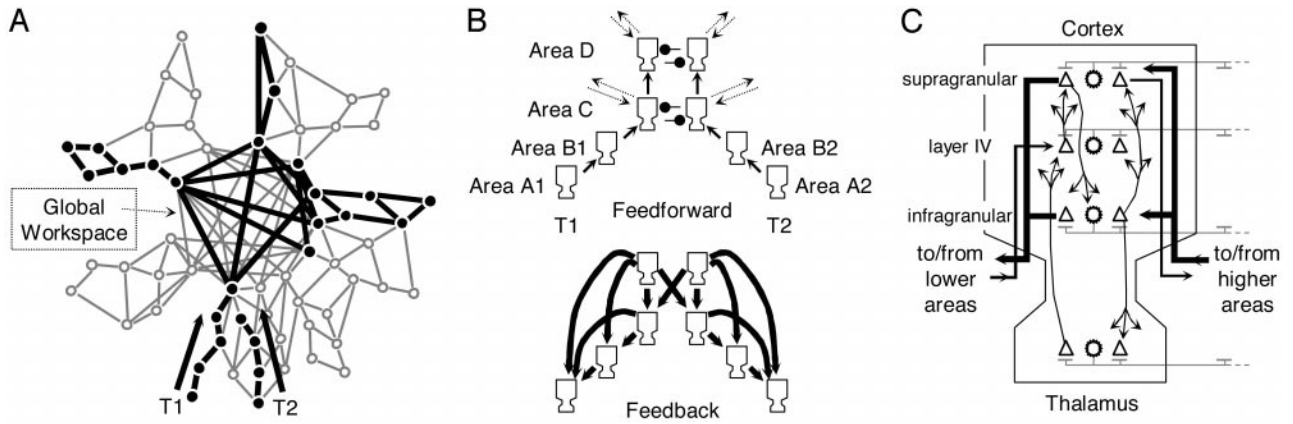


Fig. 1. Neuronal workspace model of conscious access. (A) Schematic architecture of brain areas (redrawn from ref. 14) comprising multiple specialized processors and a central network of high-level areas temporarily interconnecting them. During the attentional blink, T1 invades the workspace, where areas lock into a single assembly supporting conscious reportability. This invasion by T1 blocks the processing of T2 at a similar depth. (B) Subset of thalamo-cortical columns included in the present simulation. (C) Structure of a single simulated thalamo-cortical column, reproducing the laminar distribution of projections between excitatory neurons (triangles) and inhibitory neurons (stars; see *Materials and Methods* for details).

neuronal assemblies in two perceptual areas A and B (assemblies A1 and B1 for T1, and assemblies A2 and B2 for T2). Those assemblies do not inhibit each other because the attentional blink is known to occur between distinct visual stimuli, spatially separate visual streams, or even across the auditory and visual modalities (23). Further on, T1 and T2 reach higher association areas, also limited to two areas, C and D, where they compete for global access via reciprocal inhibitory interactions. In actual anatomy, areas A and B might correspond to a series of fast bottom-up cortical processors such as areas V1 or frontal eye field (2), whereas areas C and D would implement higher association areas of temporal, parietal, frontal, and cingulate cortex (14, 15).

Materials and Methods

Model Neurons. Neurons are modeled as single-compartment integrate-and-fire units with membrane potential V . When V exceeds a threshold of -48 mV, a spike is recorded and transmitted to other neurons with an appropriate delay, and V is reset to -80 mV for a refractory period of 4 ms. The temporal evolution of V , expressed in millivolts, is given by

$$C dV/dt = -g_{\text{Leak}}(V - V_{\text{rest}}) - I_{\text{NaP}} - I_{\text{KS}} - I_{\text{GABA}} - I_{\text{AMPA}} - I_{\text{NMDA}} - I_{\text{SRA}} - I_{\text{input}} - I_{\text{neuromodul}} \quad [1]$$

where $C = 1 \mu\text{F}/\text{cm}^2$ and currents are expressed in $\mu\text{A}/\text{cm}^2$. The leak conductance is $g_{\text{leak}} = 0.1 \text{ mS}/\text{cm}^2$, so that the membrane time constant C/g_{leak} is 10 ms.

I_{NaP} and I_{KS} are persistent sodium and slowly inactivating potassium currents, the interplay of which can generate intrinsic γ -band oscillations (18). Their simplified expression is

$$I_{\text{NaP}} = g_{\text{NaP}} m_{\text{NaP}} (V - V_{\text{Na}}), \quad [2]$$

where $m_{\text{NaP}} = (1 + e^{-(V+51)/5})^{-1}$, and

$$I_{\text{KS}} = g_{\text{KS}} m_{\text{KS}} (V - V_{\text{K}}), \quad [3]$$

where $\tau_{\text{KS}} dm_{\text{KS}}/dt = (1 + e^{-(V+34)/6.5})^{-1} - m_{\text{KS}}$, with $\tau_{\text{KS}} = 6$ ms.

I_{GABA} , I_{AMPA} , and I_{NMDA} are the total synaptic currents respectively from fast gabaergic, AMPA, and NMDA glutamergic synapses (21),

$$I_{\text{GABA}i}(t) = \sum_{j=1..n} g_{\text{GABA}i,j} h_{\text{GABA}j}(t - t_{\text{delay}i,j})(V_i(t) - V_{\text{GABA}}) \quad [4a]$$

$$I_{\text{AMPA}i}(t) = \sum_{j=1..n} g_{\text{AMPA}i,j} h_{\text{AMPA}j}(t - t_{\text{delay}i,j})(V_i(t) - V_{\text{AMPA}}) \quad [4b]$$

$$I_{\text{NMDA}i}(t) = \sum_{j=1..n} g_{\text{NMDA}i,j} m_{\text{NMDA}}(V_i(t)) h_{\text{NMDA}j}(t - t_{\text{delay}i,j}) \cdot (V_i(t) - V_{\text{NMDA}}), \quad [4c]$$

where the sum in each case is over all relevant neurons connected to postsynaptic neuron i (inhibitory interneurons for I_{GABA} , bottom-up and intracolumn excitatory neurons for I_{AMPA} , and top-down excitatory neurons for I_{NMDA}). The $g_{i,j}$ are the synaptic strengths and $t_{\text{delay}i,j}$ the transmission delays in milliseconds from neuron j to neuron i . The functions $h_j(t)$ characterize the synaptic inputs from neuron j and are given by the convolution of the spike train $s_j(t)$ with the synaptic activation profile $a(t)$,

$$h_j(t) = s_j(t) \otimes a(t), \quad [5]$$

where $a(t) = \alpha (e^{-t/\tau_1} - e^{-t/\tau_2}) / (e^{-t_{\text{peak}}/\tau_1} - e^{-t_{\text{peak}}/\tau_2})$, with $t_{\text{peak}} = \tau_1 \tau_2 / (\tau_1 - \tau_2)$, where τ_1 and τ_2 are rise and decay constants characteristic of each channel (21). In the NMDA case only, the effective conductance is scaled by $m_{\text{NMDA}}(V)$, a factor that characterizes the voltage dependence of NMDA channels (24).

$$m_{\text{NMDA}}(V) = (1 + 0.280 e^{-V/16.1})^{-1}. \quad [6]$$

I_{SRA} is an additional current used to model spike-rate adaptation (24).

$$I_{\text{SRA}} = g_{\text{SRA}}(V - V_{\text{SRA}}), \quad [7]$$

with $\tau_{\text{SRA}} dg_{\text{SRA}}/dt = -g_{\text{SRA}}$.

I_{input} is the current applied during stimulus presentation to thalamic neurons of the lowest hierarchical level to simulate a visual or auditory input.

$$I_{\text{input}}(t) = g_{\text{input}}(V(t) - V_{\text{AMPA}}) \quad [8]$$

Wakefulness and Spontaneous Activity. A final current, $I_{\text{neuromodul}}$, summarizes the known depolarizing effects of ascending activating systems such as those from cholinergic, noradrenergic, and serotonergic nuclei in the brainstem, basal forebrain, and hypothalamus (25). This parameter is used to control the level of wakefulness. When $I_{\text{neuromodul}}$ is low, membrane potentials converge to a stable asymptote in the absence of external stimulation. Beyond a threshold, individual neurons generate intrinsic γ -band membrane oscillations, first below the spiking threshold and then with an occasional occurrence of spikes (18). Because of those intrinsic properties, the network exhibits structured but semirandom patterns of spontaneous cortico-thalamic oscillatory activity, similar to empirical observations (10, 26), although our simulation, contrary to other similar work (21), is strictly deterministic. In the present work, we placed the network in a low spontaneous firing range ($I_{\text{neuromodul}} \approx -1 \mu\text{A}/\text{cm}^2$), presumably corresponding to wakefulness and attentive readiness. Work in preparation studies interactions between internal activity and stimulus processing at a higher spontaneous firing rate, which may provide a more realistic model for autonomous conscious states. We have observed that such states of elevated spontaneous activity can prevent the entry of external inputs into the workspace for a greater duration than during the blink. This effect may account for the empirical observation of long-lasting inattentive blindness in human subjects (27).

Parameter Values. The following parameter values were taken from references (18, 21, 24): $V_{\text{Rest}} = -67 \text{ mV}$, $V_{\text{Na}} = 55 \text{ mV}$, $V_{\text{K}} = -90 \text{ mV}$, $V_{\text{GABA}} = V_{\text{SRA}} = -70 \text{ mV}$, $V_{\text{AMPA}} = V_{\text{NMDA}} = 0 \text{ mV}$; $\alpha_{\text{GABA}} = 0.175$, $\alpha_{\text{AMPA}} = 0.05$, $\alpha_{\text{NMDA}} = 0.0075$; $\tau_{1\text{GABA}} = 1 \text{ ms}$, $\tau_{1\text{AMPA}} = 0.5 \text{ ms}$, $\tau_{1\text{NMDA}} = 4 \text{ ms}$, $\tau_{2\text{GABA}} = 7 \text{ ms}$, $\tau_{2\text{AMPA}} = 2.4 \text{ ms}$, $\tau_{2\text{NMDA}} = 40 \text{ ms}$, $\tau_{\text{SRA}} = 200 \text{ ms}$; $g_{\text{input}} = 0.06 \text{ mS}/\text{cm}^2$. The following parameters were generated randomly for each neuron, from a Gaussian distribution with a 5% SD and the following mean: $g_{\text{NaP}} = 0.2 \text{ mS}/\text{cm}^2$, $g_{\text{KS}} = 8 \text{ mS}/\text{cm}^2$; $I_{\text{neuromodul}} = -1 \mu\text{A}/\text{cm}^2$.

Columnar Organization and Connectivity. Each thalamo-cortical column comprised 80 excitatory and 40 inhibitory neurons organized in a three-layered structure, schematizing supragranular, infragranular, and layer IV cortical neurons, and a corresponding thalamic sector (see Fig. 1C). Within each layer, there were 20 excitatory and 10 inhibitory neurons.

Connections were established with a 60% probability. Connection parameters (synaptic strength γ , unit = mS/cm^2 , and transmission latency $t_{\text{delay } i,j}$, unit = ms) were drawn from a Gaussian distribution with 10% SD around a fixed mean. All neurons connecting to given area randomly contacted excitatory and inhibitory neurons with the same probability and strength.

Inhibition. Inhibitory neurons sent only intralaminar horizontal connections both within a column ($\gamma_{\text{GABA}} = 0.12$, $\tau_{\text{delay}} = 2$) and toward other competing columns within the same area ($\gamma_{\text{GABA}} = 0.60$, $\tau_{\text{delay}} = 2$).

Intracolumnar Connectivity. Thalamic excitatory neurons projected to layer IV ($\gamma_{\text{AMPA}} = 0.20$, $\tau_{\text{delay}} = 3$) and, with lesser strength, to infragranular neurons ($\gamma_{\text{AMPA}} = 0.10$, $\tau_{\text{delay}} = 3$). Layer IV excitatory neurons projected to supragranular neurons ($\gamma_{\text{AMPA}} = 0.15$, $\tau_{\text{delay}} = 2$). Supragranular excitatory neurons projected to infragranular neurons ($\gamma_{\text{AMPA}} = 0.10$, $\tau_{\text{delay}} = 2$). Finally, infragranular excitatory neurons projected to layer 4 ($\gamma_{\text{AMPA}} = 0.05$, $\tau_{\text{delay}} = 7$), to supragranular neurons ($\gamma_{\text{AMPA}} = 0.05$, $\tau_{\text{delay}} = 7$), and to the thalamus ($\gamma_{\text{AMPA}} = 0.075$, $\tau_{\text{delay}} = 8$). Those principles and parameter values, inspired by ref. 21, capture the major properties of translaminar connections, although they do not attempt to capture the possible functional roles of the different layers (20). Latencies were adjusted to

take into account the coarseness of the present three-layer model and the fact that, in reality, some of these pathways are disynaptic (21).

Cortico-Cortical Projections. Supragranular excitatory neurons of each area projected to layer IV of the next area ($\gamma_{\text{AMPA}} = 0.05$, $\tau_{\text{delay}} = 3$). In agreement with physiological observations (19, 22), top-down connections were slower, more numerous, and more diffuse. They connected the supra- and infra-granular excitatory neurons of a given column to the supra- and infra-granular layers of all areas of a lower hierarchical level (Fig. 1B). Strong top-down connections linked columns coding for the same stimulus ($\gamma_{\text{NMDA}} = 0.05$), whereas weaker top-down connections projected to all columns of a lower area ($\gamma_{\text{NMDA}} = 0.025$). Transmission delays increased with cortical distance ($\tau_{\text{delay}} = 5 + 3 \delta$, with $\delta = 1$ for consecutive areas, 2 for areas two levels apart in the hierarchy, etc.).

Attentional Blink Experiment. The goal of this experiment was to measure the subjective perception of targets presented during the attentional blink. Ten French students participated after giving informed consent. On each trial, participants saw a rapid serial visual presentation of distractors, amidst which two targets T1 and T2 could appear. They were asked to rate T2 visibility, then to report T1 identity. Each rapid serial visual presentation element was presented for 43 ms, followed by a 43-ms blank. Distractors were four-letter uppercase consonant strings ($4^\circ \times 1^\circ$). Except for critical items preceding and following T1 and T2, each distractor had a 20% chance of being replaced by a blank. T1, which appeared as the 7th or 10th element in the RVSP sequence, was the letter string XOOX or OXXO. T2, which could appear with a lag of 1, 2, 3, 4, 6, or 8 stimuli after T1, was a French uppercase number word (DEUX, CINQ, SEPT, or HUIT) on one-half of the trials, and a blank screen on the other one-half (384 trials total). After T2, two more distractor strings were presented, and then participants were immediately asked to rate their perception of T2 by placing a cursor on a continuous horizontal scale labeled with “not seen” at left and with “maximal visibility” at right. They then reported the identity of the middle letters of T1, by pressing a left-hand button for O and a right-hand button for X. This procedure minimized the latency between the visual presentation of T2 and the report of its subjective visibility (215 ms), thus reducing the possibility that participants forgot having seen T2.

Results

We placed the network in a parameter regime in which it exhibited spontaneous thalamo-cortical oscillations, plausibly corresponding to a state of wakefulness (10). We then exposed it to the conditions of the attentional blink: a brief (40 ms) stimulation of thalamic neurons coding for T1, followed at a variable lag (0–400 ms) by a similar stimulation of T2 neurons (Fig. 2). Consider first the activation evoked by T1. A short burst of phasic stimulus activity was first propagated across the cortico-thalamic hierarchy. The highest cortical levels generated top-down amplification signals that, $\approx 80 \text{ ms}$ later, caused sustained firing in T1 neurons, accompanied by hyperpolarization of T2 neurons in areas C and D. In a larger-scale simulation, this long-lasting dynamic state would permit the brain-scale propagation of stimulus information about T1. We hypothesize that this global broadcasting constitutes the physiological correlate of conscious reportability.

Compared to this dynamic state reliably evoked by T1 presentation, the activation evoked by T2 depended tightly on the timing of T2 presentation. T2 targets presented simultaneously with T1, or long after, elicited sustained firing supported by joint bottom-up activation and top-down amplification (Fig. 2 *Left*). However, T2 targets presented during T1-elicited global firing

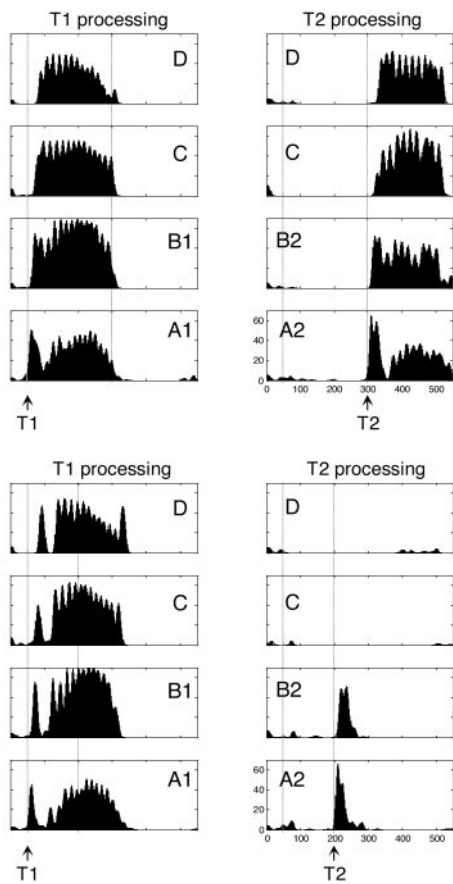


Fig. 2. Simulation of two trials of the attentional blink task. In each case, boxes show the temporal evolution of firing rate in excitatory cortical neurons within each of the eight columns of Fig. 1*B*. (*Upper*) “Seen” trial with long T1–T2 lag (250 ms). T2 is presented at a time when the sustained activity evoked by T1 has decayed. T2 is therefore able to invade all cortical levels and elicits long-lasting activity comparable to T1. (*Lower*) “Blinked” trial with short T1–T2 lag (150 ms). When T2 is presented during T1-evoked activity, it fails to evoke neural activity beyond an initial bottom-up activation in areas A and B.

elicited only bottom-up activation in assemblies A2 and B2 (Fig. 2 *Right*). T1-elicited inhibition prevented T2 activation from propagating to higher cortical levels. As a result, the second phase of top-down amplification did not occur.

By using as an index of T2 reportability the firing rate of pyramidal cells coding for T2 in areas C or D, our simulation reproduced the temporary drop in performance typical of the attentional blink (Fig. 3*A*). It also showed a global drop of power emitted in the γ -band (Fig. 3*C*) and of crosscorrelations between distant T2-coding neurons (data not shown). Thus, several indexes of firing and synchrony all pointed to a drop in global activity during the blink, particularly evident in the higher areas C and D.

As a test for the requirement of long-distance top-down connections, we ran simulations at a critical T1–T2 lag of 100 ms for which the blink normally occurs, but with the top-down connections eliminated, weakened by 50%, or reduced to nearest-neighbors. In all cases, T2 activation extended to areas C and D, and thus the blink did not occur (data not shown). This indicates that a key feature of the workspace model, the contribution of long-distance axons, is critical to the blink phenomenon.

An original property of the model is the prediction of a dynamic all-or-none bifurcation in neural activity. Across sim-

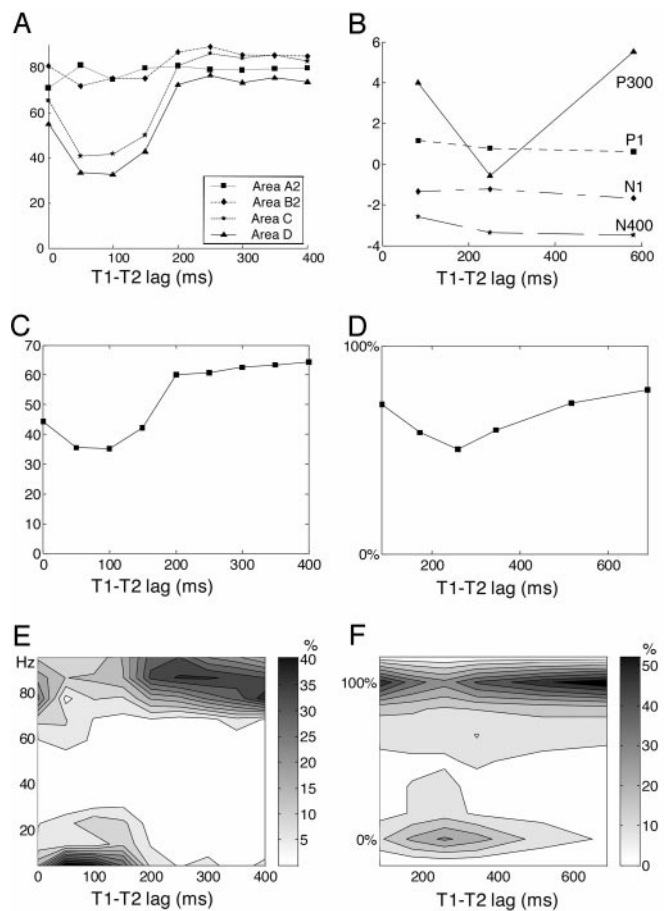


Fig. 3. Comparison of network simulations (*A*, *C*, and *E*) with experimental subjective and objective data on T2 processing (*B*, *D*, and *F*). (*A*) In the simulation, the peak firing rate of T2 neurons (shown in Hz) is unaffected by T1–T2 lag in low-level areas but shows a temporary drop at intermediate lags in higher areas C and D. (*B*) Likewise, experimentally recorded event-related potentials evoked by T2 (shown in μ V) exhibit a preservation of P1, N1, and N400 components but a drop of the P300 waveform during the attentional blink [redrawn from data by Vogel, *et al.* (12)]. (*C*) Mean power in the γ -band (20–100 Hz) emitted by simulated pyramidal neurons during a 200-ms window after T2 presentation. (*E*) Distribution of firing in area D. In the simulation, reverberant activation of workspace neurons is all-or-none such that parameters such as peak firing rate and γ -band power are distributed bimodally. Such a bimodal distribution was observed in an experiment in which we collected subjective ratings of T2 perception from 0% (not seen) to 100% (maximal visibility) (*D*, mean rating; *F*, distribution of ratings).

ulated trials, depending on random fluctuations in spontaneous activity before stimulus arrival, ascending activity could be sufficient to trigger self-amplifying recurrent activity, or it remained below threshold and only transient bottom-up activity was seen. Thus, for a fixed T1–T2 lag, simulated firing rates in higher areas and other indices of global activity (γ -band power and long-distance crosscorrelation) were found to be distributed bimodally across trials (Fig. 3*E*). The model therefore predicts that the apparent gradual drop in reportability observed during the attentional blink may be an artificial consequence of averaging across trials with full access awareness and others with no awareness.

To test this prediction experimentally, we used a modified attentional blink paradigm in which human subjects merely had to report to what extent they had seen a word (T2) within a rapid letter stream that contained another target letter string (T1). To obtain a continuous measure of subjective perception, we asked

subjects to move a cursor on a continuous scale, from “not seen” on the left to “maximal visibility” on the right. The results indicated that subjective perception during the blink is indeed all-or-none (Fig. 3*F*). At the peak of the blink, which occurred ≈ 260 ms after T1, T2 was fully perceived in approximately one-half of the trials, as indicated by a cursor position identical to “seen” trials with a long T1–T2 lag or without attention on T1. In the other one-half, T2 was totally unseen, as indicated by a cursor position identical to T2-absent trials. Participants almost never used intermediate cursor positions.

The same participants also took another set of trials with identical stimuli but with the instruction to ignore T1. We verified that the blink was much reduced in this condition, with T2 perception ratings clustering at the higher end of the scale. Two control experiments were also run to verify that participants could use the continuous rating scale (C.S. and S.D., unpublished data). In one, on each trial we presented a single T2 target for a variable duration (14, 29, 43, 57, 71, or 86 ms), followed by a single mask. In this situation, subjective ratings of T2 perception showed a unimodal distribution, the peak location of which varied continuously with T2 duration. This result indicated that subjects could use the response scale in a graded manner as a reliable indicator of subtle perceptual changes. We then combined the blink and masking experiments by presenting variable-duration T2 targets at lag 3 after T1 in the same sequence as above. The results showed a coexistence of two effects in the same trials: a continuous influence of T2 duration and an all-or-none effect of the blink. At each T2 duration, the response distribution was bimodal, with one peak at 0% visibility and the other peak at a location that increased progressively with T2 duration. Thus, even when participants show evidence of using the subjective rating scale as a graded measure of their perception, they still rate a significant fraction of blinked trials as totally unseen. These data suggest that the attentional blink in humans is associated with a stochastic all-or-none blocking of processing, which is accounted for in the model by an interaction between spontaneous activity and the nonlinear dynamics of bottom-up and top-down reverberation.

Discussion

Our model is coherent with previous proposals of a role of top-down recurrent (2), reentrant (28, 29), or resonant (10, 20) connections in the integrative processes underlying conscious perception. Nevertheless, it goes beyond those proposals by presenting a detailed simulation of a cognitive task, which tentatively links subjective reports with objective physiological correlates of consciousness on the basis of a neurally plausible architecture.

Despite the relative simplicity of the postulated neuronal architecture, simulation of the model captures the main physiological observations on the attentional blink. In the simulation, blinked T2 targets evoke unchanged bottom-up activity in areas A and B, but a much reduced activity in high-level areas C and D. This matches the empirically observed preservation of the P1, N1, and N400 components, but the drastic drop of the P300 component of event-related potentials evoked by T2 during the attentional blink (12). We hypothesize that the former components reflect the bottom-up progression of activation in a nested hierarchy of perceptual and semantic processors, whereas P300 reflects the sudden and global activation of workspace neurons.

Because the blink is attributed to competition for workspace access, the proportion of T2 targets that are blinked, as a function of time, roughly traces the inverse shape of the neural activity evoked by T1 in higher-level areas C and D. In actual experiments, similarly, there is an inverse relation between the P300 waveform evoked by T1 and the size of the blink (30). Furthermore, the functional MRI activation elicited by T1 in parietal, frontal, and cingulate areas predicts the size of the blink

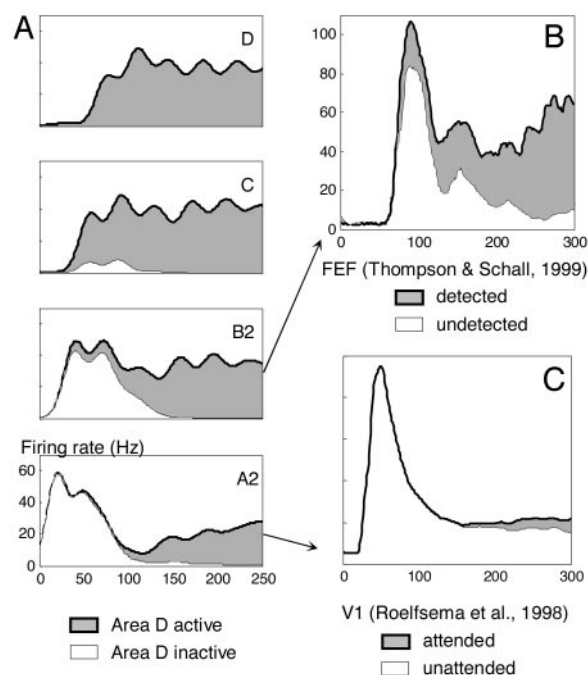


Fig. 4. Neural activity evoked by seen and unseen T2 targets in simulation (A) and actual experimental recordings (B and C). For lags between 0 and 200 ms, simulated trials were sorted as a function of the amount of pyramidal neuron activity that reached area D in a 250-ms window after T2 presentation. (A) Averages of firing rate as a function of time show that trials with area D activity are characterized by long-lasting amplification in lower areas C, B2, and A2. (B) The curves from assembly B2, showing earlier amplification and a small difference in the amplitude of the initial peak, are similar to activity evoked in monkey frontal eye field area during a masking paradigm [redrawn from data by Thompson and Schall (34)]. (C) The curves from assembly A2, with unchanged initial peak and small subsequent amplification, are similar to actual electrophysiological recordings from monkey area V1 during attention and conscious perception paradigms [redrawn from data by Roelfsema *et al.* (39)].

(31). Both observations are compatible with our hypothesis that the global neuronal workspace, formed by neurons distributed in higher areas including prefrontal cortex, operates functionally as a processing bottleneck that cannot simultaneously process two targets (13).

Although there have been no single-neuron recordings during the attentional blink, the simulated profiles of single-neuron activity to seen and blinked T2 targets can be compared to electrophysiological recordings obtained in other paradigms of conscious and unconscious processing. In perceptual areas A and B of the model, neurons fire phasically in tight synchrony with the stimulus, then show a broader period of late amplification only in seen trials, not in blinked trials. This parallels experimental recordings in areas V1 and IT under conditions of inattention, reduced contrast, masking, or anesthesia, where late amplification occurs only for reportable stimuli (1–3, 32) (Fig. 4C).

In the model, fluctuations in the state of the network before the stimulus, particularly in the spontaneous oscillations of membrane potential, create variability in the activity evoked by identical stimuli. When intrinsic fluctuations are in phase with stimulus presentation, bottom-up activation is enhanced. This coincidence has a cascading effect on subsequent areas and eventually affects the probability of the entire network falling into a global active state. Similarly, experimental recordings in V1 during the perception of ambiguous stimuli have shown that both late amplification and visibility vary in an all-or-none

manner (1) and can be partially predicted by the state of activity of V1 before stimulus arrival (33). This suggests that, as in our model, the resonance of incoming stimuli with spontaneous brain activity is essential for perception.

In our simulation, comparing seen and unseen trials also revealed a small difference in the amplitude of the first phasic peak in assembly B2 (Fig. 4). Likewise, experiments that sort trials based on the presence or absence of a subsequent conscious report occasionally show early differences in neural activity in V1 or frontal eye field (34, 35). In our interpretation, however, such early effects are only secondary correlates of consciousness because they are merely an indirect consequence of selective averaging over a fluctuating baseline. That those early differences are not associated with a conscious content is seen most clearly when they are found even before the stimulus is presented (33).

Obviously, our simulation is still simplified in many ways. First, in actual attentional blink experiments, backward masking of T2 is essential (36, 37). We suggest that masking is necessary for the blink to occur because it reduces T2-induced activation to a brief pulse (3, 32), which then acts as a “tracer” of the occupancy of the conscious workspace at a given moment. In our simulation, this could be achieved by directly stimulating T2-coding thalamic neurons with a temporally delimited pulse and thus no masking was necessary. Another difference is that, in actual experiments, the presence of the attentional blink depends on the T1 task, and disappears when subjects are instructed to ignore T1. For simplicity, our network had a fixed connectivity set to process both T1 and T2. Neural mechanisms of task switching have been proposed, however, and could be combined with the present work (14). Finally, although the present model reproduces the qualitative features of the attentional blink, there are quantitative differences in timing, particularly in the onset and duration of the blink (Fig. 2). This is due in part to the value of the propagation delays, which were extrapolated from monkey visual areas and should be slower for the more distant areas of the human brain. Blink duration, which corresponds to the duration of T1-induced activity, is jointly determined in our simulation by the strength of excitatory connections and by the amount of spike-rate adaptation. In a more realistic model, it should also

vary with the task, with additional circuits providing task-dependent reverberation, possibly associated with a subjective feeling of conscious effort (14).

According to the present model, the primary correlate of conscious access is a sudden self-amplifying bifurcation leading to a global brain-scale pattern of activity. This leads to several critical predictions for neuroimaging and pharmacological experiments. First, in experiments with a continuously varying and increasingly visible stimulus, we predict a sudden nonlinear transition toward a state of globally increased brain activity (“ignition”), particularly evident in prefrontal, cingulate, and parietal cortices, accompanied by a synchronous amplification of posterior perceptual activation and by thalamo-cortical brain-scale synchrony in the γ range (20–100 Hz). Finer scale electrophysiological or optical recordings should demonstrate that this state is selective to a subset of neurons coding for the stimulus and is accompanied in higher areas by broad inhibition of neurons coding for other stimuli. Even higher-resolution experiments should reveal a layer-specific distribution of active workspace neurons, with intense top-down activity originating from a subpopulation of cells with long axons in supra- and infra-granular layers of prefrontal, cingulate and parietal cortices. Second, we predict a bimodal distribution of activation during individual attentional blink trials, which could be demonstrated by single-event functional MRI or by distributional analysis of intracranial or even scalp measurements of the P300. Third, experimental interventions that affect top-down connections or inhibitory interactions in higher cortical areas should alter this dynamic transition, and therefore subjective perception, without necessarily affecting subliminal bottom-up processing. This prediction could be tested by using pharmacological agents such as for instance general anesthetics known to affect NMDA or GABA receptors, or effectors of cholinergic receptors that may globally gate the access to the global workspace (38).

We thank V. Lamme and Y. Rossetti for helpful comments. This work was supported by Institut National de la Santé et de la Recherche Médicale, Commissariat à l’Energie Atomique, the McDonnell Foundation, Centre National de la Recherche Scientifique, Collège de France, the French Ministry of Research, the Association pour la Recherche sur le Cancer, the Association Française contre les Myopathies, and the Commission of the European Communities.

- Super, H., Spekreijse, H. & Lamme, V. A. (2001) *Nat. Neurosci.* **4**, 304–310.
- Lamme, V. A. & Roelfsema, P. R. (2000) *Trends Neurosci.* **23**, 571–579.
- Lamme, V. A., Zipser, K. & Spekreijse, H. (2002) *J. Cognit. Neurosci.* **14**, 1044–1053.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B. & Rivière, D. (2001) *Nat. Neurosci.* **4**, 752–758.
- Moutoussis, K. & Zeki, S. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 9527–9532.
- Engel, A. K. & Singer, W. (2001) *Trends Cognit. Sci.* **5**, 16–25.
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B. & Varela, F. J. (1999) *Nature* **397**, 430–433.
- Lumer, E. D. & Rees, G. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 1669–1673.
- Kleinschmidt, A., Buchel, C., Hutton, C., Friston, K. J. & Frackowiak, R. S. (2002) *Neuron* **34**, 659–666.
- Llinas, R., Ribary, U., Contreras, D. & Pedroarena, C. (1998) *Philos. Trans. R. Soc. London B* **353**, 1841–1849.
- Tallon-Baudry, C. & Bertrand, O. (1999) *Trends Cognit. Sci.* **3**, 151–162.
- Vogel, E. K., Luck, S. J. & Shapiro, K. L. J. (1998) *J. Exp. Psychol. Hum. Percept. Perform.* **24**, 1656–1674.
- Chun, M. M. & Potter, M. C. (1995) *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 109–127.
- Dehaene, S., Kerszberg, M. & Changeux, J. P. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 14529–14534.
- Dehaene, S. & Naccache, L. (2001) *Cognition* **79**, 1–37.
- Baars, B. J. (1989) *A Cognitive Theory of Consciousness* (Cambridge Univ. Press, Cambridge, U.K.).
- Raymond, J. E., Shapiro, K. L. & Arnell, K. M. (1992) *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 849–860.
- Wang, X.-J. (1993) *NeuroReport* **5**, 221–224.
- Felleman, D. J. & Van Essen, D. C. (1991) *Cereb. Cortex* **1**, 1–47.
- Raizada, R. D. & Grossberg, S. (2003) *Cereb. Cortex* **13**, 100–113.
- Lumer, E. D., Edelman, G. M. & Tononi, G. (1997) *Cereb. Cortex* **7**, 207–227.
- Salin, P. A. & Bullier, J. (1995) *Physiol. Rev.* **75**, 107–154.
- Arnell, K. M. & Jolicoeur, P. (1999) *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 630–648.
- Dayan, P. & Abbott, L. F. (2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, MA).
- Steriade, M., McCormick, D. A. & Sejnowski, T. J. (1993) *Science* **262**, 679–685.
- Llinas, R. R. & Paré, D. (1991) *Neuroscience* **44**, 521–535.
- Mack, A. & Rock, I. (1998) *Inattentional Blindness* (MIT Press, Cambridge, MA).
- Di Lollo, V., Enns, J. T. & Rensink, R. A. (2000) *J. Exp. Psychol. Gen.* **129**, 481–507.
- Edelman, G. M. (1993) *Neuron* **10**, 115–125.
- McArthur, G., Budd, T. & Michie, P. (1999) *NeuroReport* **10**, 3691–3695.
- Marois, R., Chun, M. M. & Gore, J. C. (2000) *Neuron* **28**, 299–308.
- Kovacs, G., Vogels, R. & Orban, G. A. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 5587–5591.
- Super, H., van der Togt, C., Spekreijse, H. & Lamme, V. A. (2003) *J. Neurosci.* **23**, 3407–3414.
- Thompson, K. G. & Schall, J. D. (1999) *Nat. Neurosci.* **2**, 283–288.
- Pins, D. & Ffytche, D. (2003) *Cereb. Cortex* **13**, 461–474.
- Brehaut, J. C., Enns, J. T. & Di Lollo, V. (1999) *Percept. Psychophys.* **61**, 1436–1448.
- McLaughlin, E. N., Shore, D. I. & Klein, R. M. (2001) *Q. J. Exp. Psychol. A* **54**, 169–196.
- Granon, S., Faure, P. & Changeux, J. P. (2003) *Proc. Natl. Acad. Sci. USA* **100**, in press.
- Roelfsema, P. R., Lamme, V. A. & Spekreijse, H. (1998) *Nature* **395**, 376–381.