

Spatial structure affects temporal judgments: Evidence for a synchrony binding code

Samuel Cheadle	School of Psychology, Birkbeck, University of London, London, UK	
Frank Bauer	Department of Psychology, Ludwig-Maximilian University, Munich, Germany	
Andrew Parton	Centre for Cognition and Neuroimaging, Brunel University, Uxbridge, UK	 
Hermann Müller	Department of Psychology, Ludwig-Maximilian University, Munich, Germany	 
Yoram S. Bonneh	Weizmann Institute, Rehovot, Israel	 
Marius Usher	School of Psychology, Birkbeck, University of London, London, UK	 

Visual binding is the process by which the brain groups the elements belonging to one object, whilst segregating them from other scene elements. A computationally parsimonious mechanism of visual binding is the binding-by-synchrony (BBS) hypothesis. According to this hypothesis, detectors that respond to elements of a single object fire in synchrony, while detectors that respond to elements of different objects do not. Current psychophysical and physiological evidence are inconclusive about the role of BBS in the visual integration process. Here we provide psychophysical and computational evidence suggesting that the visual system implements a mechanism that synchronizes response onsets to object parts and attenuates or cancels their latency differences. In three experiments, observers had to judge the synchrony of two flickering Gabor patches embedded in a static Gabor contour, passing through fixation. We found that a smooth contour, as compared to a jagged one, impedes judgments of temporal synchrony between the targets, whilst facilitating non-temporal judgments of contrast on the same targets. We show that the results are consistent with a simple computational model that implements synchronization of responses via lateral interactions, enabling the visual system to pick up objects by synchrony among a temporally diverse background.

Keywords: visual-grouping, figure-ground, binding-problem, synchrony, temporal-judgment, computational-model

Citation: Cheadle, S., Bauer, F., Parton, A., Müller, H., Bonneh, Y. S., & Usher, M. (2008). Spatial structure affects temporal judgments: Evidence for a synchrony binding code. *Journal of Vision*, 8(7):12, 1–12, <http://journalofvision.org/8/7/12/>, doi:10.1167/8.7.12.

Introduction

The capacity of higher organisms to rapidly interpret visual scenes depends on their ability to solve a computationally demanding task known as the *binding problem*: grouping together elements of one object, whilst segregating them from other scene elements. The binding problem arises because of the need to integrate multiple sources of local information in a flexible way. The challenge posed by the binding problem can be illustrated (Figure 1, reproduced from Gray, 1999) by examining how the same 6 local objects (vertical line segments assumed to activate corresponding feature detectors with non-overlapping receptive fields depicted by the ellipses) can be combined

in various ways to create stimulus configurations. One option to compute these configurations on the basis of local information is to assume the existence of more complex, conjunction detectors (e.g., ‘grandmother cells;’ Barlow, 1972) corresponding to all the configurations depicted. This solution *alone*, however, runs into the severe problem of a combinatorial explosion: the number of combinations of local features increases exponentially with the number of features.

Consider the simplest binding situation where a single figure must be segregated from background. Physiological and neurocomputational studies have suggested a partial solution to figure-ground grouping for this case: the activity, or saliency, of neurons responding to visual elements that form good Gestalts is enhanced relative to

the background (Kovács, 1996). This computational scheme has been widely studied both psychophysically and physiologically, with accumulating evidence for contour integration via facilitatory lateral interactions, presumably in the primary visual cortex (Cass & Alais, 2006; Gilbert, 1993; Kourtzi, Tolias, Altman, Augath, & Logothetis, 2003; Kovács, 1996; Li, Piëch, & Gilbert, 2006; Polat & Sagi, 1993; Usher, Bonneh, Sagi, & Herrmann, 1999). Such a saliency scheme, which enhances the activation of all parts of an object, has two limitations. First, it is likely to interfere with coding of the local properties of the stimulus such as contrast, via modulation of the strength of neuronal response, which stands in contrast to experimental results showing no modulation of perceived contrast, as well as no difficulty in handling variable contrasts in contour integration paradigms (Hess, Dakin, & Field, 1998). Second, the salience enhancement scheme does not generalize to the computation of grouping in situations where multiple objects are perceived simultaneously, as illustrated in Figure 1.

A more general and efficient (though controversial; see special issue of *Neuron*, 1999, Sept.) solution to the binding problem is the ‘binding-by-synchrony’ (BBS) hypothesis, according to which detectors that respond to elements of a single object fire in synchrony, while detectors responding to elements of separate objects do not (Eckhorn et al., 1988; Milner, 1974; Singer & Gray, 1995; von der Malsburg, 1995). This synchronization, which presumably implements grouping according to such Gestalt properties as proximity, continuity, and smoothness, is thought to be mediated by lateral and feedback connections within the visual cortex (Gilbert, 1993) and to operate with a temporal resolution smaller than that of the psychological ‘present’ (von der Malsburg, 1999). Animal neurophysiological and human electrophysiological studies provide

some evidence of synchronized activity as predicted by the BBS hypothesis (e.g., Eckhorn, Frien, Bauer, Woelber, & Kehr, 1993; Fries, Roelfsema, Engel, König, & Singer, 1997; Gray, König, Engel, & Singer, 1989; Singer & Gray, 1995; Vidal, Chaumon, O’Regan, & Tallon-Baudry, 2006), but some researchers have questioned whether this synchronization signal is robust enough to encode binding (e.g., Palanca & DeAngelis, 2005).

Previous psychophysical studies testing the BBS hypothesis have mainly used paradigms that examined effects of presenting temporally structured visual elements in a grouping task (e.g., Fahle, 1993). Whilst some of these studies showed that fast temporal synchrony (>30 Hz in-phase flicker) of visual elements, which belong to a contour, facilitate its grouping and segregation from background elements (Clifford, Holcombe, & Pearson, 2004; Elliott & Müller, 1998; Lee & Blake, 1999; Leonards, Singer, & Fahle, 1996; Parton, Donner, Donnelly, & Usher, 2006; Usher & Donnelly, 1998), other studies obtained negative results (Fahle & Koch, 1995; Kiper, Gegenfurtner, & Movshon, 1996). Furthermore, alternative interpretations of these findings are that common temporal properties provide a fundamental grouping cue (Lee & Blake, 1999), or that grouping is performed on the basis of stimulus transients or other artifacts (Dakin & Bex, 2002; Farid & Adelson, 2001).

To help resolve this ambiguity, we have devised a novel approach for testing the BBS hypothesis, namely by examining the influence of *spatial structure* on *temporal discrimination*. Specifically, we asked: does task-irrelevant spatial structure modulate the perceived temporal structure? According to the BBS hypothesis, coherent spatial structure generates an internally synchronizing coupling between neural detectors that respond to different elements of this structure. Consider a task of identifying the phase relationship (synchronous or asynchronous) of two flickering “targets”, joined by either coherent or non-coherent spatial structure. As we formally show in the next section, the synchronous activity (triggered by the coherent structure) will interfere with the temporal judgment signal, resulting in a performance cost for targets linked by coherent spatial structure. This will be shown to stand in contrast to its facilitatory effects on *non-temporal* judgments of the same visual elements (Duncan, 1984; Lavie & Driver, 1996; Polat & Sagi, 1993). We start by describing a generic computational model, which implements the BBS scheme, for predicting the effects of spatial grouping on temporal judgment. We then present the results of three experiments, testing these predictions.

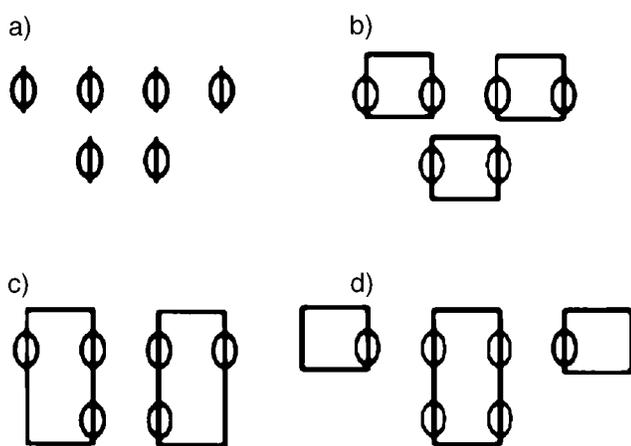


Figure 1. Local elements (line segments) can be combined in multiple ways in complex contour shapes (from Gray, 1999; reproduced from *Neuron*, permission pending). Reprinted from *Neuron*, 24, Gray, The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well. Copyright (1999), with permission from Elsevier.

Computational model

Rationale

The BBS hypothesis predicts that coherent spatial structure formed by the alignment of Gabor patches generates, via lateral connections, a synchronizing cou-

pling between neural detectors that respond to each patch (including the target patches). Even though the spatial structure is irrelevant to a temporal judgment of the targets' flicker, the temporal coupling of their activity interferes with neuronal responses induced by the flicker. This predicts a decrease in the synchrony/asynchrony discrimination threshold for 'strong' compared to 'weak' grouping conditions. To demonstrate this assertion in a formal way, we developed a simple generic neural model with which we could examine the effect of spatial coherence on temporal judgments.

Methods

We used a canonical neural model of cell assemblies (Wilson & Cowan, 1972), where each target detector corresponds to an excitatory–inhibitory assembly. To simplify, we only model the cell assemblies whose orientation and spatial–frequency preference matches that of the target elements, and we assume that their coupling depends on the smoothness of the visual elements that interpolate between them: the coupling will be zero for 'weak' grouping but greater than zero for 'strong' grouping. The activations of the cell assemblies were updated in simulation time steps according to the equations:

$$\begin{aligned} x_1(t+1) &= k x_1(t) + (1-k)[A_1 F(x_1) - B F(y_1) + A_2 F(x_2) \\ &\quad + I_1 + \text{noise}](t); \\ y_1(t+1) &= k y_1 + (1-k)[C F(x_1) - D F(y_1)](t) \\ x_2(t+1) &= k x_2(t) + (1-k)[A_1 F(x_2) - B F(y_2) + A_2 F(x_1) \\ &\quad + I_2 + \text{noise}](t); \\ y_2(t+1) &= k y_2 + (1-k)[C F(x_2) - D F(y_2)](t), \end{aligned} \quad (1)$$

where x_1 , x_2 are the activations of the target detectors (blue/red in Figure 2b) and y_1 , y_2 the activations of their inhibitory cells assembly components; $k = .98$ is a leaky integration constant; F is the firing rate nonlinearity, chosen as $F(x) = x/(1+x)$ for $x > 0$, and 0 for $x < 0$ (O'Reilly & Munakata, 2001; Usher & Cohen, 1999); and noise is a Gaussian term with zero mean and standard deviation SD_1 . A_2 is the coupling between the detectors, set to 0 for the non-smooth contour condition (no coupling) and to .3 (an intermediate magnitude, lower than the self-coupling coefficient, $A_1 = 1$) for the smooth condition. A third cell assembly, which receives no input, was also included in the simulation as a measure of baseline activation (green curve in Figure 2b).

The input to the two target detectors are step functions, which turn on/off every 200 simulation time steps, corresponding to flickering stimuli. Depending on the temporal asynchrony conditions, I_2 is time shifted relative to I_1 by 0%, 25%, 50%, 75%, or 100% of 200 time steps (the maximum shift is completely out-of-phase). The

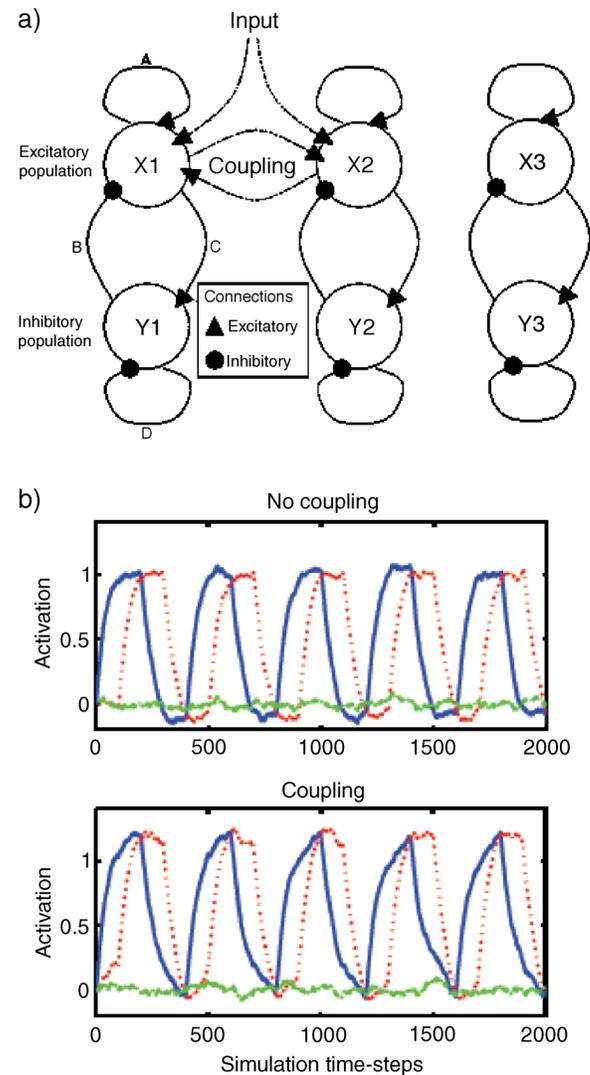


Figure 2. (a) Schematic diagram of the model with excitatory cell assemblies for targets, X1 and X2, interacting via the coupling connection. X3 (non-target unit) receives no input and has no coupling connection. Connections: A = self-coupling connection (excitatory); B = inhibitory feedback from Y to X; C = excitatory connection from X to Y; D = self-coupling connection (inhibitory). See method section for parameter values used in the model. (b) Activation profiles of the target detectors (X1—blue solid line, and X2—red dashed line), and of one detector that is not receiving input as baseline (X3—green dash-dot), in response to flickering stimuli (every 200 time steps); the time shift between the two detectors corresponds to the $1/2$ overlap condition (initially the 'blue' detector receives input from 0 to 200, and the 'red' detector from 100 to 300, etc). Top panel: no coupling. Bottom panel: with coupling; the two units' activation profiles have greater overlap, as each of them supports the activation of the other.

activation profiles of the detectors (x_1 , x_2) for the strong/weak grouping conditions (top/bottom) are illustrated in Figure 2b at a time shift of 50%. The figure shows that the coupling (strong grouping) enhanced the overlap of the activation profiles (diminishing their separation).

For each overlap condition (0%, 25%, 50%, 75%, and 100%), the activation profiles (1000 time steps) with/without coupling were used to compute a separation measure, $\int |F(x_1) - F(x_2)| dt / \int |F(x_1) + F(x_2)| dt$. This separation (which quantifies the amount of asynchrony between the two targets' responses) was then compared (subject to noise) against a criterion to decide whether the units' activations were synchronous (in-phase) or asynchronous (out-of-phase). Specifically, if the separation measure 'plus' Gaussian noise (SD_2) was larger than a criterion, the stimulus was categorized as asynchronous, and otherwise as synchronous. The response criteria were chosen such as to fit the human observers' False Alarm (FA) rates (in-phase stimuli categorized as asynchronous; those differed between the grouping conditions). 1000 trials were run for each phase difference, and signal-detection d' (which measures bias-independent discrimination sensitivity) was computed from the Hits (asynchronous stimuli correctly categorized) and FAs (synchronous stimuli categorized as asynchronous). The model parameters used were: $A_1 = 1$, $B = 1.75$, $C = .9$, $D = .1$, $SD_1 = 1$, $SD_2 = .25$, $I = .33$. See Appendix for parameters that result in synchronisation of internally generated oscillations.

Model results

The results of the simulations formally demonstrate that the presence of coupling, associated with the presence of coherent spatial structure connecting the two targets, leads to decreased discrimination performance (Figure 3). Moreover, as one can see, the difference in temporal discrimination between the weak and strong grouping

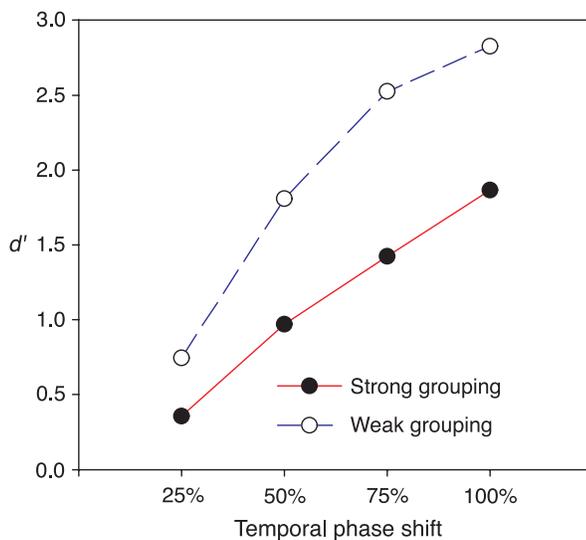


Figure 3. Pattern of data predicted by a canonical neural model. The graph depicts sensitivity (d') for 'strong' (red solid line) and 'weak' (blue dashed line) grouping for each of four incremental shifts in the relative temporal phase of the targets.

conditions reaches its maximum value (d' difference = 1.1) at an intermediate phase shift (75%), rather than at the maximal (out-of-phase) shift (100%; d' difference = .9).

Thus the model predicts a slightly larger weak-coupling superiority effect in temporal judgments at intermediate phases.

Experiment 1: The effect of visual grouping on temporal simultaneity judgments

Method

Observers

Ten observers (all naive) with normal or corrected to normal vision participated in Experiment 1.

Apparatus

Stimuli were generated with a Microsoft Windows system using a VSG 2/5 graphics board (Cambridge Research Systems) and displayed on a gamma-corrected Sony Trinitron Multiscan E450 monitor. For all experiments, the frame rate of the display was set at 140 Hz, the screen resolution at 800×600 pixels, and the background luminance at 28 cd/m^2 . Observers were tested in a dimly illuminated room, and were required to use a chin rest to minimize head movements and maintain a constant viewing distance of 57 cm. Responses were recorded using the left and right buttons of a CT3 four-button response box (Cambridge Research Systems).

Stimuli

The stimulus (see Figures 3a and 3b) consisted of a fragment of a curved contour composed of 7 odd-symmetric Gabor patches, with a spatial frequency of 3 cpd, window size of 1.2 deg, deviation of 0.18 deg, and an inter-element spacing of 1.2 deg (3.5 wave lengths). The construction and placement of the contour varied randomly across trials, with the constraints that the middle element always appeared at the center of the screen and the curvature between the elements had a constant value between 15 and 20 deg. The patches at the two ends of the contour, which served as targets, were always aligned to the contour path, while the other patches served as a connecting fragment, forming either a smooth (strong grouping condition, Figure 4a), or jagged (weak grouping condition, Figure 4b) path. A third condition, in which no connecting elements were presented between the flickering targets, was included as a baseline. The rationale behind positioning the target elements at the two ends of the contour was that we wanted to minimize any possible

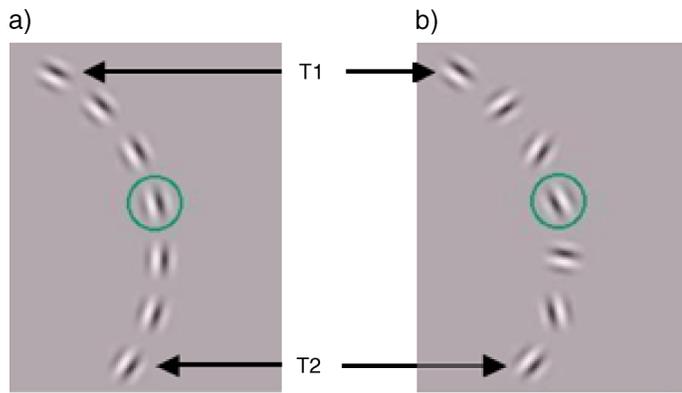


Figure 4. Typical examples of the test stimuli. The green circle indicates the fixation location. The targets were always the two end elements (T1 and T2). (a) Example contour in the strong-grouping condition. (b) Example contour in the weak-grouping condition.

role of motion signals in the temporal discrimination task. The connecting contour was static, while the target patches flickered on/off at a rate of 4.49 Hz, either in-phase, or out-of-phase at four different phase shifts.

Procedure

The first experiment employed a “yes/no” design: on each trial, observers had to make a simultaneity judgment, that is, report whether the two targets were in or out of phase. There were equal numbers of synchronous (no phase shift) and asynchronous trials, with the latter divided into 4 levels of phase shift (25%, 50%, 75%, and 100% out-of-phase). The stimulus display remained on until the observer responded, and there was a 1 second inter-trial interval. Each observer completed 180 trials for each of the three types of connecting contour (smooth and jagged, and invisible), which were presented in random order.

Results and discussion

The results of [Experiment 1](#) are shown in [Figure 5](#). We focus first on the comparison between the strong/weak grouping conditions (which are equated for the number of connecting elements, and thus free from potential confounds). The d' measure of temporal discrimination for judging simultaneity is plotted as a function of the temporal phase difference between the two flickering targets and for the two types of static contour (strong/weak grouping) connecting them.

As expected, performance improved with increased phase difference. The main result is that (for phase shifts $\geq 50\%$) the type of the static connecting contour, which was irrelevant to the task, makes a marked difference to performance: it was much easier to make the simultaneity judgment when the contour was jagged, as compared to

when it was smooth. In other words, a smooth contour interferes with the temporal judgment. Moreover, the results reflect the pattern predicted by the model ([Figure 3](#)). This was confirmed by examining these results by means of a two-way ANOVA, with grouping strength and phase shift as factors. As predicted by the model and the BBS hypothesis, there was a highly significant difference in task performance between ‘strong’ and ‘weak’ grouping conditions ($F(1,9) = 45.6, p < 0.001$). The ANOVA also revealed a significant interaction between grouping strength \times phase shift ($F(3,27) = 3.67, p = .024$), with the effect peaking at the 75% phase separation ($t(9) = 5.65, p < .001$), and falling away at the 50% ($t(9) = 2.24, p = .052$) and 100% ($t(9) = 2.22, p = .054$) separations. In the baseline condition (no connecting elements) the temporal judgment performance was similar to the weak grouping condition ($d' = 1.79$ for baseline, compared to $d' = 1.67$ for the weak grouping condition). A 4×2 ANOVA (baseline/weak-grouping \times phase) revealed no significant main effect of condition or interaction between condition and phase. Thus it appears that the mere presence of irrelevant connecting elements (unlike their spatial coherency) has little effect on temporal judgment performance.

The result of [Experiment 1](#) confirmed the model predictions, by showing the predicted effect of grouping strength on temporal discrimination performance. However, the method used (yes/no judgment) requires the observers to compare the stimulus against a response criterion, and this criterion may change (as indicated by the data) between the strong/weak-grouping conditions. Although we tried to remove such criterion effects by analyzing the results in terms of signal detection performance d' , we decided to examine whether the effect would

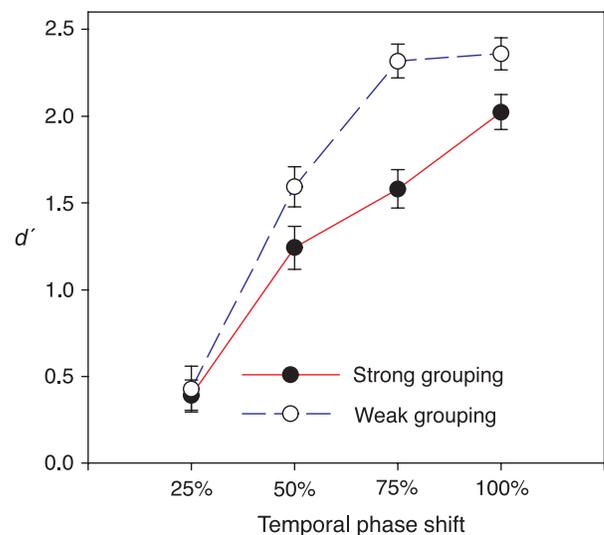


Figure 5. Sensitivity (d') plotted for strong/weak-grouping conditions, at four different temporal phase shifts. Error bars denote 1 SE (standard error) of the phase shift \times grouping strength condition means, adjusted using the method by Cousineau (2005) for within subjects designs.

also be evident when using a 2-interval forced-choice method (2IFC) which is supposed to be totally free of response criteria. To examine this, we ran a new group of 6 observers on the same task, though using a 2IFC method. On each trial, the observer was presented with a sequence of two stimuli (each presented for 1 sec), one synchronous and one asynchronous, and they then had to judge which of them (the first or the second) was synchronous. Two temporal phases were used for asynchronous stimuli, which correspond to the middle two conditions (50% and, respectively, 75% phase difference) in Figure 5, for which the grouping effect was largest. The phase conditions were blocked and the grouping conditions (strong/weak) were randomized within a block. The results confirmed the weak-grouping superiority effect for the 50% phase-shift condition (73% vs. 68%; $t(5) = .8.70, p < .001$), though not the 75% phase-shift condition (81% vs. 80%; $t(5) = .23, p = .83$). Although it is possible that this (overall) diminished effect was due to the variability introduced by the need to maintain the degree of temporal overlap of the first stimulus in memory while evaluating the second stimuli (indeed, observers experienced the task as more effortful), we decided to corroborate these findings in a further experiment (Experiment 2) designed to explore conditions in which the effect was expected to be maximized, so as to show up consistently even with the 2IFC task. To do so, Experiment 2 introduced two significant alterations to the stimuli (see Figure 6). First, the curved contours were replaced by straight ones; we reasoned that, if the effect is reliant upon collinear elements forming an object, they would exert the largest influence when the contour is straight. Second, the grouping strength was further decreased in the ‘weak’ condition, by surrounding the contour (in both grouping conditions) with background elements. In Experiment 1, the visual system may have grouped the elements to some degree also in the weak-grouping condition, simply because they were the only elements in the display and by this formed a figural object through the proximity of the elements.

Experiment 2: Judgments of synchrony using 2IFC

In Experiment 2, we tested the effect of grouping on the temporal simultaneity judgment employing a temporal 2IFC method, which is assumed to be criterion-free. Since the randomly positioned curved contours used in Experiment 1 were found to be rather difficult for observers to compare across time, Experiment 2 introduced contours arranged along a horizontal line embedded in “noise” (Figure 6). Using different stimuli also permitted us to assess the generality of the configuration effects.

Method

Observers

Four observers (2 naive) took part in Experiment 2.

Stimuli

As in Experiment 1, two flickering Gabor patches (with identical parameters as in Experiment 1) were presented either synchronously or asynchronously, and either strongly or weakly grouped (Figure 6). In contrast to Experiment 1, the following modifications were made: 1) Instead of a curved contour, a straight contour of 9 elements was presented; two additional elements were added so that the flickering targets no longer appeared at the extremities of the contour, but were embedded within it to further strengthen the grouping. 2) The contour was embedded in a 9×9 array of background elements (Gabor patches, with the same parameters as the targets); this had the effect of further reducing the grouping strength when the connecting elements were not aligned, and therefore maximizing any difference between the grouping conditions. 3) To tackle possible ceiling effects in the easier conditions, the frequency at which the targets flickered was increased to 6 Hz (in all conditions).

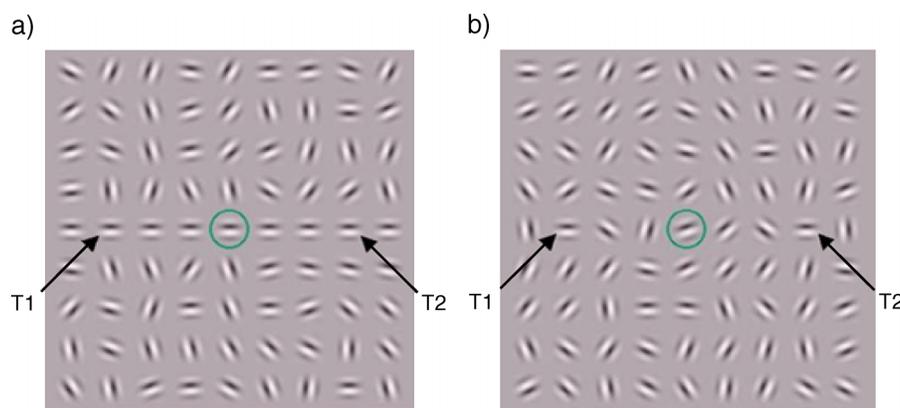


Figure 6. Example stimuli used in Experiment 2. Targets T1 and T2 flickered either in phase or out of phase. (a) Example stimulus for the strong-grouping condition. (b) Example stimulus for the weak-grouping condition.

Procedure

The second experiment employed a 2-interval forced-choice (2IFC) task. Each trial consisted of two presentations: one synchronous and one asynchronous. The observer's task was to correctly identify the order in which they appeared. Each presentation lasted for 1 second, and they were separated by 500 ms. Each observer completed 240 trials for each of the two types of connecting contour (smooth and jagged), which were blocked by phase separation, and presented in an ascending/descending sequence.

Results and discussion

Individual observers' as well as average 2IFC performance is presented in Figure 7. For all 4 observers, the performance was higher in the non-grouped compared with the grouped condition for the 2 intermediate phase shifts. This is reflected in the average data: an ANOVA revealed the main effect of grouping strength to be significant, $F(1,3) = 26.9$, $p = .014$, with maximal effects for the 50% and 75% phase separations (7% difference in both).

An interesting observation, supported by post-hoc analysis, was that the order of presentation in the 2IFC task (syn–async vs. asyn–sync) significantly influenced performance ($t(3) = 3.54$, $p = .039$): there was a systematic bias in accuracy towards asyn–sync presentations (75.9% correct) as compared to sync–async presentations (67.9% correct), that is, observers tended to decide that the second stimulus was the synchronous one. This effect shows no interaction with grouping type, and we confirmed that it was not a simple response bias (by switching the temporal structure defined as the target, from “respond which interval is synchronous” to respond “which interval is asynchronous”, for half the observers). We speculate that this effect may be attributable to a difference in attention (increased for the second interval due to cueing by the stimulus in the first interval)—which reduces temporal resolution (Yeshurun & Levy, 2003), resulting in a more ‘synchronous’ estimate for the second interval.

The results of Experiment 2 reinforce the previous results of Experiment 1, by showing that the negative effect of spatial structure on temporal judgments is not the outcome of a change in response criterion, but instead reflects a true change in temporal sensitivity. Although the findings of the two experiments are consistent with the

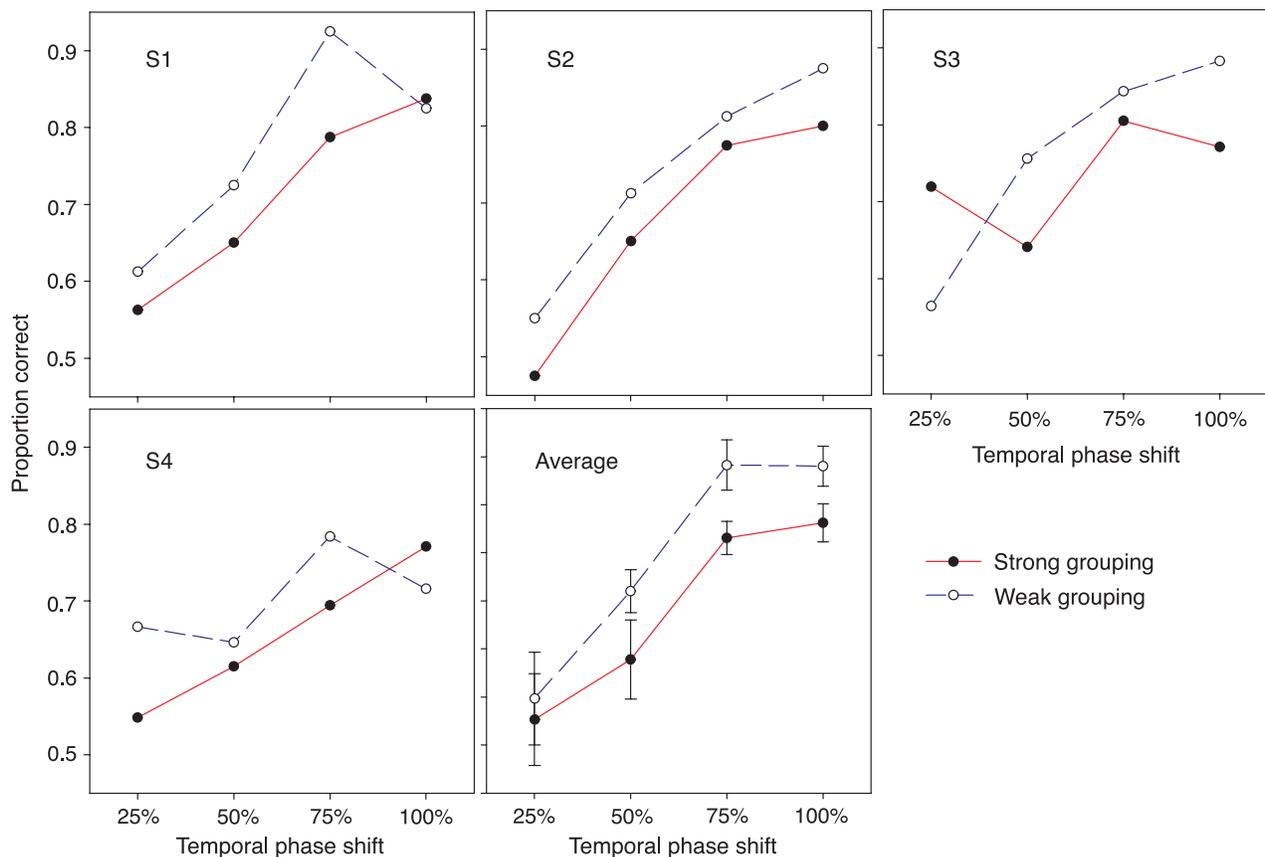


Figure 7. Percentage correct scores for both grouping conditions, at four different temporal phase shifts. A consistent advantage for the weakly grouped object is evident for the majority of separations.

BBS hypothesis, it remains a possibility that the interference effect is not specific to synchronous temporal judgments. Rather, the interference may be caused by the need to segment the targets from the rest of the contour—so it would affect judgments on *any* perceptual property of the two targets. In order to test whether the interference effect is specific to judgments of temporal synchrony, we carried out a third experiment in which 10 observers were tested on two matched tasks (temporal and non-temporal) on the same target elements. Previous studies have shown that judgments of non-temporal attributes can, in fact, be enhanced by grouping of the targets (Duncan, 1984; Lavie & Driver, 1996; Polat & Sagi, 1993). We thus predict that while strong spatial grouping diminishes performance on a temporal-judgment task, it facilitates performance for judgments of non-temporal attributes.

Experiment 3: Comparing temporal with non-temporal judgments

In [Experiment 3](#), we tested the effect of grouping on the non-temporal judgment of contrast discrimination, and compared it with a temporal judgment on the same stimuli. In order to take a conservative approach for this comparison, we used only the 100% phase separation for asynchronous trials, in the temporal-judgment condition.

Method

Observers

Ten observers (8 naive) took part in the third experiment.

Stimuli and procedure

Subjects completed both a temporal and a contrast discrimination task. Stimuli and conditions for the temporal judgment task were identical to those used in [Experiment 1](#), except that asynchronous trials always maintained 100% phase separation (anti-phase), rather than the variable separations used in [Experiment 1](#). For the contrast discrimination task, the spatial-configuration conditions remained the same, while the attribute used for the target comparison differed. Observers had to judge whether the two end elements were the same or different in contrast. A contrast increment appeared in one of the target patches in half of the trials at random. The contrast increment and temporal frequency were selected individually for each observer in a preliminary experiment that determined increment thresholds using an adaptive staircase procedure converging on ~70% correct discriminations. Each observer completed 240 trials for both the contrast and

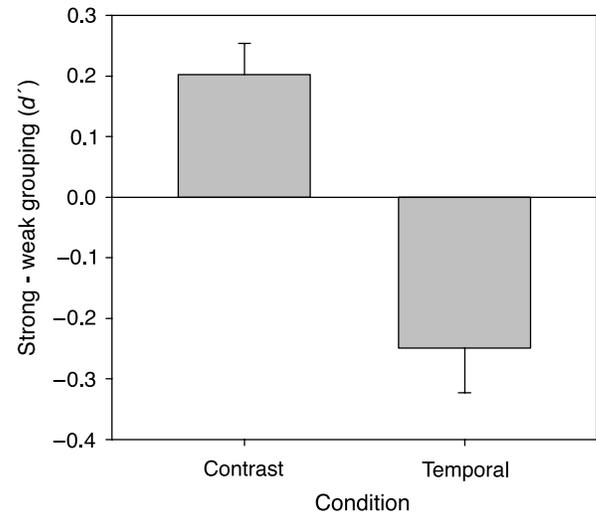


Figure 8. Difference in d' scores for 'strong' minus 'weak' grouping, separately for the contrast and temporal-phase discrimination tasks.

temporal discrimination tasks, with task order counter-balanced across observers.

Results and discussion

The results ([Figure 8](#)) show that the presence of coherent spatial structure leads to a dissociation in performance between the two tasks. For contrast judgments, 'strong' spatial structure facilitates performance, whereas for temporal judgments it impedes performance. This was reflected in the significant interaction between type of perceptual judgment (contrast vs. phase) and strength of grouping ('strong' vs. 'weak') ($F(1,9) = 15.0$, $p = 0.004$). This finding was further strengthened by comparisons between object grouping differences for each of the tasks separately. Sensitivity was higher for 'strong' compared to 'weak' grouping for the contrast judgment ($t(9) = 3.91$, $p = .004$), while there was a reversed, but still highly significant, trend for the temporal judgment (temporal $t(9) = 3.41$, $p = .008$).

General discussion

Using a computational model, we have shown that the presence of grouped (but irrelevant) spatial structure impedes the ability to detect simultaneity, or make a temporal judgment of synchrony, on two flickering targets, and that this effect is maximal at intermediate temporal phase-shifts. In a set of three psychophysical experiments, using both yes/no judgments and 2IFC, we have confirmed these predictions: irrelevant coherent spatial structure, connecting two flickering target elements, impedes

observers' ability to discriminate the targets' temporal relation (synchronous/asynchronous) and the interference is maximal at intermediate phase differences. This seemingly counter-intuitive result, however, matches the BBS predictions. An alternative interpretation of this result is that the interference is not specific to temporal synchrony judgments, but arises because of the need to segment the targets from the rest of the contour, so that any relative judgment of perceptual properties of the two targets should be affected. However, using matched stimuli we found that, while coherent spatial structure impedes temporal judgments between two targets, it enhances observers' ability to discriminate their relative contrast (Figure 8). This suggests that the effect of coherent spatial structure is fundamentally different for temporal and non-temporal judgments of the same targets. According to the BBS hypothesis, and in line with our results, whilst neural synchrony induced by coherent spatial structure facilitates grouping and the spread of attention between elements belonging to the same object, it also has a cost on the coding of other stimulus attributes. Specifically, it produces a reduction in sensitivity to differences in the temporal attributes of bound targets—due to internally generated synchrony masking externally imposed temporal relations. Such a reduction of temporal precision was, in fact, previously predicted for neural responses by BBS critics (Ghose & Maunsell, 1999).

The model used to derive these predictions is a highly schematic one, and a biologically more plausible model with multiple detectors for each location (Li, 1999) in the visual field remains to be examined. However, we believe that such complex models will validate the generic predictions we obtained here, and which follow from the simple fact that the coupling makes the responses more 'similar' due to the mutual support among the target detectors. Despite its simplicity, the model makes precise parametric predictions about the amount of interference as a function of the phase difference between the flickering elements: maximum interference at intermediate phase differences. Two important things are worth noticing: First, although the model and the experiments used periodic signals (regular flicker), its conclusion generalizes to judgments of synchrony between irregular or single-event signals, and future experiments are needed to test this. Second, the type of synchronization generated by the model (under the parameters we used here) corresponds to onset-synchronization (from each transient in the input), rather than to synchronization of internally generated oscillatory responses. We chose to focus here on the former, because it provides the simplest account for the temporal judgment performance in the task we examined and because grouping under 'stationary' stimuli may also benefit from such transient induced synchrony due to eye movements and microsaccades. The same model, however, is able (for slightly different excitation–inhibition parameters) to produce synchronization of internally generated oscillatory responses (see Appendix A

for examples of responses to steady and flickering stimuli, with and without coupling, and Niebur, Schuster, Kammen, & Koch (1991) for a model of internally generated synchronisation). The application of the oscillatory model to temporal judgment performance will be reported elsewhere (Cheadle & Usher, in preparation).

The results presented here are consistent with the finding that the presence of irrelevant temporal synchronization (the converse of the manipulation we used here) facilitates spatial grouping of elements (Elliott & Müller, 1998; Fahle, 1993; Lee & Blake, 2001; Parton, Donnelly, & Usher, 2001; Usher & Donnelly, 1998). While these previous findings are subject to a variety of alternative explanations (Dakin & Bex, 2002; Farid & Adelson, 2001), the present result—of a cost associated with spatial grouping, which is specific to judgments of synchrony between bound targets—cannot simply be explained by these proposals, and therefore strengthens support for the BBS hypothesis. Interestingly, a similar effect of degraded temporal resolution was reported as a result of spatial cueing of visual attention (Yeshurun & Levy, 2003), a cognitive process which has also been linked to synchronous neural activity (e.g., Fries, Reynolds, Rorie, & Desimone, 2001). Furthermore, the results may suggest one way in which the neural system may deal with visual inhomogeneity in object properties (contrast and luminosity conditions), which are likely to cause differential delays in cortical responses (but see Guttman, Gilroy, & Blake, 2007, for an alternative scheme). If an early visual process induces synchronization of responses to object parts (overcoming latency differences between responses caused by visual inhomogeneity), it can enable a higher visual area to pick up on grouped elements using a synchrony code (von der Malsburgh, 1995).

Nevertheless, at present, these data (and model) only provide indirect evidence towards a confirmation of the BBS mechanism in visual grouping. To further validate this support, future studies that examine the brain signature (e.g., via EEG) during temporal discrimination tasks of the type we used here would be of great importance. If the brain, indeed, encodes grouping of object elements by in-phase oscillations of their corresponding detectors, and if attentional processes are also mediated by enhanced local or global neural synchronization (Buschman & Miller, 2007; Fries et al., 2001; Vidal et al., 2006), this may provide a parsimonious explanation for the interactions between these fundamental visual computations.

Appendix A

Grouping induced synchronization in the internally generated oscillatory model

Here we present computer simulations of the same model (Wilson & Cowan, 1972), without noise, for a

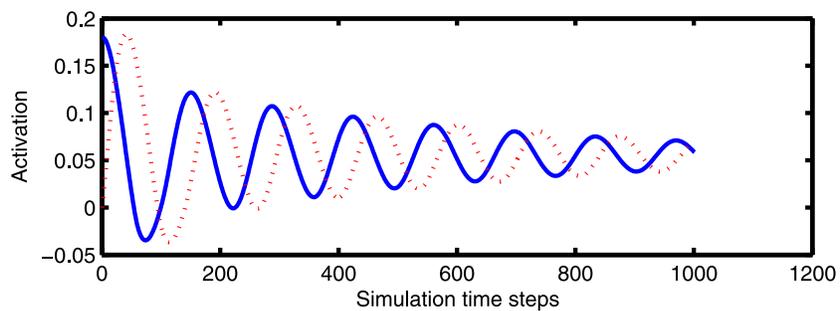


Figure A1. Constant input – No coupling ($A_2 = 0$). Damped oscillations with sustained phase shift.

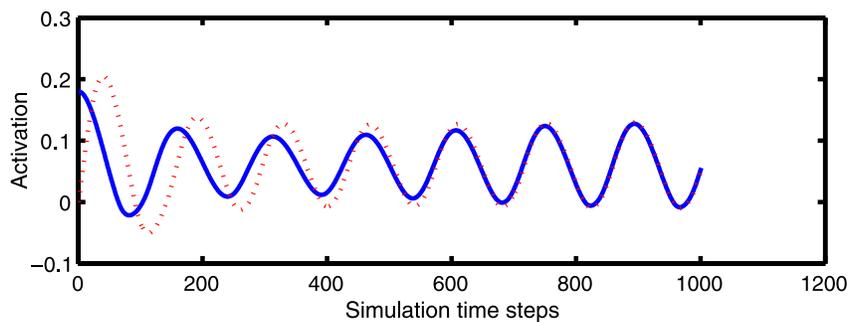


Figure A2. Constant input – Coupling ($A_2 = .5$). Sustained oscillations with diminishing phase shift.

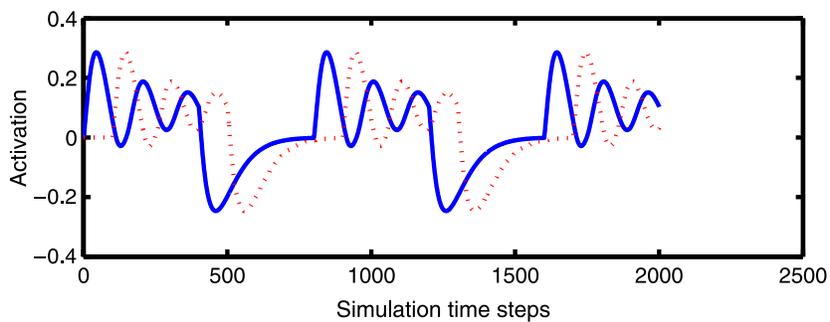


Figure A3. Square wave input – No coupling ($A_2 = 0$). Out of phase oscillations after each input onset.

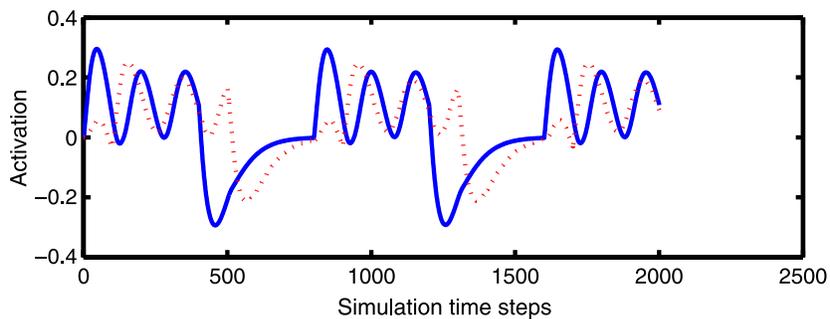


Figure A4. Square wave input – Coupling ($A_2 = .5$). Oscillations with diminishing phase shift. Note synchronisation is achieved by the third oscillation following each onset.

slight change in parameters that result in more oscillatory behavior. The model parameters are: $A_1 = 2$, $B = 3$, $C = 3$, $D = .1$, $SD_1 = 0$, $SD_2 = 0$, $I = .33$.

First we show the activation of the two detectors under steady (non-flickering input), for the uncoupled (Figure A1) and coupled (Figure A2) situation. One can see that now the detectors respond with fast internal oscillations, which for the uncoupled case are out of phase due to random starting point activations of the two detectors. With coupling the oscillatory responses get synchronized. This corresponds to grouping under steady (non-flickering input).

For all subsequent graphs, activations of unit 1 (solid blue curve) and unit 2 (dashed red curve) are plotted as a function of time.

Second, we present simulations that correspond to responses of the same two detectors to flickering input. Note the flicker period is much slower (800 simulation time steps) than the internal oscillation period (about 5–6 internal oscillations to one flicker oscillation). The two detectors are flickered out of phase (with 100 time-steps delay). One can see (Figures A3 and A4) that, like in the (non-oscillatory) model in the article, the coupling makes the responses more overlapping, thus it reduces the amount of asynchrony.

Acknowledgments

This work was supported by a grant of the German Research Council to H. Müller and M. Usher and a postdoctoral fellowship of the German Academic Exchange Service (DAAD) to F. Bauer.

Commercial relationships: none.

Corresponding author: Marius Usher.

Email: m.usher@bbk.ac.uk.

Address: School of Psychology, Birkbeck, University of London, Malet Street, WC1 7HX, London, UK.

References

- Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception, 1*, 371–394. [PubMed]
- Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science, 315*, 1860–1862. [PubMed]
- Cass, J., & Alais, D. (2006). The mechanisms of collinear integration. *Journal of Vision, 6*(9):5, 915–922, <http://journalofvision.org/6/9/5/>, doi:10.1167/6.9.5. [PubMed] [Article]
- Clifford, C. W., Holcombe, A. O., & Pearson, J. (2004). Rapid global form binding with loss of associated colors. *Journal of Vision, 4*(12):8, 1090–1101, <http://journalofvision.org/4/12/8/>, doi:10.1167/4.12.8. [PubMed] [Article]
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*, 42–45.
- Dakin, S. C., & Bex, P. J. (2002). Role of synchrony in contour binding: Some transient doubts sustained. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 19*, 678–686. [PubMed]
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501–517. [PubMed]
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., et al. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics, 60*, 121–130. [PubMed]
- Eckhorn, R., Frien, A., Bauer, R., Woelbern, T., & Kehr, H. (1993). High frequency (60–90 Hz) oscillations in primary visual cortex of awake monkey. *Neuroreport, 4*, 243–246. [PubMed]
- Elliott, M. A., & Müller, H. J. (1998). Synchronous information presented in 40-Hz flicker enhances visual feature binding. *Psychological Science, 9*, 277–283.
- Fahle, M. (1993). Figure-ground discrimination from temporal information. *Proceedings of the Royal Society B: Biological Sciences, 254*, 199–203. [PubMed]
- Fahle, M., & Koch, C. (1995). Spatial displacement, but not temporal asynchrony, destroys figural binding. *Vision Research, 35*, 491–494. [PubMed]
- Farid, H., & Adelson, E. H. (2001). Synchrony does not promote grouping in temporally structured displays. *Nature Neuroscience, 4*, 875–876. [PubMed]
- Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science, 291*, 1560–1563. [PubMed]
- Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 12699–12704. [PubMed] [Article]
- Ghose, G. M., & Maunsell, J. (1999). Specialized representations in visual cortex: A role for binding? *Neuron, 24*, 79–85. [PubMed] [Article]
- Gilbert, C. D. (1993). Circuitry, architecture, and functional dynamics of visual cortex. *Cerebral Cortex, 3*, 373–386. [PubMed]

- Gray, C. M. (1999). The temporal correlation hypothesis of visual feature integration: Still alive and well. *Neuron*, *24*, 31–47. [PubMed] [Article]
- Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*, 334–337. [PubMed]
- Guttman, S. E., Gilroy, L. A., & Blake, B. (2007). Spatial grouping in human vision: Temporal structure trumps temporal synchrony. *Vision Research*, *47*, 219–230. [PubMed] [Article]
- Hess, R. F., Dakin, S. C., & Field, D. J. (1998). The role of “contrast enhancement” in the detection and appearance of visual contours. *Vision Research*, *38*, 783–787. [PubMed]
- Kiper, D. C., Gegenfurtner, K. R., & Movshon, J. A. (1996). Cortical oscillatory responses do not affect visual segmentation. *Vision Research*, *36*, 539–544. [PubMed]
- Kourtzi, Z., Tolias, A. S., Altman, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fMRI studies. *Neuron*, *37*, 333–346. [PubMed] [Article]
- Kovács, I. (1996). Gestalten of today: Early processing of visual contours and surfaces. *Behavioral Brain Research*, *82*, 1–11. [PubMed]
- Lavie, N., & Driver, J. (1996). On the spatial extent of attention in object-based visual attention. *Perception & Psychophysics*, *58*, 1238–1251. [PubMed]
- Lee, S. H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science*, *284*, 1165–1168. [PubMed]
- Lee, S. H., & Blake, R. (2001). Neural synergy in visual grouping: When good continuation meets common fate. *Vision Research*, *41*, 2057–2064. [PubMed]
- Leonards, U., Singer, W., & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Research*, *36*, 2689–2697. [PubMed]
- Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, *50*, 951–962. [PubMed] [Article]
- Li, Z. (1999). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network*, *10*, 187–212. [PubMed]
- Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review*, *81*, 521–535. [PubMed]
- Niebur, E., Schuster, H. G., Kammen, D. M., & Koch, C. (1991). Oscillator-phase coupling for different two-dimensional network connectivities. *Physical Review A*, *44*, 6895–6904. [PubMed]
- O’Reilly, R., & Munakata, Y. (2001). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. MIT Press Cambridge, Massachusetts.
- Palanca, B. J., & DeAngelis, G. C. (2005). Does neuronal synchrony underlie visual feature grouping? *Neuron*, *46*, 333–346. [PubMed] [Article]
- Parton, A., Donnelly, N., & Usher, M. (2001). The effects of temporal synchrony on the perceived organisation of elements in spatially symmetric and asymmetric grids. *Visual Cognition*, *8*, 637–654.
- Parton, A., Donner, T. D., Donnelly, N., & Usher, M. (2006). Perceptual grouping based on temporal structure: Impact of subliminal flicker and visual transients. *Visual Cognition*, *13*, 481–452.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, *33*, 993–999. [PubMed]
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, *18*, 555–586. [PubMed]
- Usher, M., Bonneh, Y., Sagi, D., & Herrmann, M. (1999). Mechanisms for spatial integration in visual detection: A model based on lateral interactions. *Spatial Vision*, *12*, 187–209. [PubMed]
- Usher, M., & Cohen, J. D. (1999). Short term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphreys, & A. Olson (Eds.), *Connectionist models in cognitive neuroscience* (pp. 78–91). London: Springer-Verlag.
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, *394*, 179–182. [PubMed]
- Vidal, J. R., Chaumon, M., O’Regan, J. K., & Tallon-Baudry, C. (2006). Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals. *Journal of Cognitive Neuroscience*, *18*, 1850–1862. [PubMed]
- von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology*, *5*, 520–526. [PubMed]
- von der Malsburg, C. (1999). The what and why of binding: The modeler’s perspective. *Neuron*, *24*, 95–104. [PubMed] [Article]
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, *12*, 1–24. [PubMed] [Article]
- Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, *14*, 225–231. [PubMed]