Feature Binding and Object Perception. Does Object Awareness Require Feature Conjunction?

Dario Taraborelli

Feature Binding and Object Perception

Does object awareness require feature conjunction?

Dario Taraborelli

Institut Jean-Nicod, Paris
dario.taraborelli@ens.fr

Abstract : *Recent work in different fields of cognitive sciences seems to support the idea that in order to explain awareness of visually presented objects a subject needs some kind of binding mechanism, for the correct conjunction of different sensory features into a whole percept. Selective attention is commonly invoked as the key for solving this conjunction problem. Accordingly, the cognitive neurosciences have begun to investigate what kind of neural processing could underlie such a process. We analyse the evidence provided for justifying the existence of a binding problem and we argue against the claim that there is a feature binding problem to be solved in order to explain unity of object awareness. In particular, we question the definition of 'feature' and suggest some possible sources of misunderstanding related to this definition. Finally, we suggest an alternative approach in which perceived object unity does not rely exclusively on attentional conjunction of sensory features: instead of a general and unique mechanism mediating object awareness, we examine evidence that there are at least as many binding mechanisms as potential ways of interacting with visually presented objects.*

1. The ineffable unity of visually conscious objects

Every kind of activity we perform, every kind of interaction we establish with our environment seems to demand some kind of mechanism to *select* and *integrate* sensory information. In order to behave we need to package the world into units: seizing a glass, tracking the movement of a fly, pointing to a flower and, probably, even calling a person we see by his or her name are all tasks that demand some unit selection in order to be achieved. The theoretical notion of *object* is alleged to play this role. It has been shown that since early childhood humans are biased to interact with the world through entities sharing some important properties and that these entities represent the ground for the development of our object perception as adults[1]. Among the many properties objects thus conceived commonly have (they move coherently, they are usually middle sized and manipulable, they are topologically connected, i.e. they do not break apart and come together again across time) there is one that seems particularly problematic. Objects exhibit some kind of intrinsic ineffable unity, due to the fact that their features all *belong together* to form a whole[2]. A perceived object has commonly a color, a shape, a texture that are all instantiated at the same time in the same place. When we perceive an object we do not see its features apart from each other, but we perceive a spatially localized unit that has *at the same time* a given color, a given shape, a

particular texture and so on. We cannot consciously perceive an object as a whole without perceiving it as *having a value* for each of these sensory qualities[3]. Now, if we compare this feature co-instantiation property with the other properties objects usually have, we realize that the '*belonging together of features*' has something ineffable. We can easily imagine a configuration of stimuli not respecting either of the properties we mentioned above (stimuli lacking coherent movement, with no manipulable size, not connected etc…): in such cases we simply would say that there is no object, and there is rather a *collection* of objects or some kind of weird *entity* showing some visual properties and lacking others. What we are doing in such cases is checking our phenomenology to see if certain properties or configurations of stimuli result in a perceived unity or not.

The problem with feature conjunction is different. What does it mean in phenomenological terms that a consciously perceived object displays at the same time all its features and that these features are conjoined? Can we show what would be like to perceive all these features disjointly? The 'coming apart' of visual features is a phenomenological property one can check? All these questions seem to suggest that the unity we are talking of when we think of objects as conjunctions of features has a very different nature from other kinds of phenomenologically salient types of unity (unity by coherent motion, by boundedness, by topological connectedness etc.). The particular status of this type of perceptual unity has given birth to what is called the *Feature Binding Problem* ('FBP').

Current literature in psychology of perception presents both 1) theories that acknowledge feature binding as a genuine problem to be solved for explaining object awareness and 2) theories that, either by denying the relevance of sensory features for object perception or by insisting on the action-centered nature of visual awareness, reject the idea of a unitary FBP.

In what follows, we'll attempt to question the core idea of feature binding, namely that conscious perception of objects requires a mechanism of integration of sensory features, by analysing the several arguments provided to justify the existence of a FBP.

Why does feature binding represent a problem? The first argument we suggested was that feature conjunction is a problem for phenomenology, since features 'belong together' in space and we cannot perceive a localized shape without also giving it a color, a texture or other sensory features. There are at least two other ways of formulating the FBP:

A. a first class of formulations integrate phenomenological arguments with arguments drawn from known errors and difficulties in *visual performance*

B. a second class of formulations integrates phenomenological and psychological considerations with arguments drawn from particular properties of *neural coding of visual features*.

The two following paragraphs are dedicated to an analysis of both classes of arguments.

2. Feature binding and performance: from visual search to illusory conjunctions.

One way of characterizing the binding problem without solely relying on phenomenology, but rather on psychological performance, is what Frank Jackson [1977][4] called the *many-property problem*. How can we succeed in distinguishing a scene containing a red square and a green triangle from a scene containing a green square and a red triangle? The challenge consists in identifying an object by its unique combination of features, in Jackson's case: a specific conjunction of shape and color. If we were just sensitive to distinctions in color, or shape or shape-color conjunctions but unable to associate the proper feature conjunction to each visual item, we could not succeed in this task that requires the discrimination of two scenes with the same conjunction of visual attributes[5].

Recent perceptual psychology has studied a very similar problem from the perspective of visual search[6]. If we are shown a scene containing several items with different geometric shapes <triangle, square, circle>, each having a different color <red, yellow, blue, green> and we are asked to sort out only yellow triangles, we are faced with a classic problem of feature integration.. In order to point out the right target, and under the assumption that subjects apply different mechanisms to compare shapes and colors, an operation of comparison is needed between the result of the two distinct visual searches <right color items>, <right shape items>. It has been shown that whereas simple features – for instance a red line in a background of green lines – can be easily detected with no attention limits (independently of the number of distractors in the display), the detection of targets defined by a conjunction of features – such as a vertical red line in a background of vertical green and horizontal red lines – requires selective attention and strongly depends on the number of distractors presented to the subject[7].

Converging evidence on the problem of detection of feature conjunctions comes from another paradigm, that of 'illusory conjunctions'[8]. Subjects asked to verbally report properties of visually presented stimuli under high attentional load (short presentation time, large number of displayed items) make frequent errors resulting from false conjunctions of features belonging to different items. For instance, they report having seen a red square whereas the display only contained red

circles and green squares. Coherent results seem to come from both verbal reports or visual matching tasks, with letters as well as geometric shapes, and on a large variety of visual features such as color, shape, size and solidity (simple outlines vs. filled shapes). Illusory conjunctions seem also to be independent of relative distance between displayed items and of morphological difference or similarity of items whose features are switched.

These results from the illusory conjunctions paradigm as well as those from visual search paradigm provide a solid ground for a general theory of feature integration and the role of selective attention in combining features. According to this theory, binding of features is achieved by directing spatial attention serially to different locations of the visual field. Preattentively, the visual system just seem to parse the scene in feature maps quickly accessible for simple visual searches. In order to detect and correctly report conjunctions of features, selective attention is needed so as to specify feature information for each location and thus build a single 'object file' for further analysis and identification.

So far the overall picture does not present serious methodological problems, all the phenomena dealt with in these experiments result from tests of a subject's *performance* (not his *phenomenal experience*) under impoverished conditions and high computational load.

Some important consequences arise nevertheless as soon as these results are interpreted as evidence for the need of a general binding mechanism to explain 'the world as we perceive it' or our visual phenomenology. If visual objects are mere spatially localized conjunctions of features and in real world settings they are correctly identified only after attentional focus has 'bound' them, *one may be tempted to claim that to be visually aware of an object one needs a mechanisms for conjoining all its features*. In this way, binding is not only required for the correct attribution of features to an item, the detection of a specific feature conjunction or the distinction between two items sharing part of their features, but also for explaining 'coherent perception of even a single object'[9].

A solution for the FBP - which was originally required to explain failures in visual performance under difficult conditions – tacitly becomes an explanation of how visual phenomenology arises and how features come to 'belong together' in conscious object perception[10].

This shift from the domain of performance to the domain of phenomenology becomes particularly crucial when we consider its implications in the field of neuroscientific studies of object perception and visual awareness.

3. Feature binding and cortical maps: integration of distributed information

An established way of describing the architecture of the visual system in the human brain is to refer to the different 'feature modules' in the visual cortex and their reciprocal relations. Selective activity for each dimension of the stimuli has been discovered in distinct and topographically organized areas of the visual system in rats, cats and primates[11]. Non-invasive imaging techniques have shown that this modular organization is common to all vertebrates, including humans[12]. Hue, lightness, local motion, orientation, shape, spatial frequency, retinal disparity all correspond to retinotopic feature maps in the human cortex.

A large literature in neurophysiology, and in the computational and cognitive neurosciences has recently discussed the problems of information integration through different maps in the human brain. Issues about signal integration and processing have more and more often been referred to as 'binding problems'[13], thus introducing an increasing complexity in the use of this notion when applied to describe phenomenological properties, psychological processes and neural events.

Are all the FBPs the very same problem? We have already drawn a distinction between a psychological FBP and a more problematic notion of phenomenological FBP. In this paragraph we will attempt an analysis of the theoretical assumptions endorsed by the study of the FBP at the neural level.

The general formulation of the FBP in current neurosciences adds a further hypothesis to those already encountered at the psychological level: that of neural segregation of feature processing. Since the human brain processes different features in distinct retinotopic maps, and spatial information about features seems to be lost since the very first stages of cortical processing, the existence of a neural mechanism for relating sparse featural information to its common topographical coordinates seems quite natural. There are very plausible reasons to justify the fact that selective activity of neurons coding hue of a particular portion of the retina should be strictly related to that of neurons coding luminance, or that orientation related activity be associated to local motion related activity. But there is no a priori reason for identifying cortical modularity with informational segregation defined in psychological terms or vice versa. Since the development of the first Parallel Distributed Processing models we know that modular ('encapsulated') processing is compatible with distributed coding. Thus sparse coding of features of the same region of the retinal image does not necessarily require featural information to be tagged as 'belonging together', *unless we identify neural vehicle spatial location and represented spatial location*[14].

If we consider common formulations of the FBP in recent neuroscientific literature, we observe though a global commitment to the idea that distributed coding of features demands a mechanism of binding for explaining the feature's spatial co-instantiation.

'*If, as we have reason to believe, color and form are processed in separate parts of the nervous system, why does one not simply perceive circle, triangle, blue, green without knowing which form has which color ?*"

<div align="right">Attneave [1974][15]</div>

'*Most of the objects, people, and scenes we perceive produce complex, multidimensional, changing patterns of stimulation on the retina ? At least some of their attributes appear to be registered by independent neural channels, specializing in different aspects, such as orientation, color, spatial frequency, brightness... But this immediately raises the question of how the component properties are resynthesized into the correct compounds, so that we correctly see the shirt as blue and the trousers as gray, for example, rather then the reverse*'

<div align="right">Treisman [1977]</div>

'*It is now well established that in the early visual system of primates there are at least ten distinct visual areas, arranged in a branching hierarchy. Different cortical areas specialise, to some extent, in different features, one responding mainly to motion, another to colour, etc. As one proceeds to areas higher in the hierarchy the mapping of the visual field onto the cortical surface tends to become more diffuse. This is not however how we see the world. Our inner visual picture of the external world has unity. How then does the brain put together all these different activities to produce a unified picture so that, for example, for any object the right colour is associated with the right shape ?*'

<div align="right">Crick [1984]</div>

'*The various visual properties of objects in the field of view start out unified. The shape, the color, size, and direction of movement all come from the same object. They come from the same place in space and they co-occur in time. When one feature disappears typically they all disappear. But when the object is processed by the nervous system, at least some dimensions get parsed and are analyzed in different brain regions or in different cells within the same brain region. Single cells recordings and pattern of deficit following neurological disorder all suggest that different regions in the occipital, temporal, and parietal cortex process different features emanating from the same object. This distributed processing raises a problem. Suppose that two or more objects are present in the field of view, each having a different color, different shape, different location and the like. If one part of the brain codes color, for example, and another codes form, then how is it that later in the processing it is determined which color goes with which form ? What is the mechanism of reassembly ?*'

<div align="right">Keele et al. [1988]</div>

'*The physiological evidence for the binding problem comes from studies of neurons in extrastriate visual cortex of primates. One key observation is that different features of an object are processed to a certain extent by different neurons within the visual system. Logically, in order to identify the shape, color and motion of a stimulus, the visual system must somehow integrate the activity of these different shape-selective, color-selective, and motion-selective neurons. When only one stimulus is present in the visual field, this is not a difficult problem because these features can only be assigned to one possible stimulus. However, when multiple stimuli appear together in the visual field, which is the typical situation in « real-world » scenes, the visual system must assign the correct color, shape and motion signals to each object*'.

<div align="right">Reynolds & Desimone [1999]</div>

The interest of these definitions arises from the fact that, if we compare them to those mentioned in the above paragraphs concerning the psychological *explanandum*, the neuroscientific formulations of the FBP appear to be 'hybrid' insofar as they argue from properties of processing and modularity defined at a (subpersonal) neural level to properties of performance or phenomenal content defined at a (personal) psychological level, or vice versa.

The psychological (personal level) FBP may have an independent solution from that of underlying (subpersonal) neural processing: conflating them in one and the same problem violates a fundamental distinction between content and vehicle of information. Psychological modularity does not 'necessarily' require neural processing modularity, as well as psychological integration does not 'necessarily' require integration of neural signal. By '*not necessary*', we want to suggest that it may well be the case that psychological modularity is mirrored by modularity at the level of neural processing, but the former does not necessary entail the latter, as many programmatic statements seem tacitly to assume.

Millikan [1993] has recently drawn attention on the capillarity of vehicle/contents confusions in perception studies.

She points out how tempting in scientific investigations of perceptual phenomena is to apply a strategy of '**content internalization**', namely to project a set of properties of perceptual content onto properties of the alleged vehicles of this content (in this case neural activity) and to claim that *this* isomorphism of properties represents a genuine explanation of such perceptual content.

> *The error to be eradicated, then, certainly is not that of positing intermediaries. Postulation of intermediaries of some kind is essential to understanding perception and thought. The error is that of projecting, without argument, chosen properties of what is visaged or conceived onto these intermediaries and vice versa. The error is equally that of taking this sharing of properties to constitute an* explanation *of mental representing*
>
> Millikan [1993]

Some case studies, especially in the domain of temporality, have warned against the idea that simultaneity of perceived content must logically require simultaneity of stimulation or simultaneity of underlying neural activity[16].

Discrepancies between phenomenology and the alleged nature of underlying neural bases have suggested the existence of a large family of similar problems and the need for explications in terms of neural processes[17]: the existence of a blind spot in the visual field that is not consciously perceived or the fact that saccadic movements do not show up in visual experience have been

considered problems demanding special solutions (respectively, a filling-in mechanism and a process of saccadic suppression[18]). Both cases fall under the vehicle/content confusion illustrated by Millikan. Similarly, the apparent richness of our visual experience, its being 'present' and 'ongoing', its having a 'qualitative content' are all issues that traditionally urged for explication in terms of properties of underlying neural processing: such alleged explications – it has been argued – might all derive from the temptation of projecting content properties onto vehicle properties[19].

The case of feature binding, has drawn so far less attention, but it is at least as susceptible of epistemological fallacies as many other cases of neural explication of phenomenological content.

Both fallacious moves suggested by Millikan seem to be potentially present in many formulations of the FBP:

1. **Internalization of unity**: unity of features in aware object perception is projected onto unity of vehicles of these features (whereas phenomenological unity or success in feature conjunction tasks may be completely independent of unity in neural processing).

2. **Externalizion of segregation**: segregation of feature processing at a neural level are supposed to show up in perceptual content (whereas segregation of processing may be completely independent of failures in features conjunction)[20].

A large debate has addressed in the last decade the question of the *format* of neural integration of featural information and the hypothesis that large-scale synchronous activity may mediate signal processing between anatomically distant neural populations[21]. Even if this hypothesis turned to be empirically true, the claim that feature binding might explain unity of perceptual content in virtue of the existence of such large-scale synchronous activity – rather than old fashioned theories of grandmother cells or hypotheses of 'unitary loci of visual awareness'[22] – does not escape the fallacy of content internalisation: the alleged unitary vehicle is simply converted in this cases in a temporal one (spatial *co-instantiation* of sensory features is projected onto *synchronous* oscillations of cells coding different features of the same location in the visual field[23]).

An important difference in the neuroscientific formulations of the FBP is though to be acknowledged. Analogously to the psychological case, we may further distinguish among neuroscientific formulations of a FBP 1) statements that justify the existence of such a problem on

the ground of simple performance (Reynolds & Desimone [1999] – whose definition strongly remind Jackson's Multiple properties Problem) and 2) statements that on the other hand invoke phenomenological unity to justify the research of a binding mechanism (Crick [1984] clearly endorses this option when he states "*Our inner visual picture of the external world has unity*").

But again we observe that there's no logical reason to believe that if the phenomenal appearance of perceived objects has some properties of spatio-temporal co-instantiation (such as the fact that a region of the visual field cannot be consciously perceived without having a color or a given brightness value), then those properties must be directly mirrored in the dynamics of underlying neural events.

4. What is a feature?

A collateral issue in the study of Feature Binding concerns the nature of what is commonly called a 'feature'. Psychologists and neuroscientists agree that integration processes are required by most visual tasks, but even within the most robust theories of Feature Integration no comprehensive account is given about the nature of these basic 'visual features' and their possible list. A simple inspection at the set of 'features' mentioned by most authors does not seem to help. Feature Integration Theories normally identify basic features with sensory qualities derivable from psychophysical tests of discrimination. Their list of features is however not always sound. Properties that are commonly considered typically low level sensory features of stimuli are often lumped together with more high level perceptual properties (brightness vs. 3D depth). In the set of basic features we often find properties that seem to depend (at least from a classic computational perspective) from integration of others (shape vs. segment orientation). Some features are defined at the scale of a single receptive field while others are referred to larger portion of the retinal image (hue vs.size). Some works even add 'spatial location' to the list of features without realizing that the idea of binding itself is nothing but a spatial conjunction of information concerning visual attributes of the same item (Treisman & Gelade [1980], p.98).

This general lack of precision in the list of pertinent visual features might be a symptom of an important discrepancy between psychological and neuroscientific investigations of vision.

Clark [2000] recently suggested that 'feature' is an ambiguous term that might alternatively indicate:

1.1 sensory qualities defined by means of psychophysical tests,

1.2 properties of the stimuli that selectively trigger activity of early visual areas (and in some cases draw topographically organized maps of the retinal image),

1.3 general terms used in feature-placing sentences (such as 'here is green', or 'there is brighter than here').

Given this distinction, we have to take into account the possibility that:

- predicates used by subjects to describe their phenomenal experience may only loosely match with qualitative distinctions emerging from psychophysical tests

- psychophysically defined qualitative distinctions, in turn, may have no straightforward mapping to those properties of the stimuli that selectively trigger activity of given neuronal populations[24].

The very idea that sensory qualities need not correspond to those properties identified by patterns of activity in the early stage of visual processing is supported by the discovery of relatively early sensitivity to features that are normally treated as middle or high level visual properties[25].

Are these observations enough to motivate the research of an alternative account of object awareness not depending on integration of sensory features? In order to consider a possible alternative to the study of the relation between object awareness and features, let us briefly resume the global picture provided by standard Feature Integration Theories.

5. Doing without sensory features: binding and task-oriented perception

According to standard Feature Integration Theories, consciously perceiving an object is a matter of attentional integration of sensory properties. The visual system only registers featural discontinuities in the retinal image into distinct feature maps. These maps allow smart searching of single attributes (such as an item of a particular hue) across the visual field, but do not suffice for detecting conjunctions of features. Conjunctions require attention in order to be perceived: on the one hand conjunctions do not pop up in visual search as easily as simple features and require serial inspections and longer processing time; on the other hand, under impoverished conditions or high attentional load the visual system may fail to build the correct associations of features. Since objects are nothing else than spatially anchored conjunctions of features, attentional binding of these features is needed before we can consciously perceive them as unitary.

There is both psychological and neurophysiological evidence suggesting that things might not work that way. Recent experiments showed a number of cases in which sensory features such as those mentioned in much studies of attentional binding and neurophysiological research of early visual processing are not necessarily the most rapidly accessed properties of visual stimuli.

1) It has been recently shown for instance that texture perception and visual search can be based on properties such as lighting direction and surface slant – which are not considered sensory features in standard Feature Integration Theories[26].

2) Object recognition – traditionally considered the last stage of visual processing - may under particular conditions precede one of the most apparently primitive distinctions such as figure/ground judgment[27].

3) Depending on the visual routine the subject is asked to perform, features of visual search might correspond not to visual primitives but to structures formed at some higher level of processing. Visual primitives might conversely be accessible only later than more complex structures.

These results challenge the assumption that one general binding mechanism is responsible for object awareness. Being visually aware of an object seems more likely to be a task-dependent condition (or a condition strictly dependent on the visual routines the subject is accomplishing) rather then the result of an attentional processing of conjunction among low level sensory features[28].

6. Conclusions

The aim of this article was an analysis of some arguments provided in psychological and neuroscientific literature for the existence of a feature binding problem. The various justifications alleged for such a problem appear to share a common underlying assumption, drawn from a particular phenomenological property of conscious object perception, namely that sensory features appear to 'belong together' in a unitary percept. We are unable – it is argued – to perceive a localized shape without giving it at the same time a color, a texture, a luminance value etc.

This fact, together with evidence from visual performance (difficulties in visual detection of feature conjunctions, and cases of erroneous conjunctions of features belonging to distinct items under high attentional load), has led to the postulation of a mechanism responsible for the correct conjunction of features into a coherent unity. Feature Integration Theories traditionally appeal to

selective attention as the mechanism required for explaining feature conjunction. We have observed, though, that attentional focus may explain success in particular kinds of visual tasks but does not explain the 'belonging together' of features as a phenomenal property of conscious object perception.

A second class of evidence supporting the existence of a feature binding problem is drawn from properties of neural processing of sensory features in the visual system. Segregation of featural information in distinct cortical maps seems to require some mechanism of binding in order to tag the signal that codes different sensory attributes of an item to one and the same spatial location. We argued against the claim that the unitary appearance of consciously perceived objects logically requires an explanation *because of* the segregation of featural information. We showed that this conflation of content properties (unity of percept) and vehicle properties (segregation of information processing) is an instance of a common epistemological fallacy of content internalisation – which afflicts several other cases of explanation of conscious phenomena in vision science.

In order to clarify the relations between phenomenology and the psychological and neuroscientific integration problems, we examined some possible sources of confusion on the notion of 'feature'. A discrepancy between sensory qualities defined in psychophysical terms and visual attributes triggering selective activity in early visual areas showed that the identification between cortical maps and spatial properties of psychophysical features is not straightforward or mandatory.

We finally examined evidence supporting alternative theories of object perception showing that the alleged primitiveness of sensory features relative to more complex attributes of stimuli does not always hold. Object awareness and access to features are more task-related than feature integration theories seem to allow. In particular, even if attention mediates the detection of feature conjunctions when this is demanded by particular visual routines, there is no a priori reason to postulate that attentional binding is the *necessary* mechanism required to explain unitary object awareness. Again, unity of experienced content need not be explained in virtue of unity of representation or conjunction of featural information.

Abandoning justifications of the FBP that are grounded on content internalisation might free psychological and neuroscientific research from the need of explaining pseudoproblems such as those raised by discrepancies between phenomenological experience and properties of underlying information processing.

References

Briand, K. A., & Klein, R.M. (1987), Is Posner's « beam » the same as Treisman's « glue »? On the relation between visual orienting and feature integration theory, *Journal of Experimental Psychology : Human Perception and Performance*, 13, 228-241.

Briand, K. A., & Klein, R.M. (1989), Has feature integration theory come unglued? A reply to Tsal, *Journal of Experimental Psychology : Human Perception and Performance*, 15, 401-406.

Clark, A. (2000), A Theory of Sentience, Oxford, Oxford University Press.

Donk, M. (1999), Illusory conjunctions are an illusion : The effects of target-nontarget similarity on conjunction and feature errors, *Journal of Experimental Psychology : Human Perception and Performance*, 25, 1207-1233.

Enns, J.T., & Rensink, R.A. (1990), Influence of scene-based properties on visual search, *Science*, 247, 721-723.

Enns, J.T., & Rensink, R.A. (1990), Sensitivity to three-dimensional orientation in visual search, *Psychological Science*, 1, 323-326.

Felleman, D.J., & Van Essen, D.C. (1991), Distributed hierarchical processing in the primate cerebral cortex, *Cerebral cortex*, 1, 1-148.

Ghose, G.M., & Maunsell, J. (1999), Specialized representations in visual cortex : A role for binding ?, *Neuron*, 24, 79-85.

Gray, C.M. (1999), The temporal correlation hypothesis of visual feature integration : still alive and well, *Neuron*, 24, 31-47.

Green, M. (1991), Visual search, visual streams, and visual architectures, *Perception & Psychophysics*, 50, 388-403.

Hurley, S.L. (1998), *Consciousness in action*, London, England: Harvard University Press.

Keele , S.W., Cohen, A., Ivry, R., Liotti, M., & Yee, P. (1988), Tests of a temporal theory of attentional binding, *Journal of Experimental Psychology : Human Perception and Performance,* 14, 444-452.

Lennie, P. (1998), Single units and visual cortical organization, *Perception*, 27, 889-935.

Millikan, R. G. (1993), Content and vehicle. In: *Spatial Representation*, ed. N. Eilan, R. McCarthy, and B. Brewer, 256-68. Oxford: Blackwell.

Nissen, M.J. (1985), Accessing features and objects : Is location special ? in : M.I. Posner & O.S. Marin (eds.), *Attention and Performance*, vol. 11, 205-219, Hillsdale, NJ, Erlbaum.

Noë, A. & Thompson, E. (2001), *Neural correlates of consciousness and the matching content doctrine*, online paper, [http://www2.ucsc.edu/people/anoe/onlinepapers.html].

O'Regan, J.K. (1992), Solving the real mysteries of visual perception: The world as an outside memory, *Canadian Journal of Psychology*, 46 (3), 461-488.

O'Regan, J.K., & Noë, A. (2001), A sensorimotor account of vision and visual consciousness, *Behavioral and Brain Sciences*, 24 (5).

Pessoa, L., Thompson, E., & Noë, A. (1998), Finding out about filling in : a guide to perceptual completion for visual science and the philosophy of mind, *Behavioral and Brain Sciences*, 21, 723-802.

Peterson, M.A., & Gibson, B.S. (1994), Must figure-ground organization precede object recognition? An assumption in peril, *Psychological Science*, 5, 253-259.

Prinzmetal, W., Presti, D.E., & Posner, M.I. (1986), Does attention affect visual feature integration?, *Journal of Experimental Psychology : Human Perception and Performance*, 12, 361-369.

Prinzmetal. W., Diedrichsen, J & Ivry, R. B. (2000) Illusory Conjunctions are Alive and Well: A Reply to Donk (1999). *Journal of Experimental Psychology: Human Perception & Performance.*

Rensink, R.A. & Enns, J.T. (1995), Preemption effects in visual search: Evidence for low-level grouping, *Psychological Review*, 102, 101-130.

Reynolds, J.H., & Desimone, R. (1999), The role of neural mechanisms of attention in solving the binding problem, *Neuron*, 24, 19-29.

Roskies, A.L. (1999), The binding problem, *Neuron*, 24, 7-9.

Shadlen, M.N., & Movshon, J.A. (1999), Synchrony unbound : A critical evaluation of the temporal binding hypothesis, *Neuron*, 24, 67-77.

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

Spelke, E.S. (1990), Origins of Visual Knowledge, in: D. N. Osherson, S.M. Kosslyn & J.M. Hollerbach, *Visual Cognition and Action*, Cambridge, MIT Press.

Teller, D.Y. (1984), Linking propositions, *Vision Research*, 24, 1233-1246.

Teller, D.Y. & Pugh, E.N. (1983), Linking propositions in color vision, in: J.D. Mollon & L.T. Sharpe (Eds.), *Colour vision: physiology and psychophysics*, London, Academic Press.

Thompson, E., Noë, A., & Pessoa, L. (1999), Perceptual completion : A case study in phenomenology and cognitive science, in J. Petitot, J.-M. Roy, B. Pachoud & F.J. Varela (eds), *Naturalizing Phenomenology : Issues in contemporary phenomenology and cognitive science*, Stanford University Press.

Treisman, A. (1977), Focused attention in the perception and retrieval of multidimensional stimuli, *Perception & Psychophysics*, 22, 1-11.

Treisman, A. (1988), Features and objects : The Fourteenth Bartlett Memorial Lecture, *Quarterly Journal of Experimental Psychology,* 40A, 201-237.

Treisman, A. (1999), Solutions to the binding problem : progress through controversy and convergence, *Neuron*, 24, 105-110.

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

Treisman, A., & Kanwisher, N.G. (1998), Perceiving visually presented objects :recognition, awareness and modularity, *Current Opinion in Neurobiology*, 8, 218-226.

Treisman, A., & Sato, S. (1990), Conjunction search revisited, *Journal of Experimental Psychology : Human Perception and Performance*, 16, 459-478.

Treisman, A., & Schmidt, H. (1982), Illusory conjunctions in the perception of objects, *Cognitive Psychology*, 14, 107141.

Tsal, Y. (1989a), Do illusory conjunctions support the feature integration theory ? A critical review of theory and findings, *Journal of Eperimental Psychology : Human Perception and Performance*, 15, 394-400.

Tsal, Y. (1989b), Further comments on feature integration : A reply to Briand and Klein, *Journal of Eperimental Psychology : Human Perception and Performance*, 15, 407-410.

Van der Heijden, A.H.C. (1995), Modularity and attention, *Visual Cognition*, 2, 269-302.

Von der Malsburg, C. (1999), The what and why of binding : The modeler's perspective, *Neuron*, 24, 95-104.

Wandell, B.A. (1999), Computational Neuroimaging of Human Visual Cortex, *Annual Reviews in Neuroscience*, 22 (1999), 145-173.

Wolfe J.M., & Bennett S.C. (1997), Preattentive object files: Shapeless bundles of basic features, *Vision Research*, 37, 25-43.

Wolfe, J.M., & Cave, K.R. (1999), The psychophysical evidence for a binding problem in human vision, *Neuron*, 24, 11-17.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989), Guided search : an alternative to the feature integration model for visual search, *Journal of Eperimental Psychology : Human Perception and Performance*, 15, 419-433.

[1] Spelke [1990].

[2] These 'feature co-instantiation property' of visual objects corresponds to what Clark [2000] claims to be a property of perceived locations.

[3] Treisman & Gelade [1980], p.98: '*There seems to be no way we can consciously perceive an unattached shape without also giving it a color, size, brightness and location*'.

[4] See Clark [2000].

[5] Note that by *conjunction* one can either mean: a simple logical conjunction of properties (R∧S∧G∧T) or a conjunction of properties such as (R∧S) for a specified region of visual space or a spatially determined item and another conjunction (G∧T) for another region/item.

[6] Treisman & Gelade [1980]; Treisman & Sato [1990]; Nissen [1985]; Wolfe, Cave & Franzel [1989]; Wolfe & Bennett [1997].

[7] Treisman & Gelade [1980].

[8] Treisman & Schmidt [1982]; Prinzmetal, Presti & Posner [1986], Briand & Klein [1987] and [1989]; Tsal [1989], Wolf & Bennett [1997], Donk [1999], Prinzmetal, Diedrischsen & Ivry [2000].

[9] Wolfe & Cave [1999], p.11. See also: Treisman & Kanwisher [1998], p.220: '*Attention seems to be necessary for object tokens to mediate awareness*'

[10] See Van der Heijden [1999].

[11] Felleman & Van Essen [1991].

[12] Wandell [1999]– Computational Neuroimaging of Human Visual Cortex.

[13] Roskies [1999]; Treisman [1999].

[14] See Van der Heijden's [1999] commentary on Treisman [1999].

[15] See Green [1991].

[16] Dennett & Kinsbourne [1990].

[17] Noe & Thompson [2001].

[18] Pessoa et al. [1998]. O'Regan [1992].

[19] O'Regan & Noe [2001].

[20] Though such inferences are not sound according to Millikan's argument, it is an open empirical question to test if segregation or disruption of neural processing may systematically correlate to errors in feature conjunction, or if attentional processes enhance coordinate activity between feature maps, see Keel et al. [1988].

[21] Ghose & Maunsell [1999]; Gray [1999]; Shadlen & Movshon [1999]; Singer [1999]; Von der Malsburg [1999].

[22] Teller & Pugh [1983]; Teller [1984].

[23] See Hurley [1998] : '*Firing-pattern synchrony is sameness in the type of firing pattern; the suggestion is that it codes for sameness of object in content. It is of course an empirical question whether this hypothesis is true. However, we should not suppose that sameness of object represented must be encoded by sameness of firing pattern. There might be a specific systematic variation in firing pattern, without loss of information, between cell populations in different areas. Systematic difference rather than sameness of firing pattern might in principle code for the sameness of object, though it would again be an empirical question whether it ever does*'.

[24] Briand & Klein [1989], p.404. Van der Heijden [1999].

[25] See Rensink & Enns [1995] for a review of recent studies on visual primitives.

[26] Enns & Rensink [1990].

[27] Peterson & Gibson [1994].

[28] That real-world knowledge may substitute selective attention in integration tasks mediating visual awareness is partially acknowledged by the defenders of standard Feature Integration Theories. See Briand & Klein [1989], p.402.