The Binding Problem: How Your Brain Builds Reality

Authored by
mohammad looti

June 16, 2026

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

mohammad looti (2026). *The Binding Problem: How Your Brain Builds Reality*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=38076

The binding problem is a term used at the interface between neuroscience, cognitive science and philosophy of mind that has multiple meanings.

Firstly, there is the segregation problem: a practical computational problem of how brains segregate elements in complex patterns of sensory input so that they are allocated to discrete "objects". In other words, when looking at a blue square and a yellow circle, what neural mechanisms ensure that the square is perceived as blue and the circle as yellow, and not vice versa? The segregation problem is sometimes called BP1.

Secondly, there is the combination problem: the problem of how objects, background and abstract or emotional features are combined into a single experience. The combination problem is sometimes called BP2.

However, the difference between these two problems is not always clear. Moreover, the historical literature is often ambiguous as to whether it is addressing the segregation or the combination problem.

The Segregation Problem

Definition

The segregation problem is the problem of how brains segregate elements in complex patterns of sensory input so that they are allocated to discrete "objects".

Smythies defined BP1 in these terms: How is the representation of information built up in the neural networks that there is one single object 'out there' and not a mere collection of separate shapes, colours and movements? Revonsuo refers to this as the problem of "stimulus-related binding" - of sorting stimuli. Although usually referred to as a problem of binding, the computational problem is arguably one of discrimination. Thus, in the words of Canales et al.: "to bind together all the features of one object and segregate them from features of other objects and the background". Bartels and Zeki describe it as "determining that it is the same (or a different) stimulus which is activating different cells in a given visual area or in different visual areas".

Experimental Work

Most experimental work is on vision, where it is known that humans and other mammals process different aspects of perception by separating information about those aspects and processing them in distinct regions of the brain. For example, Bartels and Zeki have shown that different areas in the visual cortex specialize in processing the different aspects of colour, motion, and shape. This type of modular coding has been claimed to yield a potential for ambiguity. When humans view a scene containing a blue square and a yellow circle, some neurons signal in response to blue,

ARABPSYCHOLOGY.COM

others signal in response to yellow, still others to a square shape or a circle shape. Here, the binding problem is the issue of how the brain correctly pairs colour and shape, i.e. indicates that blue goes with square, rather than yellow.

Synchronization Theory

A popular hypothesis perhaps first suggested by Milner has been that features of individual objects are bound/segregated via synchronisation of the activity of different neurons in the cortex. The theory is that when two feature-neurons fire synchronously they are bound, while when they fire out of synchrony they are unbound. Empirical testing of the idea was given impetus when von der Malsburg proposed that feature binding posed a special problem that could not be covered simply by cellular firing rates. A number of studies suggested that there is indeed a relationship between rhythmic synchronous firing and feature binding. This rhythmic firing appears to be linked to intrinsic oscillations in neuronal somatic potentials, typically in the gamma range close to 40 Hz. However, Thiele and Stoner found that perceptual binding of two moving patterns had no effect on synchronisation of the neurons responding to the two patterns. In the primary visual cortex, Dong et al. found that whether two neurons were responding to contours of the same shape or different shapes had no effect on neural synchrony. Revonsuo reports similar negative findings.

The positive arguments for a role for rhythmic synchrony in resolving the segregational object-feature binding problem (BP1) have been summarized by Singer. There is certainly extensive evidence for synchronization of neural firing as part of responses to visual stimuli. However, there is inconsistency between findings from different laboratories. Moreover, a number of recent reviewers, including Shadlen and Movshon and Merker have raised concerns.

Shadlen and Movshon, raise a series of doubts about both the theoretical and the empirical basis for the idea of segregational binding by temporal synchrony. Firstly, it is not clear that binding does pose a special computational problem of the sort proposed by von der Malsburg. Secondly, it is unclear how synchrony would come to play a distinct role in terms of local computational logic. Thirdly, it is difficult to envisage a situation in which pre-synaptic firing rate and synchrony could be usefully interpreted independently by a post-synaptic cell, since the two are interdependent over plausible time scales.

Another point that has been raised is that within standard time frames for neuronal firing very few distinct phases of synchrony would be distinguishable even under optimal conditions. However, this would only be significant if the same pathways are potentially fed spike (signal) trains in multiple phases. In contrast, Seth describes an artificial brain-based robot that demonstrates multiple, separate, widely distributed neural circuits, firing at different phases, suggesting that synchrony may assist the establishment of discrete object-related re-entrant circuits in a system exposed to randomly timed stimuli.

ARABPSYCHOLOGY.COM

Goldfarb and Treisman point out that a logical problem appears to arise for binding solely via synchrony if there are several objects that share some of their features and not others. When viewing a display of variously coloured letters, internal representation of a red X, a green O, a red O and a green X cannot be accounted for purely by synchrony of signals for red and X shape, for instance. At best synchrony can facilitate segregation supported by other means (as von der Malsburg acknowledges).

A number of neuropsychological studies suggest that the association of colour, shape and movement as "features of an object" is not simply a matter of linking or "binding". Purves and Lotto give extensive evidence for top-down feedback signals that ensure that sensory data are handled as features of (sometimes wrongly) postulated objects early in processing. In many illusions data appear as if pre-consciously adjusted in accordance with object expectations. Pylyshyn has also emphasized the way the brain seems to pre-conceive objects to which features are to be allocated and which are attributed continuing existence even if features, like colour, change.

Feature Integration Theory

In her feature integration theory, Treisman suggested that binding between features is mediated by the features' links to a common location. Psychophysical demonstrations of binding failures under conditions of full attention provide support for the idea that binding is accomplished through common location tags.

An implication of these approaches is that sensory data such as colour or motion may not normally exist in "unallocated" form. For Merker: "The 'red' of a red ball does not float disembodied in an abstract color space in V4." If colour information allocated to a point in the visual field is converted directly, via the instantiation of some form of propositional logic (analogous to that used in computer design) into colour information allocated to an "object identity" postulated by a top-down signal as suggested by Purves and Lotto, (e.g. There is blue here + Object 1 is here = Object 1 is blue) no special computational task of "binding together" by means such as synchrony may exist. (Although Von der Malsburg poses the problem in terms of binding "propositions" such as "triangle" and "top", these, in isolation, are not propositional.)

How signals in the brain come to have propositional content, or meaning, is a much larger issue. However, both Marr and Barlow suggested, on the basis of what was known about neural connectivity in the 1970s that the final integration of features into a percept would be expected to resemble the way words operate in sentences.

The role of synchrony in segregational binding remains controversial. Merker has recently suggested that synchrony may be a feature of areas of activation in the brain that relates to an "infrastructural" feature of the computational system analogous to increased oxygen demand

ARABPSYCHOLOGY.COM

indicated via MRI. Apparent specific correlations with segregational tasks may be explainable on the basis of interconnectivity of the areas involved. As a possible manifestation of a need to balance excitation and inhibition over time it might be expected to be associated with reciprocal re-entrant circuits as in the model of Seth et al. (Merker gives the analogy of the whistle from an audio amplifier receiving its own output.)

If it transpires that synchronized activity plays at most an infrastructural role in segregative computational "binding", the question arises as to whether we need another explanation. The implication both of Shadlen and Movshon's and of Merker's analyses seems to be that there may be no special binding problem in this sense. The problem may be merely an integral part of the more general problem of the computational logic used by neurons, or what is often referred to as the "neural code". In particular it may be inappropriate to analyse binding in perception without taking into account the way features are bound in memory, as addressed by Zimmer and colleagues, and how that informs the way the brain pre-conceives objects.

ARABPSYCHOLOGY.COM

The Combination Problem

Definition

Smythies defines BP2 as How do the brain mechanisms actually construct the phenomenal object?. Revonsuo equates this to "consciousness-related binding", emphasizing the entailment of a phenomenal aspect. As Revonsuo explores in 2006, there are nuances of difference beyond the basic BP1:BP2 division. Smythies speaks of constructing a phenomenal object ("local unity" for Revonsuo) but philosophers such as Descartes, Leibniz, Kant and James (see Brook and Raymont) have typically been concerned with the broader unity of a phenomenal experience ("global unity" for Revonsuo) - which, as Bayne illustrates may involve features as diverse as seeing a book, hearing a tune and feeling an emotion. Further discussion will focus on this more general problem of how sensory data that may have been segregated into, for instance, "blue square" and "yellow circle" are to be re-combined into a single phenomenal experience of a blue square next to a yellow circle, plus all other features of their context. There are a wide range of views on just how real this "unity" is, but the existence of medical conditions in which it appears to be subjectively impaired, or at least restricted, suggests that it is not entirely illusory.

History

Early philosophers such as Descartes and Leibniz noted that the apparent unity of our experience is an all-or-none qualitative characteristic that does not appear to have an equivalent in the known quantitative features, like proximity or cohesion, of composite matter. William James, in the nineteenth century, considered the ways the unity of consciousness might be explained by known physics and found no satisfactory answer. He coined the term "combination problem", in the specific context of a "mind-dust theory" in which it is proposed that a full human conscious experience is built up from proto- or micro-experiences in the way that matter is built up from atoms. James claimed that such a theory was incoherent, since no causal physical account could be given of how distributed proto-experiences would "combine". He favoured instead a concept of "co-consciousness" in which there is one "experience of A, B and C" rather than combined experiences. A detailed discussion of subsequent philosophical positions is given by Brook and Raymont (see 26). However, these do not generally include physical interpretations. James remained concerned about the absence of a "single physical thing", other than an atom, that could be co-conscious (of A, B and C), echoing Leibniz.

Whitehead proposed a fundamental ontological basis for a relation consistent with James's idea of co-consciousness, in which many causal elements are co-available or "compresent" in a single event or "occasion" that constitutes a unified experience. Whitehead did not give physical specifics but the idea of compresence is framed in terms of causal convergence in a local interaction consistent with physics. Where Whitehead goes beyond anything formally recognized in physics is in the "chunking" of causal relations into complex but discrete "occasions". Even if such occasions can be defined, Whitehead's approach still leaves James's difficulty with finding a site, or sites, of causal convergence that would make neurobiological sense for "co-consciousness". Sites of signal convergence do clearly exist throughout the brain but there is a concern to avoid re-inventing what Dennett calls a Cartesian Theater or single central site of convergence of the form that Descartes proposed.

Descartes's central "soul" is now rejected because neural activity closely correlated with conscious perception is widely distributed throughout the cortex. The remaining choices appear to be either separate involvement of multiple distributed causally convergent events or a model that does not tie a phenomenal experience to any specific local physical event but rather to some overall "functional" capacity. Whichever interpretation is taken, as Revonsuo indicates, there is no consensus on what structural level we are dealing with - whether the cellular level, that of cellular groups as "nodes", "complexes" or "assemblies" or that of widely distributed networks. There is probably only general agreement that it is not the level of the whole brain, since there is evidence that signals in certain primary sensory areas, such as the V1 region of the visual cortex (in addition to motor areas and cerebellum), do not contribute directly to phenomenal experience.

Modern Theories

Dennett has proposed that our sense that our experiences are single events is illusory and that, instead, at any one time there are "multiple drafts" of sensory patterns at multiple sites. Each would only cover a fragment of what we think we experience. Arguably, Dennett is claiming that consciousness is not unified and there is no phenomenal binding problem. Most philosophers have difficulty with this position (see Bayne). Dennett's view might be in keeping with evidence from

ARABPSYCHOLOGY.COM

recall experiments and change blindness purporting to show that our experiences are much less rich than we sense them to be - what has been called the Grand Illusion. However, few, if any, other authors suggest the existence of multiple partial "drafts". Moreover, also on the basis of recall experiments, Lamme has challenged the idea that richness is illusory, emphasizing that phenomenal content cannot be equated with content to which there is cognitive access.

Dennett does not tie drafts to biophysical events. Multiple sites of causal convergence are invoked in specific biophysical terms by Edwards and Sevush. In this view the sensory signals to be combined in phenomenal experience are available, in full, at each of multiple sites. To avoid non-causal combination each site/event is placed within an individual neuronal dendritic tree. The advantage is that "compresence" is invoked just where convergence occurs neuro-anatomically. The disadvantage, as for Dennett, is the counter-intuitive concept of multiple "copies" of experience. The precise nature of an experiential event or "occasion", even if local, also remains uncertain.

The majority of theoretical frameworks for the unified richness of phenomenal experience adhere to the intuitive idea that experience exists as a single copy, and draw on "functional" descriptions of distributed networks of cells. Baars has suggested that certain signals, encoding what we experience, enter a "Global Workspace" within which they are "broadcast" to many sites in the cortex for parallel processing. Dehaene, Changeux and colleagues have developed a detailed neuro-anatomical version of such a workspace. Tononi and colleagues have suggested that the level of richness of an experience is determined by the narrowest information interface "bottleneck" in the largest sub-network or "complex" that acts as an integrated functional unit. Lamme has suggested that networks supporting reciprocal signaling rather than those merely involved in feed-forward signaling support experience. Edelman and colleagues have also emphasized the importance of re-entrant signaling. Cleeremans emphasizes meta-representation as the functional signature of signals contributing to consciousness.
In general, such network-based theories are not explicitly theories of how consciousness is unified, or "bound" but rather theories of functional domains within which signals contribute to unified conscious experience. A concern about functional domains is what Rosenberg has called the boundary problem; it is hard to find a unique account of what is to be included and what excluded. Nevertheless, this is, if anything is, the consensus approach.

Within the network context, a role for synchrony has been invoked as a solution to the phenomenal binding problem as well as the computational one. In his book, The Astonishing Hypothesis, Crick appears to be offering a solution to BP2 as much as BP1. Even von der Malsburg, introduces detailed computational arguments about object feature binding with remarks about a "psychological moment". The Singer group also appear to be interested as much in the role of synchrony in phenomenal awareness as in computational segregation.

ARABPSYCHOLOGY.COM

The apparent incompatibility of using synchrony to both segregate and unify might be explained by sequential roles. However, Merker points out what appears to be a contradiction in attempts to solve the phenomenal unification problem (BP2) in terms of a functional (effectively meaning computational) rather than a local biophysical, domain, in the context of synchrony.

Functional arguments for a role for synchrony are in fact underpinned by analysis of local biophysical events. However, Merker points out that the explanatory work is done by the downstream integration of synchronized signals in post-synaptic neurons: "It is, however, by no means clear what is to be understood by 'binding by synchrony' other than the threshold advantage conferred by synchrony at, and only at, sites of axonal convergence onto single dendritic trees..." In other words, although synchrony is proposed as a way of explaining binding on a distributed, rather than a convergent, basis the justification rests on what happens at convergence. Signals for two features are proposed as bound by synchrony because synchrony effects downstream convergent interaction. Any theory of phenomenal binding based on this sort of computational function would seem to follow the same principle. The phenomenality would entail convergence, if the computational function does.

Although BP1 and BP2 are different, this need not invalidate the assumption, implicit in many of the quoted models, that computational and phenomenal events, at least at some point in the sequence of events, parallel each other in some way. The difficulty remains in identifying what that way might be. Merker's analysis suggests that either (1) both computational and phenomenal aspects of binding are determined by convergence of signals on neuronal dendritic trees, or (2) that our intuitive ideas about the need for "binding" in a "holding together" sense in both computational and phenomenal contexts are misconceived. We may be looking for something extra that is not needed. Merker, for instance, argues that the homotopic connectivity of sensory pathways does the necessary work.
The nature of, and solution to, BP2 remains a matter of controversy.

ARABPSYCHOLOGY.COM