Introduction to the Special Issue on Binding

James R. Brockmole [1] **and Steven L. Franconeri** [2]

[1]*University of Edinburgh*   [2]*Northwestern University*

In many ways visual processing acts as a prism, splitting visual information from the retinal image into separately processed visual features such as colours, shapes, sizes, or orientations. *Binding* refers to the set of cognitive and neural mechanisms that re-integrate these features to create a holistic representation of the objects in the visual field. The *binding problem* in vision refers to how this integration is achieved. Because binding processes must be applied to virtually every level of visual processing, the binding problem is a ubiquitous one that applies to perception, attention, working memory, and long-term memory systems. It is more appropriate, therefore, to conceptualize the binding problem not as a singular problem, but as a constellation of interrelated problem*s*. For example, how does one perceive and represent the fact that features *a*, *b*, and *c* belong to object *X* while features *q*, *r*, and *s* belong to object *Y*? How does one perceive and represent the relationships among objects such as the fact that object *Y* is to the left of object *X* and to the right of object *Z*? How does one perceive and represent that events *x* and *y* are temporally related to each other, but are distinct from events *p* and *q*? What happens to bound representations when the relations among features, objects, or events change? How functional are these representations in the service of visually-guided tasks?

This special issue of *Visual Cognition* provides a sampling of current work being done in visual perception and visual memory to understand how human observers accomplish the appropriate binding of features within objects, the relational binding among objects, and the binding between temporally related events. Within these broad areas, the articles cover a wide range of topics including the role of attention in feature binding, the representation of static and moving multi-feature objects in memory, the binding of objects to scenes, binding processes involved in learning and long-term memory, binding of information about objects perceived in visual and non-visual modalities, the development of binding abilities, and the neural substrates involved in binding processes. The goal of this special issue is not to disseminate conclusive solutions to the various instantiations of the binding problem. Although this collection of articles converges on several theoretical points, there are also examples of clear divergence. It is our hope, rather, that this collection of work will describe the current state-of-the-science regarding binding processes in visual perception and memory, outline the critical issues that have yet to be resolved, and highlight the interconnections between the binding problem*s* and the approaches taken to solve them. In this single volume readers will encounter work with children, young adults, and patients, and work that uses traditional behavioural measures, eye movement recording, functional imaging, and transcranial magnetic stimulation.

The special issue begins by examining the role attention plays in binding. Although attention will be important in many of the articles in this volume, the first set of papers make it their primary focus and take up the classic and hotly debated question regarding the degree to which, and conditions under which, binding relies on focused attention.

Hyun, Woodman, and Luck examined the attentional requirements of explicitly binding surface features to their spatial locations in visual perception. The authors asked observers to either indicate whether a target color was present in a display, or to report the location of the target color. While the first task simply required observers to detect a feature, the second task additionally

required observers to bind a feature to a location. The principle measure of participants' relative allocation of attention resources was the lateralized N2pc ERP component, which shows a larger amplitude when a relatively greater share of spatial attention is allocated to objects in the visual field contralateral to certain posterior electrodes. A larger N2pc component was observed in the localization condition, suggesting that greater attentional resources are required to discriminate color-location bindings than to detect individual colors. Thus, attention plays a special role in the binding of features to locations and the attentional mechanism responsible for this binding operates at a level indexed by the N2pc which includes intermediate and high levels of extrastriate visual cortex.

Braet & Humphreys also emphasize the role of attention in a two-stage account of binding. In the first stage, they argue that features are weakly bound to objects and that these bindings are noisy and potentially fleeting. The second stage is the application of top-down attentional feedback which consolidates and stabilizes visual representations. After an initial feed-forward sweep of activation in visual pathways, re-entrant processes stemming from higher visual areas such as the Posterior Parietal Cortex (PPC) tune the incoming signals. If this tuning includes the stabilization of representations of bound objects, then selectively disrupting these late processes might selectively disrupt binding and produce, for example, illusory conjunctions. To test the importance of the proposed re-entrant processes in binding, the authors asked a patient with bilateral PPC damage to perform a rapid binding task in which she was to identify the color and identity of a target letter that was presented alongside a different (in both color and identity) distractor letter. In this patient, presenting the objects for long durations (300ms) – when re-entrant feedback processes should normally help to consolidate binding – led to more binding errors (illusory conjunctions) than short presentations (100ms). This pattern contrasted with performance for reporting individual features, which improved with longer presentation times. In addition, the authors disrupted re-entrant binding processes in normal observers using transcranial magnetic stimulation of the PPC. The pulse only disrupted binding performance at an above-chance level when it was delivered 150-200ms after the presentation of the binding task, suggesting that it disrupted re-entrant processes directed by the PPC. The data lend support to a multi-stage model of binding where attentional feedback, specifically from PPC, plays an important role in relatively late perceptual processes related to feature binding.

As in perception, **Fougnie & Marois** argue that spatial attention also plays an important role in maintaining bound *memory* representations of objects. The authors critique several recent papers demonstrating that observers can retain feature bindings in memory even in the face of attentional distraction. The authors suggest that the visual attention tasks used in past studies were not difficult enough, and instead use a taxing multiple object tracking (MOT) task during the retention interval of a working memory task. The MOT task impaired performance far more when the memory task required binding compared to when binding was not required, suggesting that spatial attention is indeed necessary for maintaining bindings in memory. Interestingly, however, spatial attention was only critical for binding when the objects in the memory task were presented simultaneously in different spatial locations. When the objects were presented sequentially at the center of the screen, the object tracking task had a much smaller impact on performance in the binding condition. The authors conclude that spatial attention plays an important, although perhaps not an exclusive, role for binding in visual memory.

Oakes, Messenger, Ross-Sheehy, and Luck take a developmental approach to investigating the attentional control mechanisms underlying binding in visual working memory. When shown alternating displays of colored squares, 6-month-old infants are able to detect changes to a square's color when only one is present, as demonstrated by a looking preference for the changing stream, relative to a non-changing stream. However, once multiple squares of heterogeneous color are present, and

colors cycle among the squares on each new display, 6-month-old infants do not demonstrate a preference for the changing stream, suggesting that the 6-month-olds fail to bind or appropriately compare object feature information from one display to the next. Surprisingly, even when shown displays of heterogeneously colored squares that are replaced with a set of entirely new colours, six month olds continued to show no preference for the changing stream. This finding is counterintuitive because the same population of infants can detect color changes to single object arrays, and successful binding of a single color to location would allow detection of the changes in these display containing three times the opportunity to notice an identical change. Thus, although the infants have the ability to remember a single color and compare it to a later color, increasing display complexity leads to a failure of binding, even though many more changes may occur within the display. After only one more month of development, 7-month-old infants are able to complete all of these tasks. The authors argue that this in this extra month the infants might benefit from maturation of the parietal areas important for the control of attention, and these same areas are argued by Braet and Humphreys (this issue) to support binding.

In contrast to the apparent convergence on the role of attention in feature binding in the preceding papers, the final three papers in this set challenge the conclusion that attention is necessarily required for binding. In the first of these, **Allen, Hitch, and Baddeley** examined the role of attention in the generation and maintenance of working memory representations that must bind object features that are directly perceived and those that are described verbally. Observers were asked to retain object representations in memory for either unified visual stimuli (for example, seeing a red circle), or cross-modal stimuli in which one feature was presented visually and the other auditorily (for example, seeing the outline of a circle and hearing the word "red"). Performance in a visual discrimination task, in which observers judged whether a colored shape was part of the to-be-remembered set, was broadly equivalent across the presentation conditions,

indicating that purely visual and cross-modal representations are equally effective in supporting the recognition of visual objects. Concurrent tasks that led to performance decrements equally affected the visual and cross-modal binding conditions, suggesting that the act of binding cross-modal information is no more attention demanding than binding within visual memory. The authors argue that bindings are held in a general passive store (the episodic buffer) that may be fed from a range of features and modalities with the role of focused attention limited to controlling the relative contribution made by different features to memory.

In the next paper of this set, **van Rullen** challenges the idea that all complex object recognition requires attention to bind the object's features. Instead, he proposes that two modes of binding may exist. The first is a "hardwired binding" which operates when an observer is confronted with frequently encountered natural objects. With frequent exposure, networks might develop that include detectors for common feature conjunctions, eliminating the need for selective attention to perform the binding. But a second type of "on-demand" binding, mediated by attention, is needed for arbitrary or less frequently encountered feature conjunctions. This framework can explain many results that seem inconsistent with the need for attention in binding, such as the rapid categorization of real-world objects and natural scenes.

A distinction between "hardwired" and "on-demand" binding is also discussed by **Hommel and Colzato**. Observers were presented with two objects in sequence. Each object was either a strawberry or a banana. Upon presentation of the first object observers were only required to press a pre-defined key. Upon presentation of the second object, observers made a speeded judgment of the object's identity (i.e. shape), regardless of the object's color (which varied). Although the first object was irrelevant to this classification task, response times to the second shape were fastest when shape and color of both objects matched, demonstrating priming from the first shape. Interestingly, response times were also faster when

neither shape nor color matched, relative to when only one dimension matched. That is, preceding the shape judgment with an object that matches in only one dimension actually caused a relative slowing of response time to the target object. The authors interpret this effect as demonstrating that the color and shape of the first strawberry or banana are automatically bound, because long-term memory for these objects emphasizes the importance of shape and color information for recognition of these frequently encountered objects (as opposed to location information, which is not as important for identifying fruit). Critically, when the fruit were replaced with circles and triangles, which should not produce a strong LTM influence over the relevant dimensions for recognition, this one-dimension priming cost was minimized. Together, these results suggest that experience with frequently encountered objects builds schemas of the types of features that are important to bind in the recognition of those objects.

The next two papers consider the fragility of bindings in visual working memory. When perceptual input changes, how easily can existing bindings be maintained? **Alvarez and Thompson** showed observers a set of colored dots which rotated around a central fixation point, occasionally disappearing behind and reappearing from occluders. Observers were to keep track of where each color was located throughout the trial. At the end of a trial the dots stopped behind the occluders and observers were asked to report either the color behind a specific occluder or the location of a specific color. Performance on these tasks was relatively high, compared to when observers were instead asked to indicate whether, during the trial, two colors had switched locations. These results suggest that past work measuring binding memory, which largely relies on asking observers to detect swapped features, may have underestimated one's ability to maintain object-feature bindings for moving objects. To account for this difference, the authors further show that when feature-switches are missed, observers base their recall on the updated object-feature bindings rather than the original bindings, suggesting that failures to detect feature-switches arise because updated perceptual input leads to the automatic rewriting and rebinding of otherwise fragile information.

Logie, Brockmole, and Vandenbroucke continue with this line of enquiry by using working memory tasks to explore whether memory traces of individual features and feature conjunctions are completely lost from trial to trial, or whether they can be retained in long-term memory. Across trials, either the same to-be-remembered color-shape (integrated object) combinations were consistently repeated or one feature was repeated while the other randomly varied. Performance on change-detection and cued recall tasks improved when the same integrated objects were repeatedly presented on every trial, indicating long term retention of bindings in memory. For repeated single features learning effects were markedly weaker than those observed for integrated objects. All learning effects were eliminated if just two novel object arrays were inserted between each repeated array. Consistent with the results presented by Alvarez and Thompson (this issue), these results suggest that bindings in memory are fragile and are readily displaced in memory when new bindings are created. Despite this fragility, however, longer-term residual traces can be generated to support learning, although any learning that does occur is largely based on bound objects, not individual features.

The next four papers consider binding mechanisms across time as well as space. In the first of these papers, **Makovski and Jiang** considered whether tracking multiple moving objects is easier when the target objects can be distinguished by simple features or combinations of features. While better tracking performance was observed when at least one single feature differentiated all objects (e.g. all objects had different colors or different identities), no benefit was observed when a combination of features differentiated the objects. These results suggest that object features are not properly conjoined during attentive tracking. The authors argue binding may be particularly difficult for moving objects because

motion causes the basis for binding – shared location – to be an unstable object property.

Mitroff, Arita, & Fleck explore the rules that determine whether visual entities are bound into a singular representation that is maintained and updated across object motion. In their first experiment observers saw two squares, defined by a physical contour or an illusory contour. A letter was briefly presented in each square, and then both squares moved to new locations. Following this motion, a single probe letter was displayed in one of the squares and observers judged whether or not it matched either of the preview letters. When squares were defined by physical contours, these judgements were faster if the probe letter reappeared in the same square in which it was previewed, suggesting that the letters were bound to specific squares and that this binding persisted across the object's motion. Illusory contoured objects also generated this object-specific preview benefit but only if all object contours were illusory. If physical and illusory contoured entities were intermixed, the preview benefit was eliminated for the illusory stimuli. In a second experiment, the authors presented preview letters either within or adjacent to each square. Containment always led to an object-specific preview benefit, while adjacently presented letters could remain bound to objects only if all letters were presented outside the objects. From these results the authors propose that "boundedness" and "containment" are important principles used by the visual system to determine what entities should be bound into a single representation, and whether these bindings should be maintained when one of these entities is set into motion. Importantly, however, the observed contextual influences demonstrate for the first time that the binding of object features and properties are sensitive to more than just the physical properties contained within any given visual display.

In the third paper of this set, **Saiki** investigated the degree to which memory for an object's features can efficiently substitute for missing perceptual information. In this study, a set of elongated colored bars moved horizontally until they became occluded behind squares. The remerging portion of each bar was either the same or a different color as it had been prior to occlusion. In one condition the head of the bar remerged before the tail was occluded so that the color change was perceptible. In another condition the occlusion was complete so that detection of the color change required memory. With only one bar in the display, observers performed both tasks equal well, however, when presented with two or more bars, performance in the memory condition dropped quickly, even though the number of presented objects was well within the capacity of working memory. Thus, when only one color was relevant so that binding was not necessary, memory information could be used to support task performance. Once multiple objects had to be monitored so that the binding of colors to objects was required, the use of working memory to find the non-matching object became less efficient. The results suggest that although multiple bound object representations may be maintained in working memory, very few (perhaps only 1) of these traces can efficiently function as a replacement for missing perceptual information.

Finally, **Holcombe** considers the problem of temporal binding in which features that co-occur in time must be linked. Specifically, he addresses the mechanisms used to bind the features of a single object as it changes over time. Holcombe used displays where objects rapidly alternated across two feature dimensions, color and motion direction. For example, an object might turn green at the same time as its internal texture moves to the left, and then turn red when its texture begins to move to the right. Correctly binding the color and its associated motion direction requires that the two features be marked as occurring at the same time, despite potential differences in processing latency. However, this marking does not coincide with precise temporal pairing of the features, as binding is best when changes in motion direction actually precede changes in color. The paper presents a series of studies suggesting that the visual system samples the features at a temporal marker caused by transient changes in the features, and argues that the color-motion asymmetry is caused by differences in

the way that this transient change of color and motion direction unfold over time.

The special issue closes with two papers that extend consideration of binding from features within single objects to the relations among multiple objects. First, **Ryan and Villate** examined the construction of, and subsequent access to, representations regarding the relative spatial and temporal relations among serially presented objects. Participants were sequentially shown three objects in separate spatial locations. Subsequently, a test display revealed all three objects simultaneously and participants judged whether the relative relations among the objects were maintained. The authors used observers' eye movements as their primary dependent variable. During the presentation phase, observers' eye movements transitioned between the location of the presented object and the locations that previously housed the prior study images, revealing the implicit generation and rehearsal of relational bindings in memory. During the test phase, eye movements distinguished intact displays from those in which the relations had been altered. For example, eye movements were directed towards the locations that would have contained an object had the relative spatial relations been maintained. Furthermore, objects in unaltered test displays tended to be fixated in the same temporal order in which they had been studied, an effect not observed with manipulated test images. These results demonstrate that visual representations include information regarding the relative spatial and temporal relations among objects and that eye movements may be one means by which information is bound into a lasting representation, and by which current information is compared to stored representations.

Hollingworth extends consideration of relational binding to real-world scenes and asks whether episodic scene memory includes information regarding the binding of objects to the scene's layout. Observers saw an image of a scene followed by a target probe object which was displayed in isolation at the center of the display. The scene was presented again and observers were to search for a probe object and indicate whether its orientation in the scene was the same as depicted in the probe display, or whether it was mirror reversed. Compared to a baseline condition in which no preview was given, both preview conditions (target-absent and target-present) resulted in more efficient search, indicating that memory for general context and layout facilitates search. Critically, Hollingworth also varied whether or not the target object was in the original scene. Search was more efficient if the preview scene contained the target than if it did not, demonstrating that scene memory representations that are functional in guiding search further preserve the specific binding of objects in a scene to their locations. Memory for context and memory for object-location bindings are at least partially independent.

In closing, we would like to express our thanks to all of the contributors for choosing to publish their fine work in this special issue of *Visual Cognition*. We would also like to extend our thanks to all those who so graciously provided expert reviews and commentaries on these articles. Because of their efforts, this special issue is a collection of interesting cutting-edge research articles that blends a variety of experimental approaches and theoretical points of view that we hope will stimulate the field for years to come.

Contents of the Special Issue

(In order of appearance)

Oakes, L. M., Messenger, I. M., Ross-Sheehy, S., & Luck, S. J. (in press). New evidence for rapid development of color-location binding in infants' visual short-term memory. *Visual Cognition*.

Allen, R. J., Hitch, G. J., & Baddeley, A. D. (in press). Cross-modal binding and working memory. *Visual Cognition*.

vanRullen, R. (in press). Binding hardwired vs. on-demand feature conjunctions. *Visual Cognition*.

Hommel, B., & Colzato, L. S. (in press). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. *Visual Cognition*.

Alvarez, G. A., & Thompson, T. W. (in press). Overwriting and rebinding: Why feature-switch detection tasks underestimate the binding capacity of visual working memory. *Visual Cognition*.

Logie, R. H., Brockmole, J. R., & Vandenbroucke, A. R. E. (in press). Bound feature combinations are fragile in visual short-term memory but form the basis for long-term learning. *Visual Cognition*.

Makovski, T., & Jiang, Y. V. (in press). Feature binding in attentive tracking of distinct objects. *Visual Cognition*.

Mitroff, S. R., Arita, J. T., & Fleck, M. S. (in press). Staying in bounds: Contextual constraints on object file coherence. *Visual Cognition*.

Saiki, J. (in press). Functional roles of memory for feature-location binding in event perception: Investigation with spatiotemporal visual search. *Visual Cognition*.

Holcombe, A. O. (in press). Temporal binding favors the early phase of color changes, but not of motion changes, yielding the color-motion asynchrony illusion. *Visual Cognition*.

Ryan, J. D., & Villate, C. (in press). Building visual representations: The binding of relative spatial relations across time. *Visual Cognition*.

Hollingworth, A. (in press). Two forms of scene memory guide visual search: Memory for general scene context and memory for the binding of target object to scene location. *Visual Cognition*.