What do catastrophic visual binding failures look like?

Vincent A. Billock\textsuperscript{1} and Brian H. Tsou\textsuperscript{2}

\textsuperscript{1}General Dynamics, Suite 200, 5200 Springfield Pike, Dayton, OH 45431, USA
\textsuperscript{2}US Air Force Research Laboratory, Wright-Patterson Air Force Base, OH 45433, USA

Ordinary vision is considered a binding success: all the pieces and aspects of an image are bound together, despite being processed by many different neurons in several different cortical areas. How this is accomplished is a key problem in visual neuroscience. The study of visual binding might be facilitated if we had ways to induce binding failures. A particularly interesting failure would involve a loss of the physical integrity of the image. Here, we identify conditions that induce such perceptual failures (e.g. the melting together of equiluminant colored images and the fragmentation of retinally stabilized images) and we suggest that these should studied using electrophysiological measures of binding.

In visual cortex, many neurons are specialized for detecting different visual attributes such as form, color, motion and depth. Often these cells are physically separated – in some cases, they seem to aggregate by stimulus preference in different cortical areas or at different spots of the same cortical area. So how does the visual system combine these responses to create unified and veridical percepts? A leading theory suggests that the stimulus-based responses of widely spaced neurons can be bound together by the temporal synchronization of their gamma band (30–90 Hz) activity. Although controversial, an impressive body of theoretical and electrophysiological evidence for this theory has accumulated [1–12]. Some of the most compelling binding studies compare perceptual and electrophysiological measures in humans [10–12]. To further such efforts, it is necessary to understand what perceptually constitutes a binding success or failure, so that its electrophysiological correlates can be sought. There are two especially interesting aspects of the 'binding problem': feature-binding and part-binding. Feature-binding correctly associates various features of images together, for example so that we see a yellow school bus moving towards us, rather than yellowness, curves and angles, and movement, as in a dynamic abstract painting. Part-binding refers to correct construction of spatially coherent percepts, which need to be extracted (segmented) from the noisy and often ambiguous retinal image. To study feature-binding, Treisman introduced the notion of feature-binding failures called illusory conjunctions – under some circumstances, especially in peripheral vision, two features such as shape and color can be incorrectly associated with one another – and this has been very useful for psychophysical studies of feature-binding [1]. However, there have been no comparable studies of part-binding failures (except in brain lesions [8]). For electrophysiological studies it would be useful to be able to produce part-binding failures in normal humans using reversible and noninvasive methods. So, what would a really spectacular part-binding failure look like? Has anyone ever seen one? Do they occur in normal (brain-undamaged) observers? In addressing these questions, we realized that we have seen candidates for such binding failures.

Recently, we studied the perception of retinally stabilized colored images [13]. We coupled a motorized mirror to an optical eye-tracker (which measures eye movements using infrared reflections from the cornea and lens). Subjects view the stimulus as a reflection in the mirror. By deflecting the mirror to compensate for eye movements, the reflected image remains stationary on the retina [13,14]. The textbook account of these stabilized images is that they fade (because of transient temporal responses of some retinal mechanisms), and fade they do. However, they also become unstable and lose their spatial coherence: their internal structure can flow about like figures made from melting wax, or their internal borders can fade away altogether so that the colors of the image diffuse into one another. These and similar bizarre phenomena are not explicable in terms of peripheral neural mechanisms. But what central mechanism could exhibit such behaviors? It occurred to us that some of the behaviors that we (and many earlier investigators of equiluminant and stabilized images) found were good examples of part-binding failures. Because the vast majority of these studies were made in an era innocent of binding theory, we have a rare opportunity to harness several decades of misunderstood research in the goal of advancing the current research in binding.

\textbf{Expected failure modes of part-binding}

There is an extensive literature of algorithms for part-binding. In general, these models need to bind things that belong together and segment them from backgrounds and other objects. At this level, one could imagine binding failures that result from an inability to link parts into a whole (such a binding failure will be discussed in a later section of this article). Additional possibilities can be
derived from segmentation models that rely on balancing cooperation and competition [15–18]. Competitive networks define the locations of borders and edges; cooperative networks reinforce mechanisms that agree about the presence of a border, and fill-in areas between borders. If the balance between these mechanisms breaks down, part-binding should fail in distinctive ways. If cooperative mechanisms are too strong, then the network will settle on a common value for all points in an image. If this occurs slowly enough in an already segmented image, the image would appear to fade, or gray-out, or melt. Conversely, if competition is too strong or cooperation is too weak, the image should appear to fragment. This occurs for some neural networks that use nearest neighbor connectivity [17,18].

Binding-like failures in retinal stabilization: fading and blackout

Images can be mechanically stabilized from eye movements, either by mechanical coupling to a contact lens, or via an eye-tracker [14,19]. When an image is stabilized, the most common perceptual effect is a gradual fading of the image. Similar effects can be obtained using steady fixation, afterimages or ganzfelds. If the image starts out as a rather textureless surface (a ganzfeld), then gradually all aspects of the percept are lost and the subject perceives a uniform, depthless gray (eigengrau) even if the ganzfeld was initially colored. Fading is usually attributed to the transient properties of certain retinogeniculate units – the transient M cells. This explanation is inadequate on at least two grounds: (i) ~80% of retinogeniculate units belong to the P cell pathway, which has a sustained response to both achromatic form and color stimuli [20], and (ii) most visual mechanisms (including achromatic form, color and depth) receive inputs from P cells and can be driven by P cells alone, if M cells are silenced [21]. And this information is utilized by higher mechanisms; for example, the depth and multistability of a Necker cube are just as salient when stabilized, so long as it remains visible [22].

Although P cells signal the attributes of stabilized images to cortex (a necessary condition for cortical processing), this is not always a sufficient condition for veridical perception. Consider some properties of stabilized images, paradoxical to P cell function, implicating central mechanisms in stabilized-image fading. For example, P cells have sustained responses for high spatial frequencies and high contrasts (unlike M cells, which become more transient at high contrasts [23]); yet Purkinje’s tree (the high-contrast pattern of blood vessel shadows on the retina) fades unusually rapidly, and fades fastest (as fast as 80 ms) in the parafovea, where P cells are present in high densities [24]. Similarly, although P cells have strong sustained responses to chromatic stimuli, stabilized color images fade faster and more completely than achromatic stimuli [25,26].

Actually, many attributes of visual fading are more consistent with central than peripheral mechanisms. For example, the fading of a stabilized image in one eye can be slowed or reversed by modulating the stimulation of the other eye [27]. Conversely, fading of a stabilized image in one eye makes images viewed by the other eye less visible. Remarkably, even if the stabilized image in one eye has completely disappeared, it can still be binocularly fused with a duplicate moving image in the other eye to yield a paradoxical percept of motion in depth [14]. Finally, fading of images can be slowed or reversed by stimulating attention or stimulating the other senses [19]. Blackout is a related, albeit more dramatic effect. Sometimes image stabilization, even if only in one eye, induces an extremely abrupt, binocularly simultaneous visual blackout. Recovery can be initiated by a blink. Descriptions of this ‘black field’ as ‘more black than black’ do not do it justice [19]. Our subjects describe it as ‘like someone cut my optic nerves, and I don’t have a visual system anymore’ [15]. Ditchburn attributes the black-field effect to a fading mechanism combined with a cortical inhibitory feedback loop that runs out of control [19].

Binding-like failures in retinal stabilization: image fragmentation

From a binding failure perspective, the most interesting retinal stabilization effect is fragmentation [28–30] (Figures 1,2). In general, complex patterns are more likely to fragment than simple ones, angular patterns fragment more than round ones, and parallel lines tend to disappear and reappear together [31,32]. Sometimes the fragmentation takes the form of a series of breaks in the image (Figure 1), whereas at other times, entire portions of the figure disappear and reappear (Figure 2). The probability of fragmentation versus unitary fading grows with line length and reaches equal probability for lines subtending 45°, suggesting a range for the neural processes involved [32]. Interestingly, if a stabilized image with gaps in its structure fades and then reappears, the gaps in the image are often filled-in [19]. Moreover, filling-in can occur even as the overall image is fading [33]. When images fragment, they tend to fade and revive in clusters that are cognitively meaningful, and follow Gestalt-like rules of association (Figure 2). These processes act on the central image; fragmentation can eliminate contours present only in a binocular image and not in its retinal precursors [34]. Similar fragmentation can occur during migraine attacks [35] and in strabismic amblyopia (a cortical visual defect induced by poor coordination of the eyes during visual development), especially for high spatial frequency stimuli [36] (Figure 1). The connection of fragmentation phenomena to amblyopia is especially interesting because, in cat, cortical cells driven by the amblyopic eye are poorly synchronized compared with cell populations driven by the normal eye – especially for high spatial frequency stimuli [37]. No studies of synchronization during stabilization have been made. However, some early electroencephalogram (EEG) studies report that alpha-rhythm (9–10 Hz) power increased ~0.7–1.0 s before disappearances; higher frequencies (including the gamma band, which is now implicated in binding) were suppressed until ~1 s before spontaneous reappearances [38,39].

Oddities of equiluminant images

Like stabilized images, pictures in which hues have equal luminance are less perceptually salient than natural
images. Luminance differs slightly from brightness and is usually defined operationally; for example, if two colors are shown in the same place in rapid sequence (a procedure known as heterochromatic flicker photometry), the sensation of flicker is drastically reduced for a certain ratio of the radiances of the colors – this is defined as equiluminance. Slightly harder to use in practice, but of more relevance to this discussion, is the minimum border criterion: at equiluminance, the apparent border contrast between differently colored image parts is minimized. Similarly, perception of form, depth, motion and color are all degraded for equiluminant stimuli [40–42]. Sometimes these phenomena are attributed to a lack of chromatic input to these mechanisms, but several studies show that these systems can be driven by P cells in general and by hue-specific mechanisms in particular [43]. Indeed, even some putatively chromatic functions (e.g. chromatic rivalries) are disrupted by equiluminance [41]. Others have attributed some effects to the lower acuity of chromatic mechanisms. This can explain some effects [42] but the most interesting effects are out of proportion to the acuity loss [43], or are seemingly unrelated to it [40,41]. From a binding point of view, the most significant of these are linking failures and image melting.

**Binding-like failures at equiluminance: linking failures and perceptual melting**

A variety of symptoms suggest that there are part-binding failures in equiluminance. Included in these are failures to link pieces of a surface. Consider random dot stereopsis. It is possible to build a stereogram in which the two images contain only random dots. The two images are near duplicates; one image has a group of dots shifted over several pixels. Binocular fusion of the two images yields a percept of a group of random dots floating above or below the random dot background. In an ordinary random dot
Cross-modal reinforcement: what is lost in stabilized and equiluminant images?

We have discussed how a balance of cooperation and competition is required to produce reliable segmentation networks, and how these networks sometimes fail, especially for nearest-neighbor connectivity [17,18]. One way to avoid this is to allow the activity of a second network to reinforce the first – in effect, the second network can act as a second set of nearest neighbors. In the segmentation literature, this reinforcement between different sensory modalities – called ‘cross-modal construction’ [45] – improves the precision and likelihood of both correct segmentation and feature binding [45–47]. If part-binding mechanisms evolved to exploit this reinforcement, then its loss could compromise part-binding. The evidence described in this article is consistent with reliance on cross-modal reinforcement of color segmentation mechanisms by achromatic form mechanisms, and reliance on reinforcement of achromatic form segmentation systems by mechanisms sensitive to motion or temporal modulation.

It takes two: combining equiluminance and stabilization makes everything worse

If this cross-modal construction hypothesis is correct, then we would expect that compromising more feature systems would further worsen part-binding. This seems to be the case. A familiar example is the minimum border effect – the weakened border formed by an equiluminant bipartite field can disappear completely when eye fixation is steady [44]. A more dramatic example is our own experiment using stabilized opponent colors [13]; unlike the equiluminant experiments already discussed, binding failures occur for any color combination. Perhaps the most established law in experimental psychology is that certain color combinations such as reddish green and bluish yellow are never allowed; this observation formed the basis for Hering’s theory of color opponency and much of modern color science. However, it had been reported that some subjects saw these forbidden colors when colored bipartite fields such as those shown in Figure 3 were stabilized [14]. When we attempted to replicate these results, we found that the effect depended on the combination of equiluminance and stabilization. When we stabilized non-equiluminant red–green or blue–yellow borders, we saw a variety of part-binding failures, which included effects where the border would disappear in places; the unbound features or edges fade away, they are sometimes replaced with an illusory achromatic border [40,41]).

Figure 2. Some gestalt-like attributes of image fragmentation in retinal stabilization that implicate central mechanisms. [a–d] Fragmentation of monocular stabilized images. Reproduced, with permission, from Ref. [29]. Pictures on the left show unstabilized images and pictures to the right show stabilized percepts. Note that images tend to fragment and fade, such that the visible fragments follow rules. In (a) and (b), meaningful fragments are preserved. In (c), parallel fragments are preserved. In (d), gestalt grouping rules are mostly followed. (e) Effects of stabilizing the monocular parts (left-eye image and right-eye image) of a binocular (fused) percept. Note that disappearance of image fragments is characteristic of the fused image, not the monocular forms, ruling out peripheral explanations. Reproduced, with permission, from Ref. [34].

stereogram, when a group of dots is segmented from the background, the pattern of dots defines a textured surface of common depth. Near equiluminance, depth is seen for the individual colored pixels but, unlike ordinary random dot stereopsis, the pixels seem to be separate rather than part of the same textured surface [41]. This is not a curiosity of stereopsis: 2D surfaces are not linked together when defined only by hue. An image composed of equiluminous patches does not appear unified but, rather, appears as patches of one color floating on the other [41,42]. In other cases, all sense of object and surface can be lost and the target is perceived as a ‘jumble of lines’ [41] or ‘extremely confusing and hard to describe’ [40]. Moreover, the contrast of equiluminant images can seem unstable – Gregory [40] describes such images as ‘jazzy’. In 1927, Liebmann (translated in Ref. [42]) reported that there is a ‘critical zone [where] everything flows…glimmers…most everything is soft, jelly-like, colloidal. Often…parts which belong together in the normal figure now have nothing to do with one another. [It is] a world without firm things, without solidity.’ Such effects are most severe when the borders are defined by tritan (S-cone) modulation: the borders of the images can appear to melt like hot wax and distort like pulled taffy (reminiscent of Liebmann’s comment), and the colors diffuse into one another in the same way as spilled inks, blending into continuous color gradients [13,44]. There are interesting local boundary effects: fading is sometimes localized to particular features or regions [40] and when equiluminant features or edges fade away, they are sometimes replaced with an illusory achromatic border [40,41]).
Electrophysiological testing of part-binding failures

We argue that phenomena such as border disappearance, fragmentation, diffusive color spreading, and linking failures of equiluminous and stabilized images are examples of part-binding failures, and that these particular failure conditions are clues that part-binding relies on achromatic mechanisms reinforcing segmentation by color, and on motion-driven mechanisms reinforcing segmentation by achromatic form. Testing these ideas will require some notion of what the electrophysiological correlates of binding are and how they could be detected non-invasively in intact responding humans. Because gamma-band activity is implicated in binding processes, the loss of high-frequency EEG power that accompanies stabilized-image fragmentation [38,39] is suggestive but not definitive. One potential problem is that some sources of gamma-band power in EEGs might be unrelated to binding. Another potential problem is that our binding failures might actually be incorrect bindings and therefore not distinguishable by EEG. Recent experiments have cast light on both questions, by using fragmented and camouflaged images in which subjects could eventually learn to see coherent embedded objects [10,12]. (These studies are especially relevant here because the fragmentation-to-order shift is the complement to the order-to-fragmentation phenomena we studied.) There was gamma-band activity whether subjects saw the coherent picture or not, but there were two differences in EEG when gestalt perception was obtained: the temporal output of widely spaced electrodes became more correlated (consistent with synchronization of widely spaced neural mechanisms [12]) and a second kind of gamma-band response, the ‘induced gamma response’, was recorded. On every physically identical stimulus presentation there is a gamma-band response that is phase-locked to the camouflaged or fragmented stimulus, but only on trials where the gestalt percept is seen does a second signature emerge that consists of gamma-band bursts in variable phase to the actual stimulus [10,12]. We therefore predict that EEG signatures of binding, such as the induced gamma response, will be reduced during binding-like failures induced by retinal stabilization or equiluminance, and that these signatures will be further minimized or eliminated during the more dramatic perceptual effects created by combining stabilization and equiluminance [13]. A sophisticated version of this experiment could be done using the methods of Billock et al. [13], but the simpler methods of Buck et al. [44] might well suffice.

Acknowledgements

We thank Kenneth Blum, Scott Kelso, Oliver Sacks, Wolf Singer and Anne Treisman for particularly helpful suggestions and remarks.

References

25 Weintraub, D.J. (1964) Successive contrast involving luminance and purity alterations of the Ganzfeld. J. Exp. Psychol. 68, 555–562
33 Cardu, B. et al. (1971) The influence of peripheral and central factors on the way that stabilized images disappeared. Vision Res. 11, 1337–1343