

How Attention Can Alter Appearances

Peter U. Tse, Eric A. Reavis, Peter J. Kohler,
Gideon P. Caplovitz, and Thalia Wheatley

The Class of Attentional Phenomena in Need of an Explanation

It has long been known that attention can change how things appear (e.g., Treisman, 2006). Traditional models of attention that focus on modulation of neuronal gain control, feature binding, or object tracking cannot easily account for the fact that attention can be voluntarily allocated in a way that markedly alters perceived features, including color, size, transparency, 3D shape/layout, orientation, and motion. Before delving into theory, we think it best that readers experience for themselves the problem that we feel calls for an explanation. To that end, we begin by presenting a broad range of examples where attention can change how a single input is experienced. Our goal is to present a theory of attention that can account for all of the effects summarized in the figures, as well as some new ones involving perceived colors in cases of overlapping transparent surfaces. Here, we suggest that attention alters perceived appearances (i.e., features or qualia) by defining the domain of automatic operations in the preconscious buffer—a window of time just prior to conscious experience.

Gestalt psychologists (e.g., Koffka, 1935) noted that percepts can be grouped in various ways, and that this can lead to changes in subjective appearance. In addition, which grouping is experienced is to some extent under endogenous or volitional control. For example, one can see an “X” in Figure 12.1A or a “square” or any number of other groupings, depending on how one attends to the component ellipses. Since attention can be shifted endogenously, the groupings one experiences as a result of a change in the specification of figure-versus-ground relationships can also change endogenously. In Rubin’s well-known (1915/1958) face/vase example, a version of which is shown in Figure 12.1B, one can see either a vase or two profiles facing one another. Other examples in the literature suggest that attention does more than simply flip figure and ground. Attentional “flipping” can induce changes in

Handbook of Experimental Phenomenology: Visual Perception of Shape, Space and Appearance,
First Edition. Edited by Liliana Albertazzi.

© 2013 John Wiley & Sons, Ltd. Published 2013 by John Wiley & Sons, Ltd.

EI

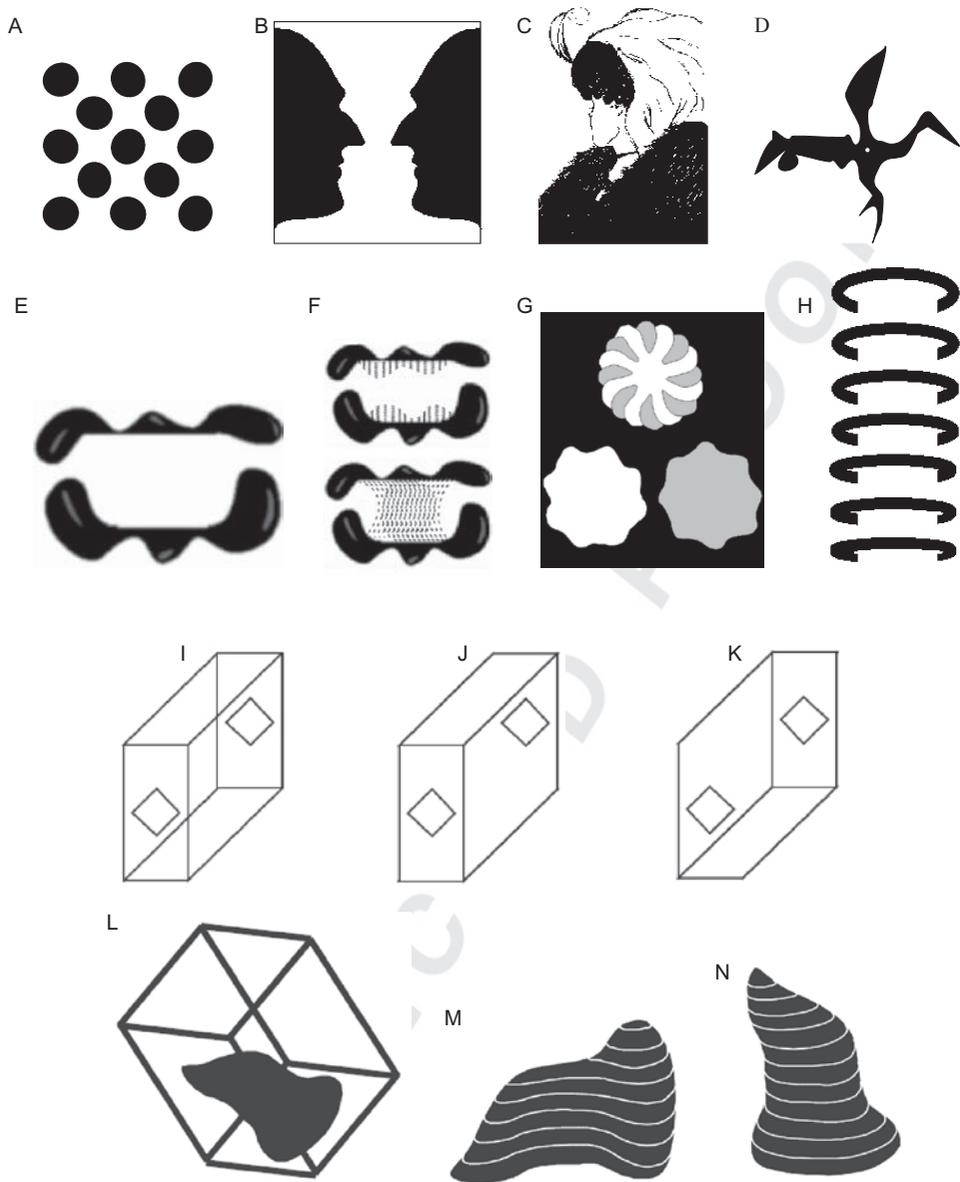


Figure 12.1 (A–D) Attention can modulate perceived groupings as in A, perceived figure–ground relations as in B, and perceived shape and meaning, as C, which can either be seen as a young woman facing away or an old woman looking down toward the left, or D where the perceived “head” can be placed at any tip of the figure at will, changing its perceived bodily structure. (E–H) Endogenous attention can modulate perceived amodal completion and shape. For example, E can amodally complete as either arrangement shown in F. Attention can also modulate perceived amodal completion in shape, as in G, which can be seen either as a white octopus hugging a gray rock from the front, or a gray octopus hugging a white rock from behind. Attention can even modulate perceived illusory contours and modal completion as in H. (I–N) Attention can modulate perceived 3D orientation as in I, which can be seen to have the perceived spatial arrangements of J or K. Note that the “Post-Its” have different perceived 3D orientations in the two possible arrangements. Similarly, the silhouette in L can be perceived to have either the 3D shape shown in M or N, depending on which Necker cube arrangement is perceived.

perceived meaning, 3D shape, size, and position. As long ago as the 19th century, drawings that could be seen in one way and then another were common. For example, in the drawing shown in Figure 12.1C (Hill, 1915), one can see a young woman looking away or an old woman facing forward (the left ear of the young woman becomes the left eye of the old woman). In Figure 12.1D, one can see at least four different bizarre animals or birds simply by willing one of the four ends to be the head.

In the past two decades, researchers have determined a number of perceptual processes that are modulated by attention. For example, attention can modulate amodal completion: the process by which perception “completes” occluded objects. One can attend to Figure 12.1E, for example, to see either of the physical arrangements shown in Figure 12.1F, suggesting that attention can modulate the outputs of the operations that underlie amodal completion of one object behind another. Similarly, one can view Figure 12.1G as a white octopus hugging a gray rock or a gray octopus hugging a white rock. Attention can modulate the outputs of modal completion as well, which is the perceptual completion of an occluding object; one can see Figure 12.1H either as black rings wrapping around a white column or as a stack of floating bracelets with their gaps facing forward. This example is remarkable because one sees illusory contours only under the first interpretation, implying that attention can even influence the formation of illusory contours (Tse, 1998, 1999a, 1999b; Tse & Albert, 1998).

Attention can also influence perceived 3D orientation (Albert & Tse, 2000) as in Figure 12.1I. Depending on how one sees the underlying Necker block, one will see the diamonds as “Post-Its” attached in the orientation of either Figure 12.1J or Figure 12.1K. And attention can influence perceived 3D shape (Tse, 2002); for example, depending on the 3D orientation of the underlying Necker cube in Figure 12.1L, one can see the 3D shape shown in Figure 12.1M or Figure 12.1N.

Attention can also influence perceived size. In Figure 12.2A, which is a modified version of the Ponzo illusion, if just one of the two sets of traintracks is attended, the “farther” red disk appears larger.

Attention can even modulate perceived transparency. In Figure 12.2B, if the two objects are attentionally attached to one versus the other of the two surfaces, one will be seen as a transparent dome and the other as an opaque bowl. When they are seen as attached now to the other surface, what was a transparent dome is now an opaque bowl, or vice versa.

Attention can even influence what scene is perceived. For example, Figure 12.2C shows an ambiguous image that can be seen as a city scene if one attends to high spatial frequencies, or as a bedroom scene, if low spatial frequencies are attended.

The influence of attention on perception is not limited to static objects, but extends to the domain of motion as well. For example, attention can modulate perceived position (Tse, Whitney, Anstis, & Cavanagh, 2011), as depicted in Figures 12.3A–D. When a pair of targets is flashed on top of two superimposed textures rotating in opposite directions, the perceived locations of the targets are shifted depending on which of the two directions of motion is attended. Because the stimulus remains unchanged as attention switches from one moving layer to the other, the effect cannot be stimulus-driven.

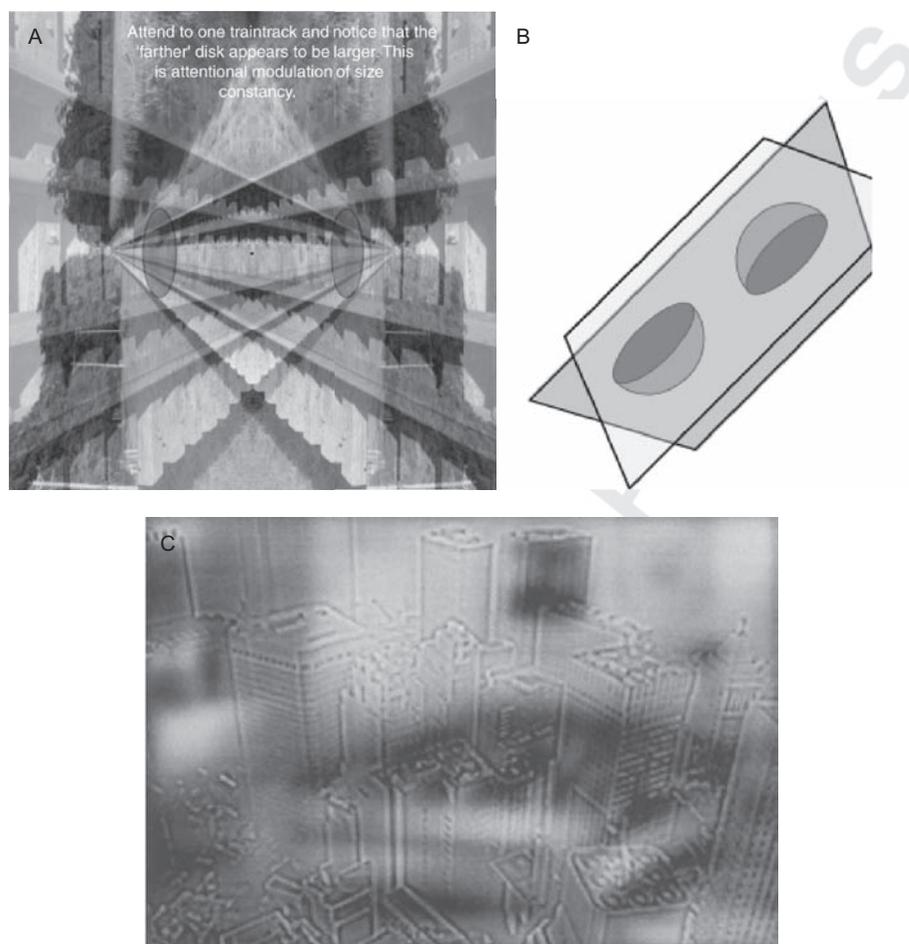


Figure 12.2 A: Attention can modulate perceived size. Depending on which set of train tracks is attended, the “farther” red disk appears larger. B: Attention can modulate perceived transparency. Depending on which of the two surfaces the two objects are perceived to be attached to, they will be seen as either transparent dome or opaque bowl. After Albert and Tse (2000). C: Attention can modulate the perceived scene. Attending to high vs. low spatial frequencies alters the scene that is experienced. From Schyns and Oliva (1994).

Another example of attention-induced changes in perceived motion can be found in the well-known “quartet” apparent motion stimulus (e.g., Kohler, Haddad, Singer, & Muckli, 2008). Two dots appear at diagonally opposite corners of a virtual square, disappear, and then reappear at the other two corner locations. This can be perceived as either horizontal or vertical apparent motion. Attention can bias which direction of motion is perceived.

In another instance (Caplovitz & Tse, 2006), a single moving object,² like that shown in Figure 12.3E, can be seen in at least four different ways, namely, as a cross-shaped figure nonrigidly morphing in size and shape (12.3H), an ellipse (12.3F)

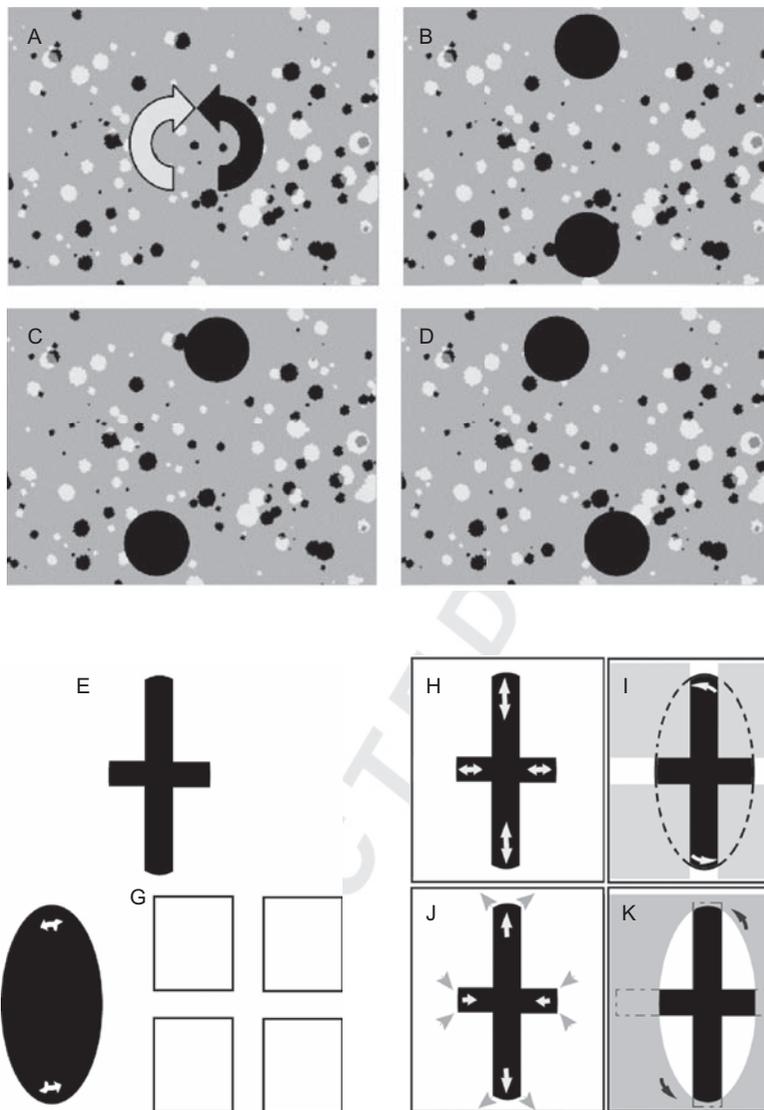


Figure 12.3 A: Arrows (not present in the actual stimulus) indicate that the transparent layer composed of black dots rotated in one direction, while the transparent layer composed of white dots rotated in the opposite direction. Rotation direction reversed for both transparent layers simultaneously every 1200 ms. B: Vertically aligned red targets appeared for 50 ms at the moment when the direction of rotation reversed. C: When the white layer was attended, the targets appeared slanted to the right. D: When the black layer was attended, the targets appeared slanted to the left.¹ (E–K) Attention can modulate perceived motion such that different types of moving objects are experienced. See text for details.

rigidly rotating behind four square occluders (12.3G) as shown in (12.3I), two independent perpendicular bars rigidly oscillating in depth (12.3J), or a stationary cross viewed through a rigidly rotating elliptical aperture (12.3K). Different possible perceptual organizations involve perception of distinct motion, shape, and depth features.

There are many more examples of static and motion-defined multistable figures, but these examples suffice to demonstrate that there is a deep problem in need of an explanation. How does attention modulate our experience of the stimulus in all of these examples? At present, there is no unifying theory that can explain the effect of attention on perception in all the cases depicted above.

Current Theories of Attention

In this section we review current models of attention and how they serve three broad classes of explananda. One class of models seeks to account for the role of attention in changing stimulus sensitivity (i.e., more precisely, as changes in response or gain control; see Desimone & Duncan, 1995; Hillyard, Vogel, & Luck, 1998; Liu, Abrams, & Carrasco, 2009; Reynolds & Heeger, 2009; Tse, Sheinberg, & Logothetis, 2002). A second class of models focuses on the role that attention plays in binding stimulus features (Allen, Baddeley, & Hitch, 2006; Ashby, Prinzmetal, Ivry, & Maddox, 1996; Engel, Fries, König, Brecht, & Singer, 1999; Oakes, Ross-Sheehy, & Luck, 2006; Reynolds & Desimone, 1999; Shipp, Adams, Moutoussis, & Zeki, 2009; Treisman, 1988, 1996; Tse, 2006b). A third set of models of attention aims to explain how an object can be bound (tracked) over time as it moves through some representational space (e.g., Kahneman, Treisman, & Gibbs, 1992; but see Pylyshyn & Storm, 1998).

Gain control, feature binding, and tracking are not mutually exclusive processes. Rather, attention involves aspects of all three. Indeed, attention is not just a single process. Rather, processes that select and deselect among inputs are likely realized via multiple mechanisms. Attention has low-level aspects that can be described as changes in sensitivity to inputs (i.e., more precisely, changes in response or gain control; see Desimone & Duncan, 1995; Hillyard et al., 1998; Liu et al., 2009; Reynolds & Heeger, 2009). Attention also has higher-level aspects, such as binding of features (Allen et al., 2006; Ashby et al., 1996; Engel et al., 1999; Oakes et al., 2006; Reynolds & Desimone, 1999; Shipp et al., 2009; Treisman, 1988, 1996), binding of operations and operands in working memory (Tse, 2006b), and tracking of a figure over time (Kahneman et al., 1992; Pylyshyn & Storm, 1998).

One dominant model of how a figure can be tracked over time even in the presence of moving distractors is “object file theory.” An object file (Kahneman et al., 1992; Tse, 2006b) is a metaphor for attentional processes that combine multiple features existing over various modalities into a common bound representation of an object. An object file is an attentionally tracked “figure” (Lamy & Tsai, 2000) integrated as a temporary episodic representation in a working memory buffer (Kahneman et al., 1992; Schneider, 1999) that maintains a coherent identity even as

the particular contents defining that identity change over time. While there are disagreements concerning what is actually involved in tracking an object (Carey & Xu, 2001; Pylyshyn & Storm, 1988; Scholl, Pylyshyn, & Feldman, 2001), theorists agree that there must be some psychological entity that keeps track of an object or figure over time, within some representational space. For example, one can listen to a symphony and track the oboe. One can then listen to the same symphony again and this time track the lead violin. In both cases the sensory input is the same. What differs is the nature of the object file one constructs.

The contents of an object file are thought to be mid to high level. That is, there is widely thought to be a preconscious stage of representation that cannot be attended and whose contents cannot be added to an object file (Wolfe, 2003; Treisman & Gelade, 1980). These are presumably the representations that are output by a preconscious stage that permits unconscious operations that lead to the mid-level representations to which attention has access. Such mid-level representations might include the results of grouping operations (Enns & Rensink, 1991; Rensink & Enns, 1995), surface completion operations (He & Nakayama, 1992; Rensink & Enns, 1998), and the outputs of processes that compute color, shape, and size constancy in order to recover the intrinsic properties of objects. Possible object file contents can thus be perceptual features that have been preprocessed to a level to which attention has access. These may be features on feature maps (Quinlan, 2003; Treisman, 1992), mid-level structures such as surfaces (He & Nakayama, 1992), abstract identity tags (Gordon & Irwin, 1996), or higher-level conceptual information (Gordon & Irwin, 2000).

But even if changes in sensitivity, gain control, binding, and tracking were perfectly accounted for by these existing models, attention-induced changes in perceived features would still need to be explained. Sensitivity modulation, feature binding, and object tracking simply cannot account for attentional alteration of high-level perceived attributes such as altered 3D shape or material. Conceiving of attention as a gain control mechanism (e.g., Reynolds & Heeger, 2009) predicts that changes in attentional allocation should at best alter the relative or absolute magnitudes, intensities or *degrees* of perceived features. The gain control model of attention cannot account for changes in the *kinds* of perceived features themselves. For example, when one sees one shape versus another in Figures 12.1B, C, D, H, K, or N, one shape is not a degree of the other. These appearances are mutually exclusive outcomes of processing that are consistent with the ambiguous input. They appear to require the application of operations like determining border ownership, generating three-dimensional surfaces, or in the case of Figure 12.1H, generating illusory contours or not. Alternatively, conceiving of attention as a “feature glue” that binds pre-consciously specified features (e.g., Treisman, 1988, 1996) cannot account for how attention can alter those features which are supposed to be determined pre-consciously. The features of a young woman (Figure 12.1C) cannot be rebound easily into the very different features of an old woman under Treisman’s feature-integration theory or its modern intellectual descendants (e.g., Wolfe, 2010; Wolfe, Vö, Evans, & Greene, 2011), because such features are the inputs to attentional operations, not their outputs.

The Preconscious Buffer

Visual experience is not of the image that is cast upon the retina. In the brief span of time between retinal activation and the world as experienced, numerous operations must transform visual input into a conscious output where key information has been completed and made explicit. The operations can be thought of as occurring within a preconscious buffer. Ambiguous or incomplete information must be appropriately completed so that it can be acted upon. Even though perception is constructed on the basis of input, it is presumably constructed to contain veridical information about the world, so that an animal can act optimally. What is made explicit at the stage of experienced objects is the kind of information that an animal can use to act in the world or plan future actions. This information includes intrinsic properties of objects (i.e., facts that would not change with viewpoint), including their shapes, pigments, materials, sizes, layout, motion directions, and motion magnitudes. For example, whereas only local levels and frequencies of light are measured at the retina, we experience brightnesses and hues that recover pigments and discount the global illuminant by comparing information non-locally. This operation is called “color constancy.” But other constancy operations must also operate in the brief duration between retinal activation and conscious experience. These operations include not only size and shape constancy, but also motion constancy (determining the true magnitude of motion in the world despite varying retinal motions introduced by, for example, perspectival projection) and material constancy (determining, for example, that something is rigid and hard, even when subject to the deformations imposed by 3D rotation, foreshortening, or the deformations introduced by intervening layers, such as glass or moving water). A great deal of work in psychophysics over decades has made apparent the complexity of many of these preconscious operations. We know that these operations must complete before visual experience, because what we visually experience is nothing like what is happening on the retina, but is instead a close approximation to what is the case in the world. For example, we do not experience a white sheet of paper at sunset as having a reddish pigment. We experience it as a white sheet of paper under reddish light. We do not experience people walking away from us as actually shrinking or a door deforming its shape as it opens.

Because of constancy operations, when we experience the reflectance of a surface, it seems to be automatically “delivered” to us: We simply see the grass to be green, to have a certain material, and to have particular shapes and distances. But none of these attributes are present in the visual input at the level of the retina. Instead, surface pigments must be computed from cues in the visual image. It therefore must be the case that these computations take place between the time that the image is activated and the time that we become conscious of high-level object and world attributes. Indeed, this can be subsumed under what Helmholtz (1866) referred to as “unconscious inferences.”

We must distinguish between “cognitive inferences” that follow perception and the “unconscious inferences” (Helmholtz, 1867/1910) that precede it. For example, when we infer that it must have rained because we see that the ground is wet, we are making a cognitive inference that follows perceptual experience in time.

However, when we see the ground as wet, we just experience it as wet. Since there is no wetness in the retinal image, seeing the ground as wet must itself be the result of operations and prior assumptions that allow the automatic and rapid construction of a visual experience of wetness, when certain criteria are met among visual inputs. Such unconscious inferences could be realized in a purely bottom-up manner, much like the simplest motion detectors (e.g., Hassenstein & Reichardt, 1956) respond to changes in luminance, and can be thought of as automatically “inferring” motion. Or the construction of visual experience of wetness and other intrinsic attributes of things and events in the world could arise from much more complex neural processes. Much human psychophysics has established beyond a doubt that such processes play a role in the construction of visual experience prior to the having of that experience.

The existence of a preconscious buffer can be shown in a number of ways. A compelling case is offered by apparent motion. When two non-overlapping static images of a luminance-defined spot are alternated in succession within a certain range of spatiotemporal offsets (Korte, 1915), they appear to comprise a single object jumping back and forth smoothly in apparent motion. Because no object actually moves, the appearance of continual motion must be a construction of the visual system. Moreover, because the position of the second spot cannot be known in advance, it must be the case that perceived motion is fully constructed only *after* the onset of the second image. Thus the apparent motion from position 1 to position 2 must be an instance where the visual system constructs the motion that “must have happened” prior to the appearance of the spot at position 2 (Beck, Elsner, & Silverstein, 1977; Choi & Scholl, 2006; Eagleman & Sejnowski, 2000; Tse, Nakayama, & Cavanagh, 1998; Tse & Logothetis, 2002). This implies that our perceptual experience is not of events as they are happening now, but of events as they happened in the recent past constructed on the basis of past and present input, as well as expectations of future input. This further implies that there must be a short-term perceptual buffer within which past and present inputs are integrated with predictive models, before a “commitment” is made to how immediately past events gave rise to present inputs. This perceptual buffer permits the influence of stages of form analysis (Tse, 2006a; Tse & Caplovitz, 2006), mid-level surface representations (He & Nakayama, 1994), and high-level expectations (Bar, 2003, 2004; Bar et al., 2006; Tse & Cavanagh, 2000) or internal models such as those that govern biological motion processing (e.g., Johansson, 1973; Shiffrar & Freyd, 1990, 1993; see also Jastorff & Orban, 2009; Neri, Morrone, & Burr, 1999). Tse and Logothetis (2002) demonstrated that this buffer compares form and motion inputs over the past 100 ms, and Eagleman and Sejnowski (2000) estimated a comparable 80 ms of comparison between past and present inputs, during which perceptual events are constructed. A consequence of this information-processing architecture is that the experience of events will lag their occurrence in the world by at least ~100 ms, and potentially by as much as a few hundred milliseconds. Of course, if this delay were too long, we could never make an appropriate decision about how best to act. We could not experience the world quickly enough to respond appropriately to a tennis ball, for example. By the time we experienced it, it would be long gone. The various preconscious operations that map visual inputs onto later visual experiences have therefore

evolved under selective pressure to be both rapid and sufficiently veridical (i.e., corresponding to events as they really are in the world) to permit the actions required for survival.

Computations made within the preconscious buffer can be influenced by top-down factors. There are different ways in which top-down factors operate. Ambiguous bottom-up inputs can be interpreted in light of the constraints imposed by knowledge of how things tend to move (e.g., Tse & Cavanagh, 2000). A particularly compelling example of this occurs in the case of biological apparent motion (Shiffrar & Freyd, 1990, 1993; see also Chatterjee, Freyd, & Shiffrar, 1996; Johansson, 1973; Neri *et al.*, 1999), where two static images of bodies are alternated, separated by a blank. Shiffrar and Freyd (1993) considered apparent motion over photographs of human bodies, and found that observers tended predominantly to see biologically plausible motion, such as a hand moving around a body, when the stimulus onset asynchrony (SOA) between photographs was higher than approximately 350 ms. Below this, observers tended to see biologically implausible motion, such as a hand passing through a body. For short SOAs, the biases of the low-level motion system presumably dominate, and these appear to be indifferent to considerations of biological plausibility. Instead, the low-level motion processing system matches on the basis of motion-energy (e.g., Hassenstein & Reichardt, 1956; van Santen & Sperling, 1984), spatial frequency (Green, 1986; Ramachandran, Ginsburg, & Anstis, 1983), and minimization of path distance (Ullman, 1979). For longer SOAs, implicit knowledge of how bodies move can influence the motion trajectories that are computed in the preconscious buffer, and that are then experienced as biologically possible motion. We cannot infer from this that the upper bound on the duration between retinal activation and conscious experience of a motion path is ~350 ms. It could be the case that it takes time to invoke the appropriate internal biological motion model, but that, once in place, the computation of motion path is faster than this. Although no one to date has been able to precisely determine the duration of time between retinal activation and visual experience, it would seem to require a minimum of ~80 ms and a maximum of a few hundred milliseconds.

Attention Influences Subsequent Preconscious Processes

Whereas preconscious operations are fast, automatic, mandatory, dedicated, impenetrable, encapsulated, and inflexible (Fodor, 1983, 2001; Gardner, 1993; Karmiloff-Smith, 1995), endogenous attention appears to operate over mid-level representations in a manner that has just the opposite attributes. However, there is evidence that endogenous attention can influence processes outside the domain of consciousness.

Certainly attention cannot reach down to the level of the earliest representations. There are not even axonal fibers that would permit feedback down to the level of the retina. However, there are axonal feedback connections from the cortex to the thalamus, and, in the case of vision, a strong case can be made that feedback connections from V1 and elsewhere to the lateral geniculate nucleus (LGN) of the thalamus may play a role in cortical modulation of subsequent cortical input (e.g., Saalmann & Kastner, 2009). For example, if one is looking for a red object, one could imagine that such cortical feedback would both turn up the “gain” on red

signals sent on to cortex for further processing, and perhaps also turn down the gain on irrelevant signals. Attention thus appears to be capable of operating not only over the contents of consciousness, but of operating on representations and processes within the preconscious buffer, which is likely realized in cortical processing rather than thalamic. Indeed, recent work has shown that attention can be allocated entirely in the absence of any conscious experience of the attended objects or locations at all (Jiang, Costello, Fang, Huang, & He, 2006).

One fundamental role of attention is to serve as a salience map that determines what processed inputs pass a threshold making them “worthy” of access to consciousness. It follows that attention, at least in this capacity as a “gatekeeper” to what is worthy of binding into object representations, must by definition be operative pre-consciously. That is, at least in this limited sense, attention must operate on inputs prior to and outside of consciousness. But we must distinguish between the pre-conscious and unconscious assessment of salience, and the consequences of passing the threshold for consciousness. The assessment of salience might occur, for example, in a low-level, local, retinotopic, or space-based frame of reference, whereas consciously attended contents might involve much more complex operations, such as tracking of objects over time (Kahneman et al., 1992). Consciously attended representations might occur in higher-level frames of reference than mere salience over locations, such as, for example, object-based coordinates. Alternatively, the preconscious operations of attention might be more complex than previously thought. While the findings of Jiang et al. (2006) implied that there could be enhanced salience to certain classes of images relative to scrambled images in the absence of any conscious experience of those images, their findings were consistent with an enhancement of salience at a location, because of either cortical or subcortical analyses of those images. More recent work, however, has shown that unconscious attentional operations involve object-centered representations, not just space-based operations (Chou & Yeh, 2012). It remains an open question whether even more complex attentional operations, such as object tracking, can occur in the absence of consciousness.

If we cannot endogenously attentionally access certain types of information, it would seem by definition that we cannot attend to them endogenously. How could we decide to allocate our attention to things to which we have no access? How would we even know where they were? Thus, endogenous attention, in contrast to exogenous or stimulus-driven attention, would seem necessarily to operate over those types of representations that can be reported, namely, the contents of conscious experience. Even though endogenous attentional allocation can probably not operate unconsciously (but see Kanai, Tsuchiya, & Verstraten, 2006), it might nonetheless be the case that endogenous attention to conscious contents can alter the contents of conscious experience by influencing how subsequent preconscious operations are executed which then result in experiences that are the products of those operations.

Endogenous attention can be allocated at will over limited spatial regions or layers of the visual field. This in turn can influence how the visual field is experienced. Indeed, supposed low-level features like color and brightness can change as a function of how attention is allocated in space, particularly within visually defined boundaries. Here we argue that attention accomplishes changes in phenomenal features by delimiting the domain over which preconscious operations will generate their outputs.

These preconscious operations include, but are not limited to, filling-in, computations of constancy (i.e., recovery of intrinsic size, shape, distance, pigment, transparency, luminosity, and material), computation of 3D surface, shape, distance, segmentation, and layout, motion matching, contour closure, contour completion, and illusory contour formation. Limiting the application of preconscious operations to only the feature primitives extant within an attended figure may have evolved as a way to save the expense of carrying out such operations and to speed their completion by focusing limited resources only where they are immediately needed. When the outputs of these operations differ as a function of the spatial domain over which attention is allocated, perceived features, such as experienced pigments, can appear to change dramatically as a function of how attention is endogenously allocated.

Specifying the Domain of Preconscious Operations

Evidence from color constancy

To make the above theoretical points concrete, we will focus on the specific case of endogenous attentional modulation of perceived reflectance or pigments. By way of background to the phenomena that we will consider, note that a given patch of gray will appear brighter against a dark background and darker against a bright background. The earliest models of brightness perception attempted to explain such illusions in terms of lateral inhibition occurring in the retina (Cornsweet, 1970) or cortex, where the activation of one cell inhibits the activation of its neighbors. Such models failed to explain how higher-level perceptual factors, such as inferred three-dimensional shape (Adelson, 1993), layout (Gilchrist, 1977), or curvature (Knill & Kersten, 1991), could influence brightness perception. In particular, the visual system must determine what portion of a single luminance or hue value detected at a location on the retina arises from each of several possible causes of that value in the world, such as surface reflectance, shadow, illumination, self-luminance, smoke, mist, or an intervening transparent layer. Models attempting to explain these effects have gone well beyond earlier models based solely on lateral inhibition among adjacent neurons. More recent models incorporate both low-level, localistic factors, such as lateral inhibition, and mid-level factors, such as the global geometric analyses that may underlie the decomposition (Watanabe & Cavanagh, 1993) of the image into contributions from reflectance, illumination, shadow, and transparency.

Models of perceived brightness or hue have emphasized that global context can influence the local outcomes of constancy operations. Bayesian models of color constancy (Brainard & Freeman, 1997; Brainard et al., 2006; Geisler & Kersten, 2002; Kersten & Yuille, 2003; Knill & Richards, 1996; Mamassian, Landy, & Maloney, 2002; Purves & Lotto, 2003; Rao, Olshausen, & Lewicki, 2002; Stocker & Simoncelli, 2006; von Helmholtz, 1866; Weiss, Simoncelli, & Adelson, 2002) emphasize the way that a scene illuminant can be estimated using a Bayesian operator with built-in prior “knowledge” about likely image to world mappings. Once computed, the illuminant can be effectively discounted in order to recover estimates of reflectance. A Gestaltist psychological approach emphasizes the global shape and configurational

operations that allow for scission of a scene into multiple layers, which can then be discounted in order to recover reflectance. For example, the reflectance that an object appears to have depends critically on measurements taken outside the object; in Anderson and Winawer's (2005) striking demonstrations, perceived white reflectance could be turned into perceived black reflectance by changing the context in which identical grayscale patches were embedded. Their theoretical claim that perceived surface albedo (lightness) results from the decomposition of the image into multiple layers can be bolstered by generalizing it to colored surface reflectances. For example, in the color case shown in Figure 12.4A, the left two disks are identical (except for a 90° rotation), and the right two disks are also identical. However, the left two disks appear to be brown versus green, and the right two disks purple versus blue. This is because the contextual differences afforded by interposing different transparent layers leads to a different interpretation of the underlying reflectance of the object, just as their theory suggests. Processes involving either Bayesian estimation via learned or innate priors or Gestaltist "scission" or segmentation into multiple layers may operate on the visual image prior to conscious vision, so that what we consciously perceive is the set of reflectances that are most likely in fact present in the world.

Both approaches have in common the discounting of spurious factors (e.g., the illuminant, intervening transparent layers, shadows) in order to recover what is intrinsically true about surfaces, including their pigments. However, both Bayesian and Gestaltist models (e.g., Gove, Grossberg, & Mingolla, 1995; Singh & Anderson, 2002) are nonetheless primarily "bottom-up" theories in the sense that perceived brightness is ultimately thought to be driven by the stimulus rather than some internal factor, such as attention. While these theories may not be incorrect, they would be incomplete to the extent that they disregard the roles that attention can play in the construction of perceived lightness and hue. The focus here is the manner in which endogenous attention can influence the outcome of the operations that culminate in a perceived brightness or hue. We argue that attention can effectively change the figural context or spatial domain over which operations such as filling-in and color, lightness, or size constancy operate automatically. This can in turn lead to attentionally modifiable perceived sizes, hues, lightnesses, and so forth.

Endogenous attention plays a role in the perceived brightness of overlapping transparent surfaces (Tse, 2005). Here we extend this finding from the domain of brightness to the domain of perceived hue, and provide simple demonstrations to make theoretical points. These effects suggest a mechanism whereby filling-in occurs within a figural boundary by effectively averaging feature information within the attended closed boundary, if that boundary can be interpreted as belonging to an individual transparent or opaque object, whether in the world or arising as a result of an afterimage. This filling-in process appears to coincide with mid-level visual operations that recover the intrinsic reflectances of surfaces, the relative depths of surfaces, and whether surfaces are seen as transparent or opaque. That is, what appears to get filled in is not just a hue, but a reflectance or pigment of a surface at a particular depth and a particular opaqueness and transparency.

A brightness version of figure-based attentional modulation of perceived features can be seen in Figure 12.4B. In the original report (Tse, 2005), eye movements were monitored and perceived brightnesses were assessed. The observer fixates a fixation

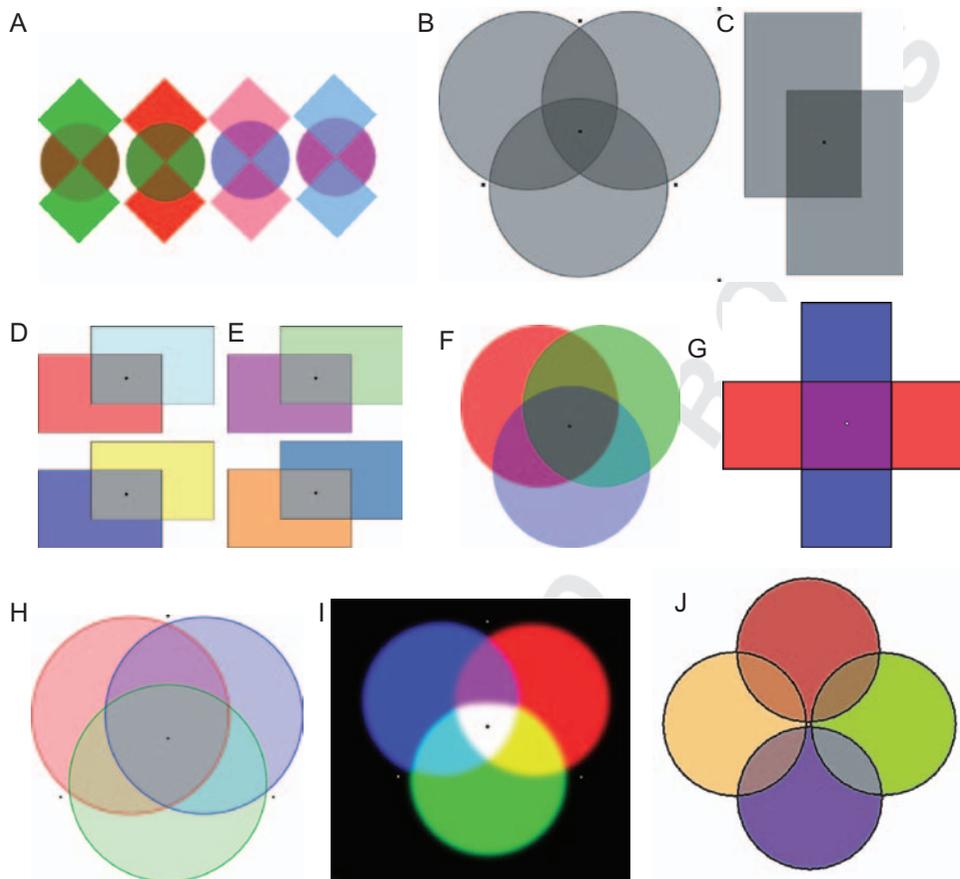


Figure 12.4 The left (right) two disks in A are identical except for a 90° rotation, yet appear to have a brown or green (blue or pink) reflectance, because the context implies occlusion by different transparent layers. Attending to one disk in B or one rectangle in C leads to a perceived darkening of that figure, but not a change in hue, because the overlap regions have the same hue as the surrounding regions. D: The overlap region is the same gray in each of the images. Choose one pair and fixate on the center spot, then attend to one colored rectangle or the other. This will alter the appearance of the central region. For example, attention to the ‘blue rectangle’ in panel E gives the appearance of a blue rectangle that is partially occluded by a transparent orange filter. Alternatively, attention to the “orange rectangle” gives the appearance of an orange rectangle that is partially occluded by a transparent blue filter. Attending to the overlap region alone leads to the perception of gray. The overlap region need not be grayscale in order for its perceived color to vary as a function of the figure that is attended. The only requirement is that the overlap region be consistent with an interpretation of a transparent object that occludes another transparent or opaque object. While a range of overlap hues is allowed, in general, the hue of the overlap region must lie between the hues of the two overlapping objects in color space (F). G: Here the central region is in fact the average color of the red and blue colors present in the periphery. Attending solely to the horizontal (vertical) rectangle leads to a perception that it is a uniformly pigmented rectangle occluded by a transparent layer of a different color.

point and then attends to a single disk. The attended disk was found to appear to darken. The effect requires that disks be consistent with the image that would be projected by transparent surfaces occluding a background, or as the backmost and now opaque layer occluded by transparent surfaces. The effect is robust; any combination of luminance values appears capable of creating the illusion, as long as the appearance of occluding transparent layers is preserved. When such an interpretation is not possible because key image cues for transparency (Metelli, 1974; Singh & Anderson, 2002) are absent, or because distinct objects consistent with attended boundaries cannot be individuated, perceived brightness is not modulated by attentional allocation. The simplest arrangement where the effect occurs is shown in Figure 12.4C, involving just two layers. Whichever rectangle is attended while maintaining fixation on the center dot appears to darken. There is also a tendency for the attended layer to become the backmost layer, and to be perceived as matte and opaque.

There have been surprisingly few studies to date showing that attention modulates stimulus appearance. The most direct test of this was carried out by Carrasco, Ling, and Read (2004; see also Cameron, Tai, & Carrasco, 2002; Carrasco, Penpeci-Talgar, & Eckstein, 2000). Similarly, Tse (1998; Figure 12.1H here) showed that one could attend to a given image one way, and experience illusory contours, or attend to a different figural organization of the same image and not see any illusory contours. Prior to these experiments, there were studies that came to the opposite conclusion, namely that attention cannot modulate perceived brightness at all (Prinzmetal, Nwachuku, Bodanski, Blumenfeld, & Shimizu, 1997), or that attention reduces the perceived contrast between a stimulus and its background (Tsal, Shalev, Zakay, & Lubow, 1994), contrary to the effect reported by Tse (2005).

In Carrasco and colleagues' experiment (2004), changes in perceived brightness were subtle and not consciously noticed by observers, though they were statistically significant as measured using points of subjective contrast equality specified by psychometric functions. The presumed mechanism for the type of contrast enhancement described by Carrasco et al. was attentional modulation of neuronal response gain in early visual areas (Martinez-Trujillo & Treue, 2002; McAdams & Maunsell, 1999; Reynolds & Desimone, 2003; Reynolds, Pasternak, & Desimone, 2000; Treue, 2000). The brightness illusions shown in Tse (2005; like Figures 12.4B and C here) demonstrated for the first time in a consciously noticeable and voluntarily manipulable manner that attention can modulate perceived brightness. This would seem to require higher-level attentional mechanisms than local contrast gain control, because such mechanisms are essentially local in nature, do not involve surface-, figure-, or object-level representations, or the specification of features in light of computations of reflectance or transparency. Whatever attentional mechanism can account for these observations, it would appear to operate at a level where surfaces are computed and divided into multiple layers, where figural boundaries are formed, and where figural objects or surfaces can interact, potentially inhibiting one another.

In Figures 12.4D and E the same basic phenomena (Tse, 2005) are extended from the domain of attention-induced alterations in brightness to the domain of attention-induced alterations in perceived hue. In these demonstrations, there is a grayscale overlap region among two or three colored figures that meet transparency conditions. Endogenously attending to one rectangle while maintaining fixation on the central

fixation spot tends to lead to filling-in of the hue from the attended figure into the central grayscale region, so that this grayscale region appears phenomenologically to take on the hue of the peripheral region. One may feel that one is seeing a uniformly colored rectangle through another overlapping, transparent rectangle that has a different hue. When one fixates and attends solely to the gray region, one indeed does perceive gray in this region. However, when one attends solely to a colored large rectangle, the entire rectangular region appears to have the hue of the outer portion of the attended rectangle. Sometimes it appears to take on a matte colored reflectance, and at other times it appears to be a colored transparent layer, depending on whether the figure is perceived to be behind or in front of the non-attended figure(s), respectively. The effect is somewhat paradoxical, in that as soon as one attends to only the central gray region, it appears as the gray color that it is. The color spreading effect only occurs if one attends to one of the large rectangles in its entirety. Note that in this and other cases of overlapping transparent surfaces, the different interpretations tend to undergo monocular rivalry, switching back and forth among possible percepts, even in the absence of voluntary shifts of attention.

The same effect can be experienced in the three examples of overlapping transparent disks shown in Figures 12.4F, H, I, and J. In each case, fixating on one of the fixation points while attending to only one disk will lead to the hues of the peripheral portions of the attended disk spreading into the overlap region that is in fact colorless. Figure 12.4I is perhaps most surprising, because the overlap region that is in fact white can appear bluish, greenish, or reddish, depending on how attention is allocated. If filling-in involves an area-weighted averaging of supporting features within a figural boundary (Hsieh & Tse, 2009), there will be less hue information in the filled-in end-state of that process.

While the overlap region between two overlapping surfaces, at least one of which is assumed to be transparent, need not be gray (see Figures 12.4G and J), the gray case is particularly interesting, because one perceives hue information to be present at a location where in fact there is none.

How does attention alter phenomenological experience in the central gray region? Attention appears to strengthen the attended boundary over non-attended or ignored figural boundaries, as indicated by the darker boundary of the attended disk in Figure 12.5. Because filling-in of features is carried out within closed figural boundaries (Grossberg & Mingolla, 1985a, 1985b; Hsieh & Tse, 2006, 2009, 2010) that here define foreground transparent objects, or background opaque or transparent objects, the peripheral hues are filled into the overlap region and perceived to take on the peripheral hue of the attended figure. This is indicated by the arrows in Figure 12.5. Filling-in can also account for the initial finding (Tse, 2005) of attentional modulation of brightness. In that case, the overlap region, which in Figure 12.4B is darker, spreads within the attended figure, presumably in a manner that averages brightness within a figure in an area-weighted fashion. Thus the attended figure appears darker in Figures 12.4B and C.

This process of averaging hues upon filling-in within a boundary is not unique; other features such as motion, luminance, and texture are also averaged within a figural boundary, and if a boundary undergoes perceptual fading, features can blend across boundaries present in the image (Hsieh & Tse, 2009). Featural filling-in

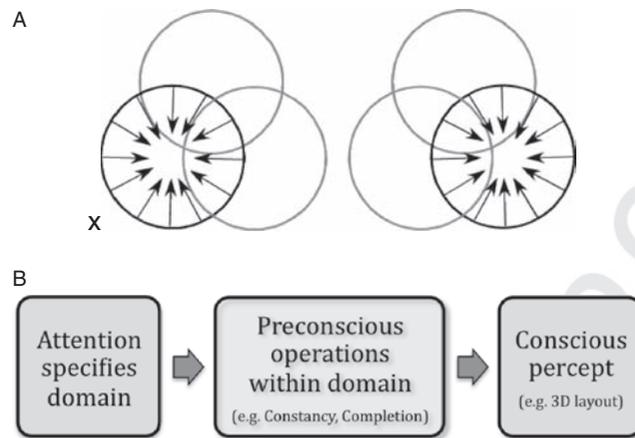


Figure 12.5 A: One of the three disks is voluntarily attended, strengthening its borders. Filling-in and associated feature averaging take place within these boundaries. If a different disk is attended, different features are filled into and averaged within the attended object. B: Schematic of theory: Attention specifies the domain within which preconscious operations such as filling in, motion matching, constancy operations, surface completion, and segmentation will take place. This results in the conscious experience of these outputs, typically as a 3D object/scene.

therefore proceeds within perceived boundaries as opposed to boundaries that are in the image. This process of figure specification, with its concomitant featural extension within that figural boundary, appears to be part of the mid-level processes that individuate possible objects at particular relative depths, that specify whether an object is opaque or transparent, and that compute what the intrinsic reflectance of a surface is.

The colors chosen for the overlap region in Figures 12.4G or J are not arbitrary. They fall on either side of gray in color space, such that their average would be gray were there really two overlapping layers with those colors, at least one of which is transparent. A process of scission could therefore create two independent layers, each with one of the two corresponding colors. That this is not the case when the overlap region has a hue that cannot be the average of the two peripheral hues suggests that more is going on than mere filling-in or averaging of colors within an attended boundary. It appears that scission into two layers occurs in light of color constancy operations that permit the recovery of surface reflectances, and that filling-in occurs within such post-scission surface representations.

In Figures 12.4G and J, it is shown that the overlap region need not be gray; it can be the average of the two hues present in the periphery. The color/luminance of the overlap region must be between the color space positions of the two overlapping surfaces in order for the color spreading effect to occur. This is because only when this condition of hue averageness in the overlap region is met can there be an ecologically valid scission into multiple overlapping transparent surfaces that have different hues. Attending to just the horizontal (vertical) surface in Figure 12.4G

while fixating the central dot can lead to a perception of a uniformly colored red (blue) layer, whether in front, in which case it appears to be transparent, or behind, in which case it appears to be opaque. Moreover, when this uniform state is reached, the perceived color within an entire attended rectangle is phenomenally more purple than either the pure red or pure blue present in the periphery. This is consistent with the weighted color “vector averaging” reported previously for filling-in in other studies (Hsieh & Tse, 2006, 2009, 2010).

The initial brightness effect (Tse, 2005) and present hue effect for the case of visible overlapping transparent surfaces can be explained with a simple series of processing steps: First, voluntarily attending to a given figure leads to strengthening of that figural boundary. Second, assuming the attended boundary can be individuated into a unique object with a uniform reflectance, filling-in of features occurs within those boundaries. The regions that are in fact grayscale take on a perceived hue because filling-in effectively averages (Hsieh & Tse, 2006, 2009, 2010) features within an attentionally defined figural boundary. The effect only happens in cases where transparency can be perceived, because only in the transparency case can multiple overlapping boundaries undergo scission into multiple, simultaneously existing colored surfaces, even if the backmost object is typically perceived as opaque. In the absence of transparency, scission into multiple overlapping colored objects fails, and filling-in cannot operate within a single object to the exclusion of other objects that could have been attended.

These demonstrations add to mounting evidence that attention can modulate perceived appearances, features, or qualia in a manner strong enough to be experienced firsthand. To date, attention has been shown to modulate the phenomenal experience of brightness (Tse, 2005), shape (Albert & Tse, 2000), illusory contours (Tse, 1998), and location (Tse et al., 2011). The present demonstrations show that voluntary attention can also modulate perceived hue.

But there need be no visible boundaries at all for attention to select one domain over another, within which preconscious operations will be carried out. In the case of the multistable Ponzo illusion shown in Figure 12.2A, an entire layer is brought to the fore attentionally. And in the case of apparent motion quartets, one can limit one’s attention to, say, space only encompassing the left pair of apparently moving dots. This will bias matching operations to occur within this spatial domain, leading to a perception of vertical apparent motion. Similarly, attending to the bottom spatial domain in the case of quartets’ apparent motion will bias one to see horizontal apparent motion. On this account, this occurs because the operation that matches an object to “itself” across space and time is biased to operate within an attended domain.

Conclusion

Attention influences processing in a top-down manner that is more profound than permitted by traditional notions of neuronal contrast gain control or feature binding. The theory presented here argues that attention can modulate perceived appearances because preconscious operations occur within attentionally specified domains. Typically this involves three steps (see Figure 12.5).

At time t_1 an observer attends to a spatial domain, effectively demarcating surfaces and figural boundaries, or a layer in the case of transparency, or a region over which matching can occur in the case of apparent motion. At t_2 preconscious operations are carried out automatically within the attended boundary, layer, or region. Preconscious operations that are influenced by figural specification include, but are not limited to: filling-in, motion matching, constancy operations that recover intrinsic properties such as reflectance, size, and shape, as well as surface completion and segmentation operations. This in turn generates, at t_3 , an experience that results from the outputs of these preconscious operations, typically a three-dimensional object/layout with particular intrinsic properties of pigment, size, shape, luminosity, transparency, and motion. On this account, attention does not alter the various preconscious operations, because these operate automatically within an attended domain. Instead, by delimiting the figural domain within which automatic preconscious operations are carried out on subsequent inputs, attention can influence the outcomes of those operations in a way that can be experienced firsthand.

Notes

- 1 A movie of this stimulus can be found at <http://www.journalofvision.org/content/11/3/12.long>.
- 2 A movie of this stimulus can be found at http://www.perceptionweb.com/perception/misc/p5568/Caplovitz_Demo.mov.

References

- Adelson, E. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042–2044. doi:10.1126/science.8266102.
- Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*(2), 298–313. doi:10.1037/0096-3445.135.2.298
- Albert, M., & Tse, P. (2000). The role of surface attachment in perceived volumetric shape. *Perception*, *29*, 409–420. doi:10.1068/p3040
- Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature*, *434*(7029), 79–83.
- Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, T. (1996). A formal theory of illusory conjunctions. *Psychological Review*, *103*, 165–192. doi.apa.org/psycinfo/1996-01716-007.
- Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*(4), 600–609. www.mitpressjournals.org/doi/pdf/10.1162/089892903321662976
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629. doi:10.1038/nrn1476
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences U S A*, *103*(2), 449–454. doi:10.1073/pnas.0507062103
- Beck, J., Elsner, A., & Silverstein, C. (1977). Position uncertainty and the perception of apparent movement. *Perception & Psychophysics*, *21*(1), 33–38.

- Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, *14*(7), 1393–1411.
- Brainard, D. H., Longère, P., Delahunt, P.B., Freeman, W.T., Kraft, J.M., & Xiao, B. (2006). Bayesian model of human color constancy. *Journal of Vision*, *6*(11), 1267–1281. doi:10.1167/6.11.10
- Cameron, E.L., Tai, J. C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, *42*(8), 949–967. [http://dx.doi.org/10.1016/S0042-6989\(02\)00039-1](http://dx.doi.org/10.1016/S0042-6989(02)00039-1)
- Caplovitz, G. P., & Tse, P. U. (2006). The bar-cross-ellipse illusion: Alternating percepts of rigid and non-rigid motion based on contour ownership and trackable feature assignment. *Perception*, *35*(7), 993–997. www.perceptionweb.com/perception/perc0706/p5568.pdf
- Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*(1–2), 179–213. [linkinghub.elsevier.com/retrieve/pii/S0010-0277\(00\)00154-2](http://linkinghub.elsevier.com/retrieve/pii/S0010-0277(00)00154-2)
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*(3), 308–313.
- Carrasco, M., Penpeci-Talgar, C., & Eckstein, M. (2000). Spatial attention increases contrast sensitivity across the CSF: Support for signal enhancement. *Vision Research*, *40*(10–12).
- Chatterjee, S. H., Freyd, J. J., & Shiffrar, M. (1996). Configural processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 916–929. Retrieved December 10, 2012, from <http://dynamic.uoregon.edu/~jjf/articles/cfs96.pdf>
- Choi, H., & Scholl, B. J. (2006). Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. *Perception*, *35*, 385–399. Retrieved December 10, 2012, from <http://perception.research.yale.edu/papers/06-Choi-Scholl-Perception.pdf>
- Chou, W.-L., & Yeh, S.-L. (2012). Object-based attention occurs regardless of object awareness. *Psychonomic Bulletin and Review*, *19*(2), 225–231. doi:10.3758/s13423-011-0207-5.
- Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. Retrieved December 10, 2012, from <http://wexler.free.fr/library/files/desimone%20281995%29%20neural%20mechanisms%20of%20selective%20visual%20attention.pdf>
- Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science*, *287*, 2036–2038. doi:10.1126/science.287.5460.2036
- Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cognition*, *8*(2), 128–151. Retrieved December 10, 2012, from <http://www.ini.ethz.ch/~peterk/OwnPapers/engel.cc.99.pdf>
- Enns, J. T., & Rensink, R. A. (1991). Preconscious recovery of three-dimensional orientation from line drawings. *Psychological Review*, *98*(3), 335–351. doi:10.1080/13506280042000135.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, J. (2001). *The mind doesn't work that way: The scope and limits of computational psychology*. Cambridge, MA: MIT Press.
- Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Geisler, W. S., & Kersten, D. (2002). Illusions, perception, and Bayes. *Nature Neuroscience*, *5*, 508–510. Retrieved December 10, 2012, from <http://www.cns.nyu.edu/ftp/eero/Kersten-Geisler-NN-news+views.pdf>

- Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, 195, 185. doi:10.1126/science.831266
- Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception & Psychophysics*, 58(8), 1260–1277.
- Gordon, R. D., & Irwin, D. E. (2000). The role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 136–150. doi:10.1037/0278-7393.26.1.136.
- Gove, A., Grossberg, S., & Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, 12(6), 1027–1052. doi:http://dx.doi.org/10.1017/S0952523800006702.
- Green, M. (1986). What determines correspondence strength in apparent motion? *Vision Research*, 26, 599–607. http://dx.doi.org/10.1016/0042-6989(86)90008-8.
- Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics*, 38(2), 141–171. doi:10.3758/BF03198851
- Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, 92, 173–211. doi:10.1037/0033-295X.92.2.173.
- Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenbewertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung*, 11b, 513–524.
- He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, 359(6392), 231–233. Retrieved December 10, 2012, from <http://cw.tsolab.org/jclub/20030915/He%20and%20Nakayama.pdf>
- He, Z. J., & Nakayama, K. (1994). Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature*, 367(6459), 173–175. Retrieved December 10, 2012, from <http://visionlab.harvard.edu/members/ken/Ken%20papers%20for%20web%20page/068HeKNNature1994.pdf>
- Helmholtz, H. von. (1866). *Treatise on physiological optics*. New York: Dover.
- Helmholtz, H. L. von. (1910 [1867]). *Handbuch der physiologischen Optik*. Leipzig: L. Voss. (Repr. in A. Gullstrand, J. von Kries, & W. Nagel [Eds.]. [1910]. *Handbuch der physiologischen Optik* [3rd ed.]. Hamburg: L. Voss)
- Hill, W. E. (1915). My wife and my mother-in-law. *Puck*, 16, 11.
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society London B: Biological Science*, 353(1373), 1257–1270. doi:10.1098/rstb.1998.0281
- Hsieh, P.-J., & Tse, P. U. (2006). Illusory color mixing upon perceptual fading and filling-in does not result in “forbidden colors.” *Vision Research*, 46(14), 2251–2258. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/hsiehVR.pdf>
- Hsieh, P.-J., & Tse, P. U. (2009). Feature mixing rather than feature replacement during perceptual filling-in. *Vision Research*, 49(4), 439–450. <http://www.dartmouth.edu/~petertse/featuremixing.pdf>
- Hsieh, P.-J., & Tse, P. U. (2010). “Brain-reading” of perceived colors reveals a feature mixing mechanism underlying perceptual filling-in in cortical area V1. *Human Brain Mapping*, 31(9), 1395–1407. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/HsiehTseHBM.pdf>
- Jastorff, J., & Orban, G. A. (2009). Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. *Journal of Neuroscience*, 29(22), 7315–7329. doi:10.1523/JNEUROSCI.4870-08.2009

- Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences*, *103*, 17048–17052. doi:10.1073/pnas.0605678103
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211. doi:10.3758/BF03212378
- Kahneman, D., Treisman, A. M., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175–219. [http://dx.doi.org/10.1016/0010-0285\(92\)90007-O](http://dx.doi.org/10.1016/0010-0285(92)90007-O)
- Kanai, R., Tsuchiya, N., & Verstraten, F. A. (2006). The scope and limits of top-down attention in unconscious visual processing. *Current Biology*, *16*(23), 2332–2336. <http://dx.doi.org/10.1016/j.cub.2006.10.001>
- Karmiloff-Smith, A. (1995). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*, 150–158. [http://dx.doi.org/10.1016/S0959-4388\(03\)00042-4](http://dx.doi.org/10.1016/S0959-4388(03)00042-4)
- Knill, D., & Kersten, D. (1991). Apparent surface curvature affects lightness perception. *Nature*, *351*, 228. Retrieved December 10, 2012, from <http://gandalf.psych.umn.edu/users/kersten/kersten-lab/papers/knill91.pdf>
- Knill, D., & Richards, W. (1996). The perception of shading and reflectance. In D. Knill & W. Richards (Eds.), *Visual perception: Computation and psychophysics* (pp. 409–423). New York: Cambridge University Press.
- Koch, C. (2004). *The quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts and Company.
- Koffka, K. (1935). *Principles of Gestalt psychology*. London: Lund Humphries.
- Kohler, A., Haddad, L., Singer, W., & Muckli, L. (2008). Deciding what to see: The role of intention and attention in the perception of apparent motion. *Vision Research*, *48*(8), 1096–1106. <http://dx.doi.org/10.1016/j.visres.2007.11.020>
- Korte, A. (1915). Kinematoscopische Untersuchungen. *Zeitschrift für Psychologie*, *72*, 193–206.
- Lamy, D., & Tsal, Y. (2000). Object features, object locations, and object files: Which does selective attention activate and when? *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1387–1400. doi:10.1037/0096-1523.26.4.1387
- Liu, T., Abrams, J., & Carrasco, M. (2009). Voluntary attention enhances contrast appearance. *Psychological Science*, *20*(3), 354–362. doi:10.1111/j.1467-9280.2009.02300.x
- Mamassian, P., Landy, M. S., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. P. N. Rao, B. A. Olshausen, & M. S. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (pp. 13–36). Cambridge, MA: MIT Press.
- Martinez-Trujillo, J. C., & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron*, *35*, 365–370. [http://dx.doi.org/10.1016/S0896-6273\(02\)00778-X](http://dx.doi.org/10.1016/S0896-6273(02)00778-X)
- McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, *19*, 431–441. <http://www.neuro.cjb.net/content/19/1/431.full.pdf>
- Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4), 91. doi:10.1038/scientificamerican0474-90
- Neri, P., Morrone, C., & Burr, D. (1999). Seeing biological motion. *Nature*, *395*(6705), Oct. 29, 894–896. Retrieved December 10, 2012, from http://www.abdn.ac.uk/~mbi337/pn_nature1998.pdf
- Oakes, L.M., Ross-Sheehy, S., & Luck, S. J. (2006). Rapid development of feature binding in visual short-term memory. *Psychological Science*, *17*(9), 781–787. doi:10.1111/j.1467-9280.2006.01782.x

- Prinzmetal, W., Nwachuku, I., Bodanski, L., Blumenfeld, L., & Shimizu, N. (1997). The phenomenology of attention. 2. Brightness and contrast. *Conscious Cognition*, 6(2–3), 372–412. <http://dx.doi.org/10.1006/ccog.1997.0313>
- Purves, D., & Lotto, R. B. (2003). *Why we see what we do: An empirical theory of vision*. Sunderland, MA: Sinauer.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197. <http://dx.doi.org/10.1163/156856888X00122>
- Quinlan, P.T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin*, 29(5), 643–673. doi:10.1037/0033-2909.129.5.643
- Ramachandran, V. S., Ginsburg, A. P., & Anstis, S. M. (1983). Low spatial frequencies dominate apparent motion. *Perception*, 12, 457–461. PMID: 6672741
- Rao, R. P. N., Olshausen, B. A., & Lewicki, M. S. (Eds.). (2002). *Probabilistic models of the brain: Perception and neural function*. Cambridge, MA: MIT Press.
- Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, 102(1), 101–130. doi:10.1037/0033-295X.102.1.101
- Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research*, 38(15–16), 2489–2505. [http://dx.doi.org/10.1016/S0042-6989\(98\)00051-0](http://dx.doi.org/10.1016/S0042-6989(98)00051-0)
- Reynolds, J., & Heeger, D. (2009). The normalization model of attention. *Neuron*, 61(2), 168–185. <http://dx.doi.org/10.1016/j.neuron.2009.01.002>
- Reynolds, J. H., & Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 24(1), 19–29, 111–25. Retrieved December 10, 2012, from <http://snl-r.salk.edu/publications/rd.pdf>
- Reynolds, J. H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron*, 37, 853–863. [http://dx.doi.org/10.1016/S0896-6273\(03\)00097-7](http://dx.doi.org/10.1016/S0896-6273(03)00097-7)
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, 26, 703–714. [http://dx.doi.org/10.1016/S0896-6273\(00\)81206-4](http://dx.doi.org/10.1016/S0896-6273(00)81206-4)
- Rubin, E. (1958). Figure–ground perception. In *Readings in perception* (M. Wertheimer, Trans.). Princeton, NJ: Van Nostrand. (Original work published 1915)
- Saalman, Y. B., & Kastner, S. (2009). Gain control in the visual thalamus during perception and cognition. *Current Opinion in Neurobiology*, 19(4), 408–414. <http://dx.doi.org/10.1016/j.conb.2009.05.007>
- Schneider, W. X. (1999). Visual-spatial working memory, attention, and scene representation: A neuro-cognitive theory. *Psychological Research*, 62(2–3), 220–236. doi:10.1007/s004260050052
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, 80(1–2), 159–177. [http://dx.doi.org/10.1016/S0010-0277\(00\)00157-8](http://dx.doi.org/10.1016/S0010-0277(00)00157-8)
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science*, 5, 195–200. doi:10.1111/j.1467-9280.1994.tb00500.x
- Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science*, 1, 257–264. doi:10.1111/j.1467-9280.1990.tb00210.x
- Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, 4(6), 379–384. doi:10.1111/j.1467-9280.1993.tb00585.x
- Shipp, S., Adams, D. L., Moutoussis, K., & Zeki, S. (2009). Feature binding in the feedback layers of area V2. *Cerebral Cortex*, 19(10), 2230–9. doi:10.1093/cercor/bhn243
- Singh, M., & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review*, 109(3), 492–519. doi:10.1037/0033-295X.109.3.492

- Stocker A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 9, 578–585. doi:10.1038/nrn1669
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, 40A, 201–236.
- Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychology*, 47(7), 862–875. doi:10.1037/0003-066X.47.7.862
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6(2), 171–178. [http://dx.doi.org/10.1016/S0959-4388\(96\)80070-5](http://dx.doi.org/10.1016/S0959-4388(96)80070-5)
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14(4–8), 411–443. doi:10.1080/13506280500195250
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136. [http://dx.doi.org/10.1016/0010-0285\(80\)90005-5](http://dx.doi.org/10.1016/0010-0285(80)90005-5)
- Treue, S. (2000). Neural correlates of attention in primate visual cortex. *Trends in Neuroscience*, 24, 295–300. [http://dx.doi.org/10.1016/S0166-2236\(00\)01814-2](http://dx.doi.org/10.1016/S0166-2236(00)01814-2)
- Tsal, Y., Shalev, L., Zakay, D., & Lubow, R. E. (1994). Attention reduces perceived brightness contrast. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 47(4), 865–893. doi:10.1080/14640749408401100
- Tse, P. (1998). Illusory volumes from conformation. *Perception*, 27(8), 977–994. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/tseIllusoryVols.pdf>
- Tse, P., & Albert, M. (1998). Amodal completion in the absence of image tangent discontinuities. *Perception*, 27, 455–464. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/tsealbert.pdf>
- Tse, P., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. In T. Watanabe (Ed.), *High-level motion processing: Computational, neurobiological and psychophysical perspectives* (pp. 249–266). Cambridge, MA: MIT Press.
- Tse, P. U. (1999a). Volume completion. *Cognitive Psychology*, 39, 37–68. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/TseVolumeCompletion.pdf>
- Tse, P. U. (1999b). Complete mergeability and amodal completion. *Acta Psychologica*, 102, 165–201. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/TseCompleteMerge.pdf>
- Tse, P. U. (2002). A contour propagation account of surface filling-in and volume formation. *Psychological Review*, 109(1), 91–115. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/TsePsychReview.pdf>
- Tse, P. U. (2005). Voluntary attention modulates the brightness of overlapping transparent surfaces. *Vision Research*, 45(9), 1095–1098. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/tseatttransp.pdf>
- Tse, P. U. (2006a). Neural correlates of transformational apparent motion. *Neuroimage*, 31(2), 766–773. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/TAMpublished.pdf>
- Tse, P. U. (2006b). How the evolution of symbolic cognition transformed human morality. In W. Sinnott-Armstrong (Ed.), *Moral psychology. Volume 1: The evolution of morality* (pp. 269–314). Cambridge, MA: MIT Press.
- Tse, P. U., & Caplovitz, G. P. (2006). Contour discontinuities subserve two types of form analysis that underlie motion processing. *Progress in Brain Research*, 154, 271–292. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/TC.pdf>
- Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition*, 74, B27–B32. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/chinese.pdf>

- Tse, P. U., & Logothetis, N. K. (2002). The duration of 3-D form analysis in transformational apparent motion. *Perception and Psychophysics*, *64*(2), 244–265. Retrieved December 10, 2012, from <http://www.dartmouth.edu/~petertse/Tse3D.pdf>
- Tse, P. U., Sheinberg, D. L., & Logothetis, N. K. (2002). Fixational eye movements are not affected by abrupt onsets that capture attention. *Vision Research*, *42*, 1663–1669.
- Tse, P. U., Whitney, D., Anstis, S., & Cavanagh, P. (2011). Voluntary attention modulates motion-induced mislocalization. *Journal of Vision*, *11*(3), 12. Retrieved December 10, 2012, from <http://www.journalofvision.org/content/11/3/12.long>
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*, 451–473. PMID: 6726493
- Watanabe, T., & Cavanagh, P. (1993). Surface decomposition accompanying the perception of transparency. *Spatial Vision*, *7*(2), 95–111. <http://dx.doi.org/10.1163/156856893X00306>
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*, 598–604. Retrieved December 10, 2012, from <http://www.cns.nyu.edu/~msl/courses/2223/Readings/Weiss-NatNeurosci2002.pdf>
- Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Science*, *7*(2), 70–76. [http://dx.doi.org/10.1016/S1364-6613\(02\)00024-4](http://dx.doi.org/10.1016/S1364-6613(02)00024-4)
- Wolfe, J. M. (2010). Visual search. *Current Biology*, *20*(8), R346–R349. PMID: 21749949
- Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Science*, *15*(2), 77–84. <http://dx.doi.org/10.1016/j.tics.2010.12.001>

UNCORRECTED PROOF

UNCORRECTED PROOFS