
Discussion

Contrasting theories of White's illusion

Howe (2001) presents a new variant of White's illusion and asserts that it demonstrates the failures of a number of models of White's illusion, including my own (Anderson 1997). He also argues that the FACADE model of Grossberg can explain this effect, as well as the other published variants of White's illusion (Grossberg 1997; Kelly and Grossberg 2000). My response has two parts. First, although Howe's display does reveal the need to go beyond the local properties of T-junctions, I contend that it does not provide a critical test of the two competing explanations of White's effect (namely, a contrast theory, and a scission theory). Second, I will argue that although the FACADE model can provide some insight into this particular variant of White's illusion, it actually fails to account for White's original effect or variants of the illusion I have published previously (Anderson 1997).

Howe's demonstration consists of a very weak version of White's effect containing a single target bar. The main change to White's stimulus is the addition of extended black and white horizontal bands that cut across the image. In his display, the contours of these bands are carefully aligned with the stems of the targets' T-junctions. Howe finds that observers report that the illusion is either greatly reduced or inverted by this manipulation. He argues that this impacts negatively on T-junction-based theories of this illusion, since the local T-junctions are unchanged in the top and bottom figure. This seems fair enough, and points to a need to consider other, nonlocal properties to account for the effect. The main change introduced by Howe's variant of White's illusion is the presence of aligned contours at the T-junctions from the horizontally extended bars in his pattern. Although Howe did not actually demonstrate that the alignment of these contours is critical for his effect, it can be easily demonstrated; see figure 1. When the horizontal bands are aligned with the stems of the T-junction, the illusion is greatly reduced (as Howe reports); however, when they are misaligned, the illusion returns to normal (which he did not). Why would this be so? The simplest and most obvious answer is that the extended horizontal contours effectively transform the local T-junctions into X-junctions, which is precisely how Howe tries to explain the new effect. The question is whether this provides any conceptual leverage into the two competing theories of this illusion that Howe discusses: contrast theory, and scission theory [see also Gilchrist et al (1999) for an anchoring account of this illusion].

Contrast-based models attribute White's effect to a selective contrast between the gray target bars and the stripes on which the targets are embedded. Scission theory (Anderson 1997) attributes this illusion to mechanisms that are responsible for decomposing images into multiple layers. The theoretical problem created by White's illusion for contrast theories is that the perceived lightness difference is in the wrong direction. One way to 'solve' this problem is to assert that the edges that would drive the illusion in the wrong direction are somehow discounted. Such models essentially reduce White's illusion to a form of simultaneous contrast: a gray bar in a white stripe becomes equivalent to a gray target on a white background; and a gray bar in a black stripe becomes equivalent to a gray target on a black background.

My scission-based account was motivated (in part) by a significant shortcoming of such theories. By reducing White's illusion to a variant of simultaneous contrast, such theories place an upper bound on the strength of the illusion: White's effect can only be as large as a comparable simultaneous-contrast display. But even White's original

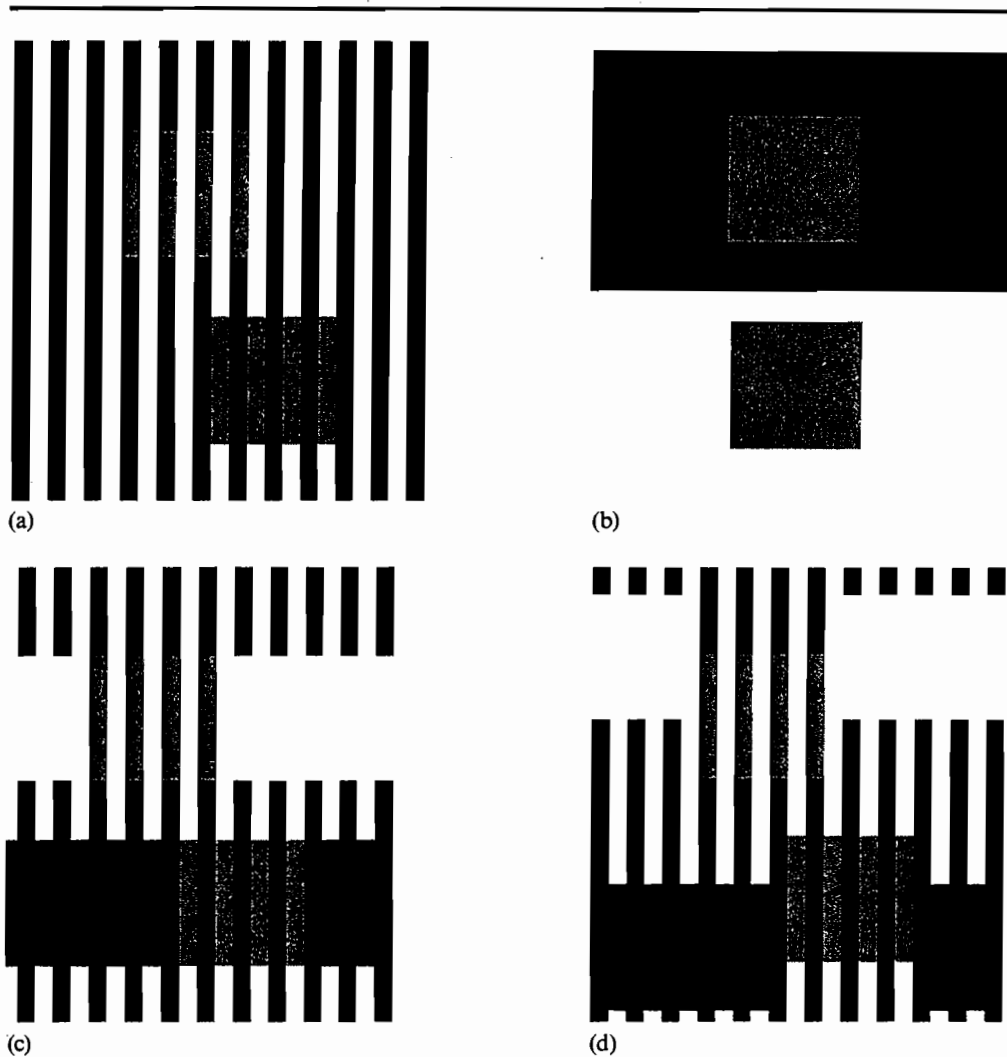


Figure 1. Contrast based theories restrict White's effect to be smaller in magnitude or equal to simultaneous contrast (b). However, as White originally reported, this effect (a) can be much larger than simultaneous contrast (b). A variant of Howe's variant of the illusion (c) greatly reduces or reverses the magnitude of White's effect, but the effect is restored when the horizontal contours are misaligned (d). It is argued that Howe's result can be understood with virtually any model of White's effect, so it does not provide a critical test of competing theories.

papers showed that this effect can be much *stronger* than simultaneous contrast, which greatly limits the theoretical force of these models. The importance of this point for theories of this illusion continue to be ignored, and contrast-based models are left wanting because of this deficiency. The issue raised by Howe is whether his new demonstration can be understood from a scission-based account. The intuitive content of my scission account is that the color of the stripe on which the target is embedded is removed from the target (ie attributed to an underlying layer). If the underlying layer is black, then the target should appear lighter (since some of its darkness is attributed to an underlying layer); the converse holds if the underlying layer is white. In Howe's demonstration, there are two possible interpretations of the underlying layer: either the long thin bars on which the target is embedded; or the extended thicker bars adjacent to the target (that have the opposite color as the extended bars). Since it is unclear whether the underlying layer is black or white, it is not particularly surprising

that the illusion is abolished. The same ambiguity holds for contrast-based theories that would serve as the background for the target, so this does not at all seem like a critical test of these two accounts.

Howe concludes by suggesting that the FACADE model of Grossberg and colleagues provides a better account of this and previously reported versions of White's effects. Grossberg and colleagues' model provides two explanations of White's effect. The primary explanation of the effect is similar to other contrast-based theories: "in the Benary and Munker-White displays, the contrastive illusion is explained by analyzing how the visual system interprets whether the gray patch is solely on a white or a black background, thereby discounting the effect of other spatially congruent regions" (Kelly and Grossberg 2000, page 1608). These authors note that "the Munker-White illusion is considerably stronger than the Benary illusion"; and speculate that "... this may be because, unlike the case of the Benary cross, amodal completion of the gray patches occurs in this display" (page 1610). There are a number of problems with this explanation. First, the amodal completion that occurs in their model depends critically on the alignment of the T-junction stems in White's display, but the magnitude of White's illusion is not affected when the stems are misaligned (Anderson 1997; cf Todorović, this discussion). Second, even if the amodal completion could (somehow) 'explain' the difference between White's and Benary's illusions, it cannot explain why White's illusion is larger than its simultaneous-contrast analogue (wherein the target bars are assumed to complete and form a single gray rectangle), which forms the basis of Grossberg and colleagues' explanation (see figure 1).

The other explanation of the illusion offered by Grossberg and colleagues is similar to the scission explanation I proposed. However, to achieve this percept in their model, the authors have to strengthen the contrast of the T-junction stems by hand to prevent the occluded interpretation from arising in their model (which they attribute to some form of attentional bias; see Kelly and Grossberg 2000, page 1612). However, here, too, the model's success depends critically on the alignment of the T-junction stems, which is (again) not observed psychophysically. Moreover, their model predicts that the three gray target bars that they used in their test display should not appear equal in lightness, which is also not observed psychophysically (see figures 19e and 19f in Kelly and Grossberg 2000). In sum, there are a number of substantial discrepancies between the predictions of Grossberg et al's model of White's illusion and human perception, so there is little to support Howe's claim that this model currently has an advantage over other related models in accounting for White's illusion.

Acknowledgement. This work was supported by NIH grant R29-EY11487-01A1.

Barton L Anderson

Brain and Cognitive Sciences, Massachusetts Institute of Technology, E10-E135, Cambridge, MA 02139, USA; e-mail: bart@psyche.mit.edu

References

- Anderson B L, 1997 "A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions" *Perception* **26** 419-453
- Gilchrist A, Kossyfidis C, Agostini T, Li X, Bonato F, Cataliotti J, Spehar B, Annan V, Economou E, 1999 "An anchoring theory of lightness perception" *Psychological Review* **106** 795-834
- Grossberg S, 1997 "Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures" *Psychological Review* **104** 618-658
- Howe P D L, 2001 "A comment on the Anderson (1997), the Todorović (1997), and the Ross and Pessoa (2000) explanations of White's effect" *Perception* **30** 1023-1026
- Kelly F, Grossberg S, 2000 "Neural dynamics of 3-D surface-perception: Figure-ground separation and lightness perception" *Perception & Psychophysics* **62** 1596-1618
- Todorović D, 1997 "Lightness and junctions" *Perception* **26** 379-395

Lightness, junctions, and depth

Howe (2001) has constructed a variant of White's display which fulfills the stimulus conditions for a T-junction-based lightness rule that I formulated earlier (Todorović 1997), but which does not exhibit the predicted perceptual consequences. Cataliotti and Bonato (1999) found that the flanking regions in White's display exert a greater effect on the lightness of target figures than previous authors had assumed. Such data indicate that my earlier claim that lightness of target figures *predominantly* depends on the luminance ratio with collinear regions is in need of amendment. The geophotometric structure of junctions is an important stimulus condition of lightness, but it is only one among several conditions whose relative contributions may vary in different circumstances and which still remain to be elucidated in detail.

Howe's theory of the lightness effect in his figure 1a is based on work by Kelly and Grossberg (2000). An important condition in his account is a particular depth stratification of regions in that figure: neural interactions induced by T-junctions (see my

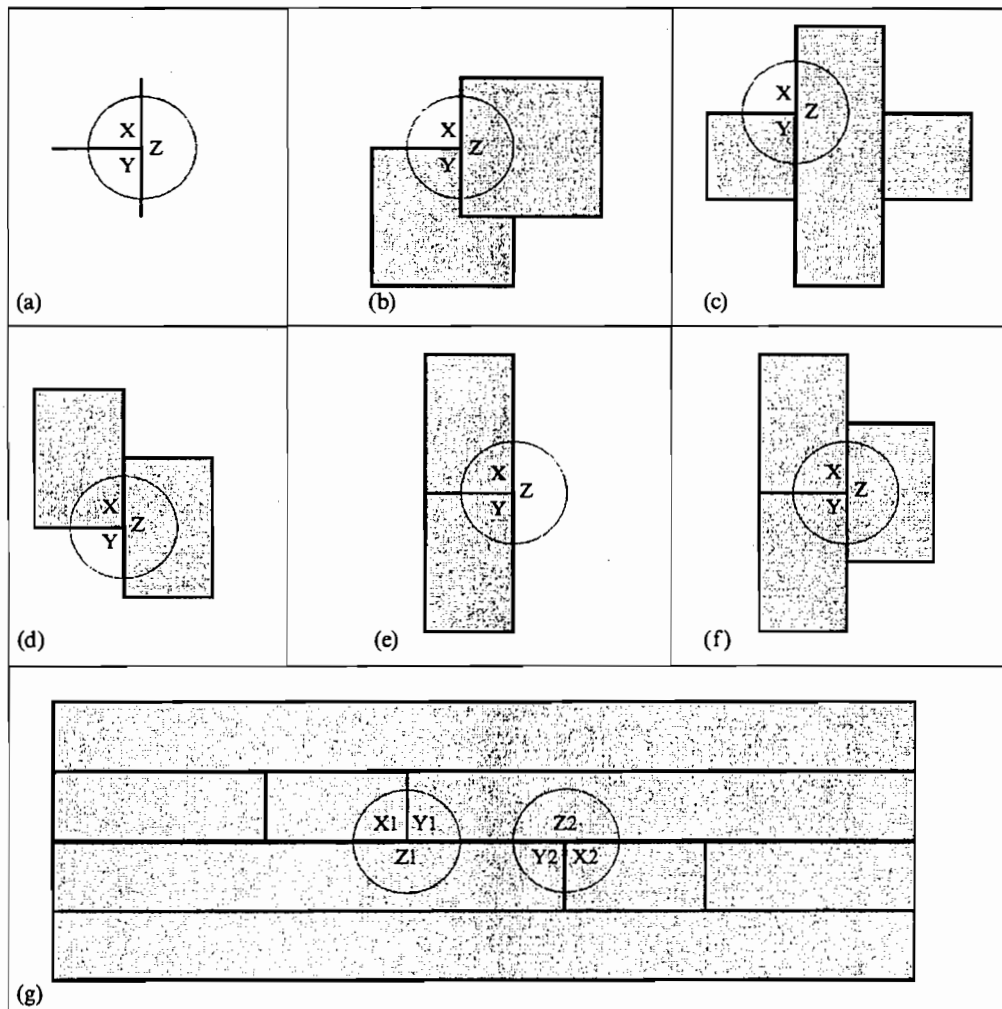


Figure 1. (a) A T-junction involves a flanking region, Z, and two collinear regions, X and Y. (b, c) Figural arrangements in which region Z is perceived as in front and occluding regions X and Y. (d, e, f) Figural arrangements in which Z is not perceived as in front or as occluding both regions X and Y. (g) A case of potentially inconsistent depth assignments (see text).

figure 1a) are assumed to cause the flanking region, Z, to be represented at a nearer depth plane and perceived to be occluding the collinear regions, X and Y, which are represented at a further depth plane. In this theory, subsequent neural processing, causing the differential lightness effects, essentially depends on such a depth stratification. The problem I see here is that T-junctions do not necessarily induce such depth assignments, either generally or in this particular case.

In order to study which conditions favor occlusion (depth stratification) and which favor juxtaposition (same-plane percepts), researchers have constructed a large variety of figural arrangements involving different types of region junctions (Boselie and Wouterlood 1989; Wouterlood and Boselie 1992; Boselie 1994; van Lier et al 1994). This research has shown that T-junctions may be associated with a number of different percepts, depending on factors such as figure closure, good continuation, symmetry, etc. Figure 1 presents some relevant examples involving line patterns representing opaque 2-D figures. Cases such as figures 1b and 1c are in accord with the depth stratification assumed by Howe, in that region Z is indeed perceived as nearer than occluding regions X and Y. However, cases such as figures 1d, 1e, and 1f are counterexamples, because region Z is not perceived as nearer or as occluding both of the two other regions. In the study by Boselie and Wouterlood (1989), which used stimuli similar to figures 1d and 1e, 100% of the subjects interpreted as *juxtaposed* regions X and Z in figure 1d, and regions X and Y in figure 1e. As for figure 1f, although a similar arrangement appears not to have been used in published studies, it readily evokes an interpretation of three juxtaposed regions: X, Y, and Z. In figure 1g, involving several constellations similar to figure 1f, and relevant to White's effect, if depth were assigned as in figures 1b and 1c, region Z1 would be assigned to be in front of Y1, and Z2 in front of Y2; however, Z1 is the same region as Y2, and Y1 the same as Z2 (see Todorović 1997). The fact that identical T-junctions, embedded in different figural arrangements, may be associated with different or potentially inconsistent depth assignments, creates problems for theories that assume that T-junctions automatically trigger particular depth stratifications, and that lightness in 2-D displays depends on depth specified in this manner.

To assess the role of T-junctions and depth stratification, it is instructive to compare two versions of White's display. In figure 2a, which is similar to the one used

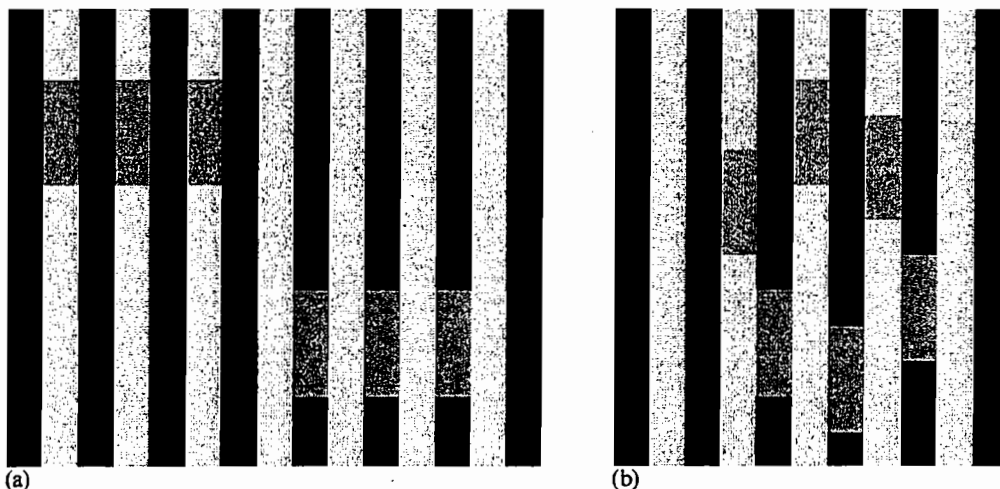


Figure 2. Two variants of White's effect with similar strength: (a) targets on separate stripes, involving consistent potential depth assignments; (b) targets on interleaved stripes, involving potentially inconsistent depth assignments.

in the simulations by Kelly and Grossberg (2000), the stems of the T-junctions of nearby gray targets in both target groups are aligned. Such alignment involves line patterns similar to those in figure 1c, and may induce the corresponding depth stratification; furthermore, Kelly and Grossberg (2000) noted that attentional shifts may induce alternate depth assignments, causing the gray targets in some cases to be perceived as in front. In contrast, in figure 2b the alignment of stems is disturbed, providing no support for either depth stratification. Also, in figure 2a the two groups of gray targets are well separated, so that depth assignment as in figure 1c (Z in front of X and Y) would be consistent for all four T-junctions in all six targets. In contrast, in figure 2b the two sets of targets are positioned on alternating, interleaved stripes, in a constellation similar to figure 1g, so that local depth assignment of this type could be globally inconsistent. Targets on interleaved stripes were used in the original version of White's effect (White 1979), and in several subsequent studies.

As noted, in his theoretical account Howe claims that both gray bars in his figure 1a, together with the stripes they lie on, are perceived as being occluded by the flanking stripes, which would agree with the assumed depth stratification. However, such a percept does not appear salient. Howe himself does not mention such occlusion in his introduction, but describes the figure as consisting of "alternating white and black stripes onto which two gray bars have been superimposed". In figure 2b here, with multiple targets on interleaved stripes, involving potentially inconsistent depth assignments, the theoretically required occlusion is also not very apparent. A tendency for such stratification might conceivably be manifested in bi-stable or two-layered depth percepts, or undulation of black and white stripes in depth. In fact, the display may be better described simply as a flat mosaic. In sum, depth stratification is neither generally unambiguously demanded by the stimulus structure of such displays, nor salient in the particular percepts; should it, then, be so prominently invoked in the theory? Although depth plays a role in other circumstances, its relevance for lightness in flat, nonstereoscopic, homogeneously illuminated displays is not well established (see Todorović 1997).

Although quite different in terms of potential depth stratification, the two variants of White's effect in figure 2 appear to have the same strength. To test this, two displays similar to figures 2a and 2b were presented to twenty-seven subjects on computer screens, in four orientations each (original, horizontal inversion, vertical inversion, both inversions), in random order. The task was to match the central members in both groups of gray targets on a Munsell scale, consisting of sixteen paper chips ranging from Munsell value 2 to 9.5, in 0.5 increments. The lightness judgments for the two variants were virtually identical. The mean Munsell matches for the lighter-appearing targets were 6.58 for figure 2a and 6.60 for figure 2b, and for the darker-appearing targets they were 5.42 and 5.46, respectively. A 2×2 ANOVA found the lightness effect significant ($F_{1,26} = 403.25, p < 0.0001$), whereas the type of display and the interaction were not significant ($F < 1$). Thus the strength of the phenomenon is unaffected by differences in T-junction stem alignment (as noted already by Anderson 1997), and potential depth assignment. It would be interesting to compare how the Kelly and Grossberg (2000) model would handle the two variants, with respect to both depth assignment and strength of the lightness effect.

Dejan Todorović ¶

Department of Psychology, University of Belgrade, Čika Ljubina 18-20, 11000 Belgrade, Serbia, Yugoslavia; ¶ Current address: Department of Psychology, University of Toronto at Scarborough, 1265 Military Trail, Scarborough, Ontario M1C 1A4, Canada; e-mail: dejan@utsc.utoronto.ca

References

- Anderson B L, 1997 "A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions" *Perception* 26 419–453
- Boselie F, 1994 "Local and global factors in visual occlusion" *Perception* 23 517–528
- Boselie F, Wouterlood D, 1989 "The minimum principle and visual pattern completion" *Psychological Research* 51 93–101
- Cataliotti J F, Bonato F P, 1999 "The direct influence of the neighboring stripe in White's effect" *Investigative Ophthalmology & Visual Science* 40(4) S747
- Howe P, 2001 "A comment on the Anderson (1997), the Todorović (1997), and the Ross and Pessoa (2000) explanations of White's effect" *Perception* 30 1023–1026
- Kelly F, Grossberg S, 2000 "Neural dynamics of 3-D surface perception. Figure-ground separation and lightness perception" *Perception & Psychophysics* 62 1596–1618
- Lier R van, Helm P van der, Leeuwenberg E, 1994 "Integrating local and global aspects of visual occlusion" *Perception* 23 883–903
- Todorović D, 1997 "Lightness and junctions" *Perception* 26 379–394
- Wouterlood D, Boselie F, 1992 "A good-continuation model of some occlusion phenomena" *Psychological Research* 54 267–277
- White M, 1979 "A new effect of pattern on perceived lightness" *Perception* 8 413–416
-

Theory versus speculation in visual perception

Howe (2001) showed that more than T-junctions are needed to explain White's effect. His example is theoretically interesting because Barton Anderson and Dejan Todorović had published strong claims about the role of T-junctions. Howe's demonstration contradicted these claims. Howe also noted that the FACADE model can explain his effect, as well as properties of the traditional White's effect.

Anderson's reply (Anderson 2001) mischaracterizes FACADE and its relationship to his own hypotheses. Perhaps this is because Anderson has not himself developed a computational vision model of these phenomena. He has proposed speculative interpretations of visual data without mechanistic realizations or computational analyses.

No vision model is complete. If a model does not explain a particular fact, it may still be correct as far as it goes. The FACADE model has qualitatively explained and quantitatively simulated more vision data than other current models. To evaluate FACADE, one needs to weigh the experimental evidence that supports each model mechanism and to ask if other models also explain these data. Anderson's discussion does not do this. Examples of these concerns will now be summarized.

Howe (2001) correctly asserts that Anderson's published hypotheses about the role of T-junctions cannot explain the Howe percept. In reply, Anderson (2001) notes that Howe's criticism is "fair enough" and points to a need to consider other, nonlocal properties. He then adds: "The simplest and most obvious answer is that the extended horizontal contours effectively transform the local T-junctions into X-junctions, which is precisely how Howe tries to explain the new effect. The question is whether this provides any conceptual leverage into the two competing theories of this illusion that Howe discusses: contrast theory, and scission theory." Anderson hereby implies that his scission concepts *can* explain the Howe percept. However, although there are T-junctions, there are no X-junctions in Howe's display. Anderson's suggestion therefore does not mechanistically clarify how the brain responds to Howe's display. Anderson also states that Howe (2001) compares contrast theory and scission theory. Howe's goal was more limited: to show that his display contradicts published T-junction claims of Anderson and Todorović. Howe also sketched how FACADE concepts of boundary and surface interaction can mechanistically explain his result, as well as other data about White's effect. These mechanisms were, moreover, derived from perceptual principles that have a neural interpretation and that were developed to explain different data than White's effect.

Anderson's classification of all models into contrast and scission models leads him to classify FACADE as a contrast model, which he distinguishes from "my scission-based account". Anderson notes, however, that FACADE is also a scission model when he writes that "the other explanation of the illusion offered by Grossberg and colleagues is similar to the scission explanation I proposed". Actually, FACADE proposed scission concepts before Anderson's 1997 article was published (Grossberg 1994), and has used them regularly in its explanations of figure-ground percepts; eg Grossberg (1997) and Kelly and Grossberg (2000). I believe that a classification of models into contrast and scission models is too coarse to reveal important differences between models.

Within FACADE, scission properties arise because of deeper design principles. One such principle concerns how perceptual boundaries and surfaces, which obey *complementary* computational rules, work together to generate a *consistent* percept due to their mutual interactions. Another key issue concerns the conflicting requirements that are needed to *recognize* partially occluded objects, including their occluded completions, yet to *see* only their unoccluded parts, except during transparency conditions. The mechanisms that realize these requirements can parse boundary groupings and surface filling-in generators to lie on different depth planes. Such boundary-surface interactions can cause a 'scission' of color or brightness [see Grossberg (1999a) for how they respond to a neon color spreading display]. These boundary and surface properties clarify properties like the size-disparity correlation, the way in which multiple spatial scales interact to code multiple depths, and the asymmetry between near and far. FACADE hereby proposes an explicit mechanistic explanation of how and why scission may sometimes occur.

In Anderson's "scission account ... the color of the stripe on which the target is embedded is removed from the target (ie attributed to an underlying layer). If the underlying layer is black, then the target should appear lighter (since some of its darkness is attributed to an underlying layer) ..." These hypotheses are unsupported by mechanistic analyses of why and how a scission operation could derive from deeper perceptual principles. Anderson also remarks that "in Howe's demonstration, there are two possible interpretations of the underlying layer" and thus "it is not particularly surprising that the illusion is abolished", without explaining why or how.

Kelly and Grossberg (2000) provided computer simulations of how two different perceptual interpretations of the White display can be generated; see figures 15-19. Anderson criticizes the simulations of our 'scission' interpretation by claiming that we "strengthen the contrast of the T-junction by hand". In fact, we explained how an attention shift, among other activity fluctuations, could induce the simulated strengthening. In Grossberg (1999b) and Grossberg and Raizada (2000), the same attentional mechanism was used to qualitatively explain and quantitatively simulate neurophysiological data about the influence of attention on the activity of cells in cortical areas V1, V2, and V4.

Anderson further asserts that the FACADE model's success depends critically on the alignment of the T-junction stems, which is not observed psychophysically. He restates this criticism in his comparison of how Kelly and Grossberg (2000) simulated White's effect and the Benary cross display. Kelly and Grossberg also simulated stratification percepts, checkerboard percepts, and Bregman-Kanizsa figure-ground separation using the same mechanisms. Earlier articles explained many more data, ranging from da Vinci stereopsis data to the Weisstein effect, 3-D neon color spreading, and correlations between Kanizsa square brightness and depth. Kelly and I noted that amodal completion may play a role in explaining White's effect, using the same perceptual grouping mechanism that Howe later used to explain his new effect. Anderson writes that "amodal completion that occurs in their model depends critically on the

alignment of the T-junction stems in White's display, ... but White's illusion is not affected when the stems are misaligned".

Anderson's assertion is incorrect. FACADE has helped to explain data about a variant of White's effect in which there is neither alignment nor even existence of T-junctions. This explanation uses the same mechanisms that were used to explain White's effect data where alignment of T-junction stems does occur. In this latter case, amodal completion seems to play a role, as Anderson acknowledges in his remark about how Howe's display functionally transforms T-junctions into X-junctions. Although there are no explicit X-junctions in the Howe display, they can be implicitly formed if the brain uses amodal completion, which leads to the type of 'extended horizontal contours' that Anderson seems to accept in his reply to Howe. FACADE provides a principled mechanistic account of when and how such horizontal contours can be formed.

Why does Anderson gloss over these properties of FACADE? Two words may summarize the problem: mechanism and context. Anderson seems unable to infer how FACADE mechanisms may respond to different displays. Each display generates a different perceptual context and thus a potentially different perceptual outcome. FACADE mechanisms reflect this context-sensitivity. For example, in FACADE, T-junctions do not act as a depth cue as such. Rather, they may initiate boundary-breaking in response to which interacting boundary and surface mechanisms *sometimes* yield depthful figure-ground separation, depending on the context, as Todorović (2001) has nicely illustrated.

FACADE is an emerging perceptual theory that offers unified explanations and predictions of many data about visual perception and its brain substrates. The process of building, and experimentally testing, such a theory requires that one balance multiple experimental and theoretical constraints, and that one understand how a few interacting mechanisms may explain many data because their emergent properties are highly context-sensitive. Anderson's critique is misinformed about the properties of FACADE and consequently offers an unbalanced evaluation of its explanatory value.

Acknowledgements. Supported in part by the Air Force Office of Scientific Research (F49620-01-1-0397), the Defense Advanced Research Projects Agency, and the Office of Naval Research (ONR N00014-95-1-0409; ONR N00014-95-1-0624).

Stephen Grossberg

Department of Cognitive and Neural Systems and Center for Adaptive Systems, Boston University, 677 Beacon Street, Boston, MA 02215, USA; e-mail: steve@bu.edu; web: <http://www.cns.bu.edu/Profiles/Grossberg>

References

- Anderson B L, 1997 "A theory of illusory lightness and transparency in monocular and binocular images: the role of contour junctions" *Perception* **26** 419–453
- Anderson B L, 2001 "Contrasting theories of White's effect" *Perception* **30** this discussion
- Grossberg S, 1994 "3-D vision and figure-ground separation by visual cortex" *Perception & Psychophysics* **55** 48–120
- Grossberg S, 1997 "Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures" *Psychological Review* **104** 618–658
- Grossberg S, 1999a "A comment on 'Assimilation of achromatic color cannot explain the brightness effect in the achromatic neon effect' by Marc K Albert" *Perception* **28** 1291–1302
- Grossberg S, 1999b "How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex" *Spatial Vision* **12** 163–186
- Grossberg S, Raizada R D S, 2000 "Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex" *Vision Research* **40** 1413–1432
- Howe P D L, 2001 "A comment on the Anderson (1997), the Todorović (1997), and the Ross and Pessoa (2000) explanations of White's effect" *Perception* **30** 1023–1026
- Kelly F, Grossberg S, 2000 "Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception" *Perception & Psychophysics* **62** 1596–1618
- Todorović D, 2001 "Lightness, junctions, and depth" *Perception* **30** this discussion