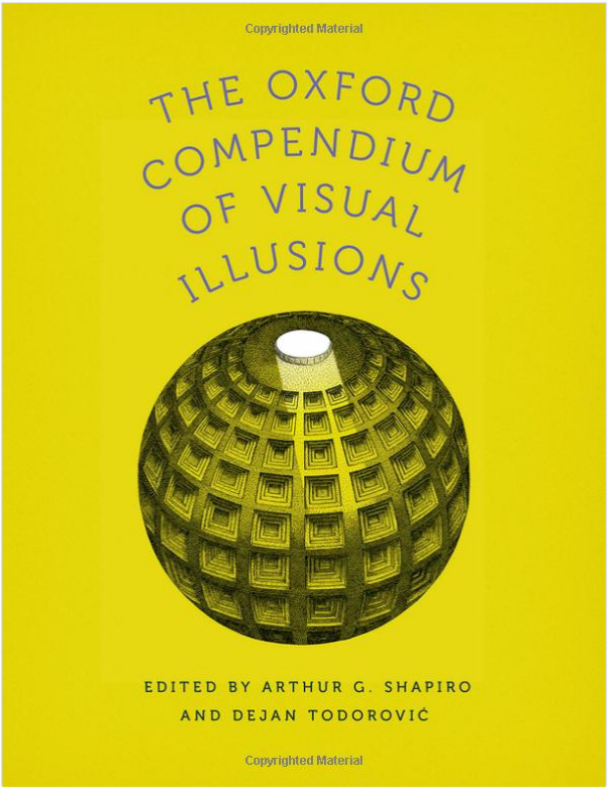
Chapter 96. The Scramble Illusion: Texture Metamers

C. Chubb, J. Darcy, J., M. S. Landy, J. Econopouly, J-H. Nam, D. Bindman, & G. Sperling

in



THE OXFORD COMPENDIUM OF VISUAL ILLUSIONS

EDITED BY ARTHUR G. SHAPIRO
AND DEJAN TODOROVIĆ

Chubb, C., Darcy, J., Landy, M. S., Econopouly, J., Nam, J-H, Bindman, D., & Sperling, G. (2018). The Scramble Illusion: Texture Metamers. In Shapiro, A. & Tedorovic, D. (Eds). The Oxford Compendium of Visual Illusions. New York, NY: Oxford University Press. Chapter 96, Pp. 668-672.

Chapter 96
The Scramble Illusion
Texture Metamers

*Charles Chubb, Joseph Darcy, Michael S. Landy, John Econopouly,
Jong-Ho Nam, Dan Bindman, and George Sperling*

Challenge: Take a look at Figure VI.96-1. Can you explain what makes this figure interesting? This is a question that we have posed to many audiences since we first discovered the scramble illusion. No one has ever come up with the right answer.

As we shall show, the phenomenology of visual texture perception is strangely misleading. Wherever we look within a field of texture, we can see all the details of the texture filling that region. For example, when we look at the texture shown in Figure VI.96-1, we can see the specific gray levels of the squares in whichever region we choose to inspect. This gives us the impression that when we view a texture, our eyes and brain take a "visual inventory" of the texture and that this inventory defines the impact exerted on our vision by the texture. Under this naïve theory, two textures should look different if they differ significantly in the relative proportions of the various features they contain.

That the visual-inventory of features theory is false is dramatized by Figure VI.96-2. The left and right sides of Figure VI.96-2 are filled with two different types of texture called "scrambles." A scramble is composed of small squares of different gray levels mixed together in specified proportions and randomly arranged in space. The scrambles shown in Figure VI.96-2 appear to be homogeneous. However, careful scrutiny reveals that the scrambles on the left and right sides differ dramatically in ways that human viewers utterly fail to notice without explicit prompting. The actual histograms of the left and right sides in Figure VI.96-2 are shown in the inset bar graphs. The histograms give the proportions of different gray levels in the two scrambles. As shown by the right-hand bar graph, the scramble filling the right half of Figure VI.96-2 contains equal proportions of 17 different gray levels that increase from black to white. By contrast, the left-hand bar graph shows that the scramble filling the left side of this figure consists of a mixture of just three gray levels whose proportions have been carefully balanced to achieve several conditions that are described later.

Now that we have seen the two different scrambles in Figure VI.96-2, take another look at Figure VI.96-1. This figure also comprises two different scrambles of the same sorts as are used in Figure VI.96-2. Can you find the boundary between them?

THE SCRAMBLE ILLUSION AND PREATTENTIVE VISION

Figures VI.96-1 and VI.96-2 tell us a great deal about human preattentive vision. The word "preattentive" refers to those processes that operate, prior to any effort of attention on the part of the viewer, to make manifest the qualitative differences in the visual input that exist across space. White paint on a dark red background produces a visible difference that we cannot fail to see. Any difference of this sort that our vision reveals to us immediately and spontaneously is considered to be preattentive. It is clear that human preattentive vision is not at all sensitive to any of the physical differences that exist between the scrambles on the left-hand versus the right-hand side of Figure VI.96-2. We assume that in order for human vision to be preattentively sensitive to some particular physical property, there must exist a class of neurons operating in human preattentive vision that are differentially activated by variations in this statistic. In other words, any difference that is immediately evident to human vision implies the existence of some population of neurons sensitive to that difference. We take as a working hypothesis that preattentive human vision is devoid of any neurons sensitive to any of the many physical differences that exist between the right and left sides of Figure VI.96-2.

For example, we can immediately rule out the possibility that human preattentive vision contains a class of neurons that is specifically sensitive to any of the three gray levels that are used to make up the scramble on the left side of Figure VI.96-2. Suppose, in particular, that human preattentive vision contained a retinotopically organized array of neurons each of which was highly activated by gray level #15 in our set (third from right) and not at all by any other gray levels. The output from such an array would compose a "neural image" (Robson, 1980) providing

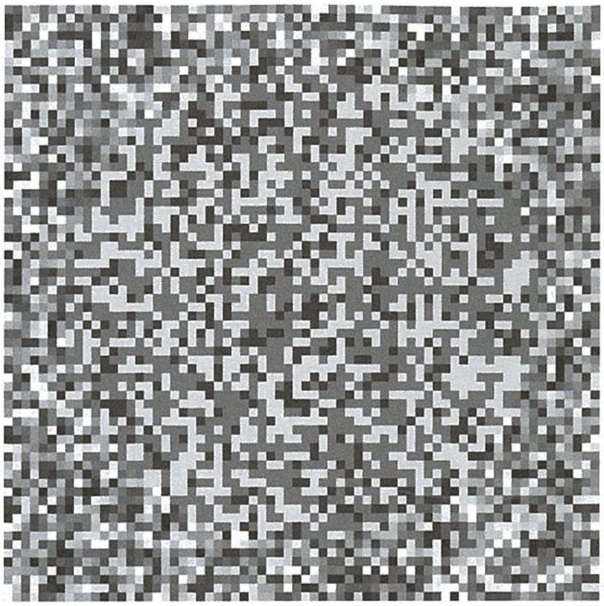


Figure VI.96-1. Can you guess what makes this figure interesting?

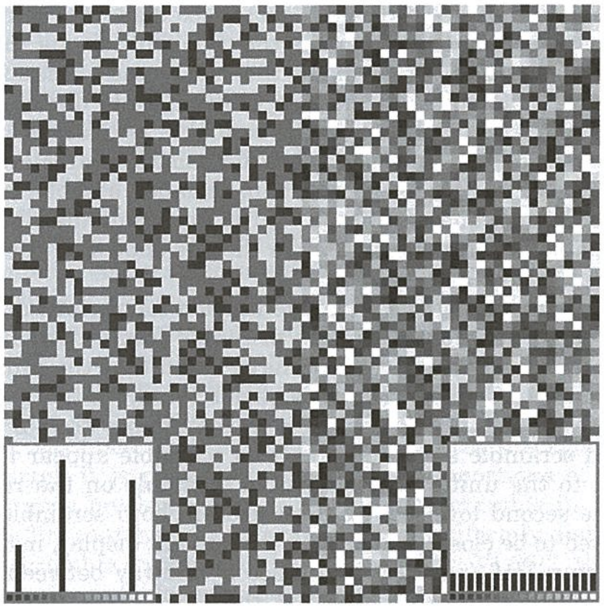


Figure VI.96-2. Two adjacent scrambles separated by a vertical boundary. The histogram insert at the lower left indicates the relative proportions of the three gray levels in the left half of the figure. The histogram insert at the lower right indicates that all 17 gray levels in the right half of the scramble have equal proportions.

a spatial map of gray level #15. Using black to code no activation and white to code the activation produced by gray level #15, the neural image produced in this array by Figure VI.96-2 would look like Figure VI.96-3. The large difference in average activation between the left and right sides of Figure VI.96-3 indicates that if human vision did possess such a neural array tuned specifically to gray level #15, then we would immediately see the difference between the left and right sides of Figure VI.96-2. We conclude that human preattentive vision contains no such neural array.

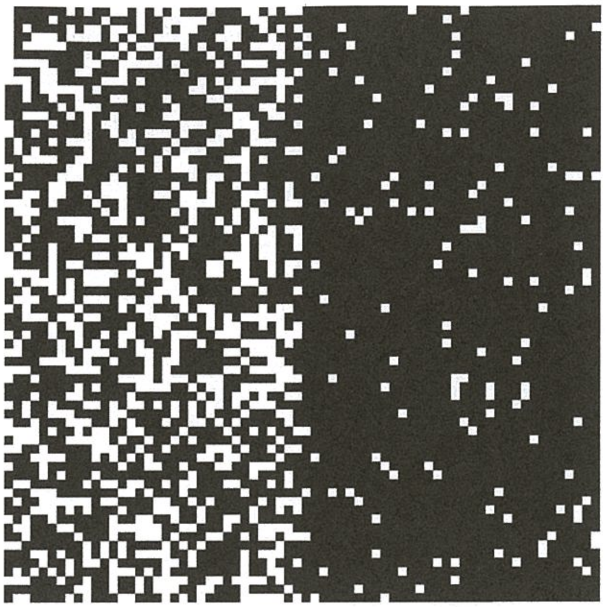


Figure VI.96-3. The scrambles of Figure VI.96-2 with gray level 15 (almost white) represented as white and all other gray levels represented as black.

Figures VI.96-1 and VI.96-2 also make it clear that human vision has little or no preattentive sensitivity to the *entropy* of the probability distribution characterizing a scramble. The entropy of a probability distribution p reflects the overall "randomness" of p. Specifically, the entropy of p is the expectation of $-\log_2(p[X])$, where X is a random variable with distribution p. Thus if p assigns probability 1 to a particular value (and zero to all others; i.e., a uniform field), then p's entropy is zero. If p assigns equal probability to each of 2^n values (and zero to all others), then p's entropy is n. The entropy of the uniformly distributed scramble on the right in Figure VI.96-2 is 4.09 whereas the entropy of the three-gray-level scramble on the left is 1.48. Thus if human vision were preattentively sensitive to the entropy of a scramble's distribution, there should be a strong preattentive difference between the scrambles on the left and right in Figure VI.96-2.

To take a third example, it has been proposed that human vision contains a "boundary contour" system that registers the locations of boundaries between homogeneous regions (Grossberg & Todorović, 1988). This system is hypothesized to play a central role in the process through which brightnesses are assigned to homogeneous regions in the visual field. Suppose that human vision does indeed possess such a boundary contour system and that this system registers the existence of a boundary with a degree of strength that is invariant with respect to the contrast difference across the boundary.[1] In this case, the response of the boundary contour system to Figure VI.96-2 should resemble the image shown in Figure VI.96-4. The fact that we easily see the difference between the right and left sides of Figure VI.96-4 due to the greater number of edges on the right suggests that if human vision does embody a boundary contour system, either this system does not operate as we have hypothesized here or else the output of this system is not used to inform preattentive vision.

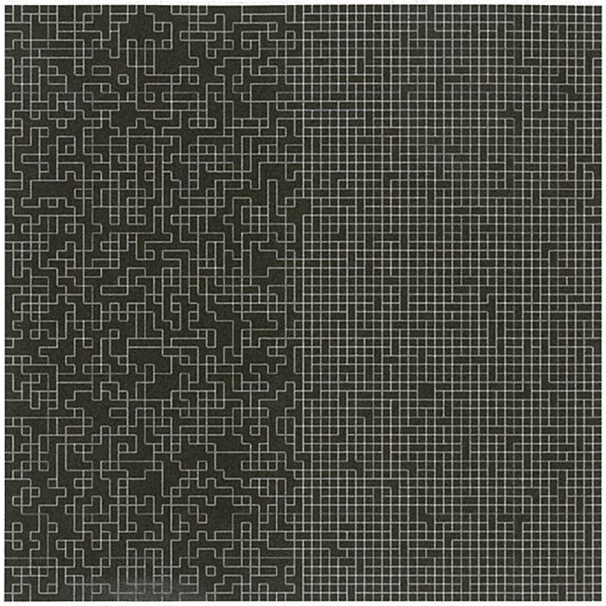


Figure VI.96-4. The boundary map of the scrambles of Figure VI.96-2. Every white line in this figure marks a change in gray level between homogeneous regions in the scrambles of Figure VI.96-2.

HISTORY AND EXPLANATION

Figure VI.96-1 was prompted by the experiments of Chubb, Econopouly, and Landy (1994) who investigated preattentive discrimination of gray-level scrambles. In their experiments, participants judged the orientation (vertical vs. horizontal) of a briefly presented square wave whose bars were defined by scrambles with different histograms. In particular, all the scrambles used by Chubb et al. had mean and variance equal to those of the uniformly distributed scramble shown on the right in Figure VI.96-2; scramble histograms varied only in higher-order moments. They found that for scrambles in this restricted class, sensitivity to the orientation of the scramble-defined square wave could be understood in terms of a single preattentive mechanism that was differentially sensitive to very dark pixels, subsequently called the *blackshot mechanism.* Chubb et al. (1994) provided a partial characterization that sufficed to predict the differential activation of the blackshot mechanism by scrambles in the class to which their study was restricted. Chubb, Landy, and Econopouly (2004) later extended these results to complete the characterization of the blackshot mechanism.

The function that gives the sensitivity of the blackshot mechanism to different gray levels is plotted for three observers in Figure VI.96-5. As these plots make clear, the blackshot mechanism is highly activated by squares with Weber contrasts near –1 (i.e., with gray levels near black) but discriminates very poorly between texture elements with Weber contrasts greater than –0.95.

The results of Chubb et al. (1994) suggested that if two scrambles were equated in mean, variance and blackshot (although the blackshot sensitivity function was not fully characterized until 2004), then they should be preattentively indiscriminable. It was this observation that led us (Chubb, Darcy, & Sperling, 1993) to hunt down the scramble shown on the left-hand side in Figure VI.96-2. Specifically, we conducted a brute force search through the space of scrambles made of only three gray levels to see if we could find such a three-gray-level scramble with mean, variance, and blackshot equal to those of the uniformly distributed scramble.

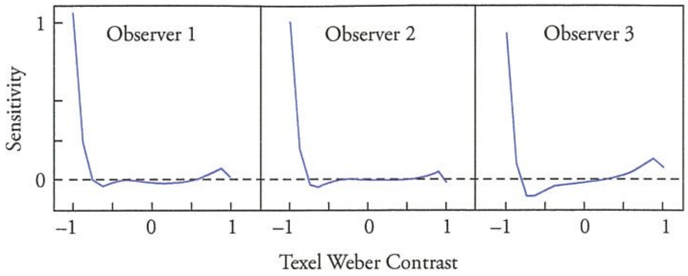


Figure VI.96-5. Sensitivity of the blackshot texture mechanism as a function of the Weber contrast of the scramble texture elements. (From Chubb et al., 2004).

It was not a foregone conclusion that such a scramble would exist. For any fixed set of three gray levels, once one assigns probabilities to two of the gray levels, the probability of the third gray level is fixed as well. For the same reason, once one adjusts the probabilities of a three-gray-level scramble to make the mean and variance of the scramble equal to those of the uniformly distributed scramble, the blackshot of the scramble is fixed as well. As it turned out, however, there did exist a scramble with only three gray levels whose mean, variance, and blackshot were nearly equated to those of the uniformly distributed scramble, a scramble similar to the one shown on the left side of Figure VI.96-2.

In reproducing the scramble illusion for various different demonstrations—some in printable format, others on different display devices—we have found that the illusion is easier to achieve (i.e., it is easier to adjust the probabilities of the three gray levels in the Figure VI.96-2 left-hand scramble so as to make this scramble appear identical to the uniformly distributed scramble on the right) if the second lowest gray level used in both scrambles is shifted to be closer to black. In the current display, instead of being 1/16 (which would place it midway between the first and the third gray levels), the second gray level has been set to 1/32. In the scramble on the left, all gray levels have probability zero except the second, the seventh, and the fifteenth, which have probabilities 0.1642, 0.4517, and 0.3841, respectively.

Given that we arrived at these probabilities by adjusting the proportions of gray levels numbers 2, 7, and 15 so as to make the 3- and 17-gray-level scrambles as similar in appearance as possible, it is reasonable to ask: How do the mean, variance, and blackshot of the 3- and 17-gray-level textures compare? The answer is that the 3-gray-scale scramble has 2.6% higher mean, 2.1% higher variance, and approximately 4.3% higher blackshot than the 17-gray-level scramble.

These disparities are not unexpected. Although we know that the blackshot mechanism is very sharply tuned for the darkest scramble elements, the characterization of the sensitivity function provided by Chubb et al. (2004) is a seventh-order polynomial approximation that may not be perfectly accurate across the range of gray levels near black where the function is steep. We therefore cannot rule out the possibility that the disparity in blackshot between the 3-gray-level and 17-gray-level scrambles may be due to error in our estimate of the blackshot sensitivity function. The disparities between the means and variances of the 3- and 17-gray-level scrambles, although slight, are also consonant with previous results. As shown by Chubb, Nam, Bindman, and Sperling (2007), the preattentive sensitivity of human observers to scramble mean and variance is conferred by mechanisms that are actually sensitive to statistics that differ slightly from mean and variance.

In color theory, two patches that appear identical in color but have different wavelength compositions (different spectra) are called metamers. By analogy, two scrambles that appear preattentively identical but that have different gray-level histograms are texture metamers. Human color vision is three-dimensional. This implies that any color can be matched with an appropriate combination of three primary colors. Tridimensionality is the basis of color photography and color video. Similarly, we (Chubb et al., 2007) have found that preattentive human vision for scrambles of the type shown here is three-dimensional (excluding scrambles whose histograms have very low entropy). The analogy to color thus suggests that one should be able to match the appearance of any scramble by taking an appropriate mixture of some fixed set of three well-chosen gray levels. However, this is not quite true: as soon as one assigns the proportions of the first two gray levels, the proportion of the third gray level is also determined (because the probability distribution characterizing any scramble is constrained to sum to 1). This means that with only three gray levels, one does not have enough freedom to construct a scramble that matches a given scramble in all three of mean, variance, and blackshot. In order to ensure the ability to obtain such a match, a fixed set of not three but rather *four* gray levels is required to match the mean, variance, and blackshot of an arbitrary scramble. (Although it is possible to adjust the proportions of the three gray levels used in the left-hand scramble in Fig. VI.96-2 to match the appearance of the uniformly distributed scramble on the right-hand side of Fig. VI.96-2, there are many scrambles whose mean, variance, and blackshot cannot be matched by adjusting the proportions of these same three gray levels). In color metamers, no amount of selective attention will reveal the difference between metamers because there are only three kinds of color-sensitive receptors (retinal cones). In scramble perception, selective attention can bring into play visual processes other than the mechanisms typically used in texture perception. That is, selective attention to internal details of a texture is not a typical texture-perception process. However, for monochromatic scrambles that are perceived preattentively, texture perception is a three-dimensional process analogous to the three-dimensional perception of color patches.

We can reveal the physical difference between the scrambles in Figure VI.96-1 by applying an appropriate transformation of gray level. To see what is really going on in Figure VI.96-1, we apply a transformation that maps gray level #2 onto the value 0.83, gray level #7 onto 1.0, gray level #15 onto 0.98, and all other gray levels onto zero. This particular function has been chosen to maximize the signal-to-noise ratio in the output image between the texture element values in the region of the 3-gray-level scramble versus the region of the 17-gray-level scramble. The resulting image is shown in Figure VI.96-6. This function maps regions filled with the 3-gray-level scramble onto regions of high average value and regions filled with the 17-gray-level scramble onto regions of low average value.

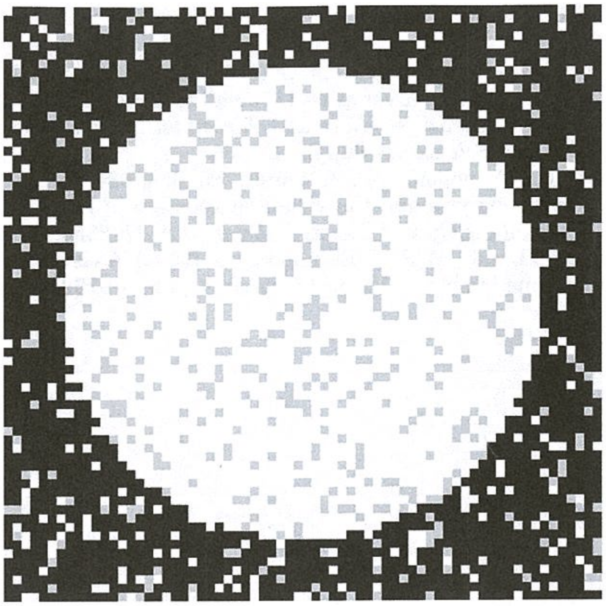


Figure VI.96-6. The organization of the 3- and 17-level scrambles in Figure VI.96-1 is revealed by applying a function to Figure VI.96-1 that maps gray level #2 onto the value 0.83, gray level #7 onto 1.0, gray level #15 onto 0.98, and all other gray levels onto zero. This function maximizes the signal-to-noise ratio in the output image between the texture element values in the region of the 3-gray-level scramble versus the region of the 17-gray-level scramble. As this transformation shows, Figure VI.96-1 is composed of a central disk of the 3-gray-level scramble set against a background of the 17-gray-level scramble.

The general point illustrated by the scramble illusion is that we cannot always see what is visible, even when we look directly at it. The physical differences between the scrambles on the left and right sides of Figure VI.96-2 are visible: When asked how many gray levels are used in the scramble on the left, we easily see that the answer is three; and when asked how many are used in the scramble on the right, we quickly affirm that the answer is substantially more. Nonetheless, we do not notice these differences without explicit prompting, which raises an interesting question: What else are we missing that is right before our eyes?

NOTE

1. In the simulations presented by Grossberg and Todorović (1988), the boundary contour system does not, in fact, have this property. Rather, boundary strength increases with the contrast difference on the two sides of the boundary.

REFERENCES

Chubb, C., Darcy, J., & Sperling, G. (1993). Metameric matches in the space of textures comprised of small squares with jointly independent intensities [Abstract]. *Investigative Ophthalmology and Visual Science*, *34*(Suppl. 4), 1289.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A*, *11*, 2350–2374.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research*, *44*, 3223–3232.

Chubb, C., Nam, J.-H., Bindman, D. R., & Sperling, G. (2007). The three dimensions of human visual sensitivity to first-order contrast statistics. *Vision Research*, *47*, 2237–2248.

Grossberg, S., & Todorović, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics*, *43*, 241–277.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris (Ed.), *Visual coding and adaptability* (pp. 177–214). Hillsdale, NJ: Lawrence Erlbaum.