BRILL *Art & Perception (2020)* DOI: 10.1163/22134913-bja10002 brill.com/artp

When Scenes Look Like Materials: René Magritte's Reversible Figure–Ground Motif

J. Brendan Ritchie[1] **and Benjamin van Buren**[2]

[1]Laboratory of Biological Psychology, Brain and Cognition Unit, KU Leuven, Leuven, Belgium

[2]Department of Psychology, The New School, New York, NY, USA

Received 9 September 2019; accepted 11 February 2020

Abstract

We draw attention to a frequent motif in the work of the Belgian surrealist René Magritte (1898–1967). In the motif, a scene is depicted that contains a silhouette, which itself contains another depicted scene. The silhouette is bistable, appearing either as a figural region whose positive space is covered, or filled, with the interior scene texture, or as a ground region providing a window onto a more distant scene. We call this the 'reversible figure–ground motif'. Because the stimulus does not change when our percept changes, the motif's appearance at any particular moment cannot be explained by its local or global image statistics. Instead principles of perceptual organization, and in particular image segmentation and figure–ground assignment, appear crucial for determining whether the interior of the silhouette is processed as a material vs. a scene — which in turn reflects the fundamental role of visual segmentation in material and scene perception more generally.

Keywords

René Magritte, figure–ground segmentation, materials, scenes

1. Introduction

We are surrounded by curtains. – (Magritte, 1979/2016, p. 214).

The Belgian surrealist René Magritte (1898–1967) often suggested that we see the world as if through a flat curtain. His observation highlights what many vision scientists consider to be a central function of visual perception — transmuting protean and flat retinal images into experiences of stable and

* E-mail: j.brendan.w.ritchie@gmail.com

 DOI: 10.1163/22134913-bja10002

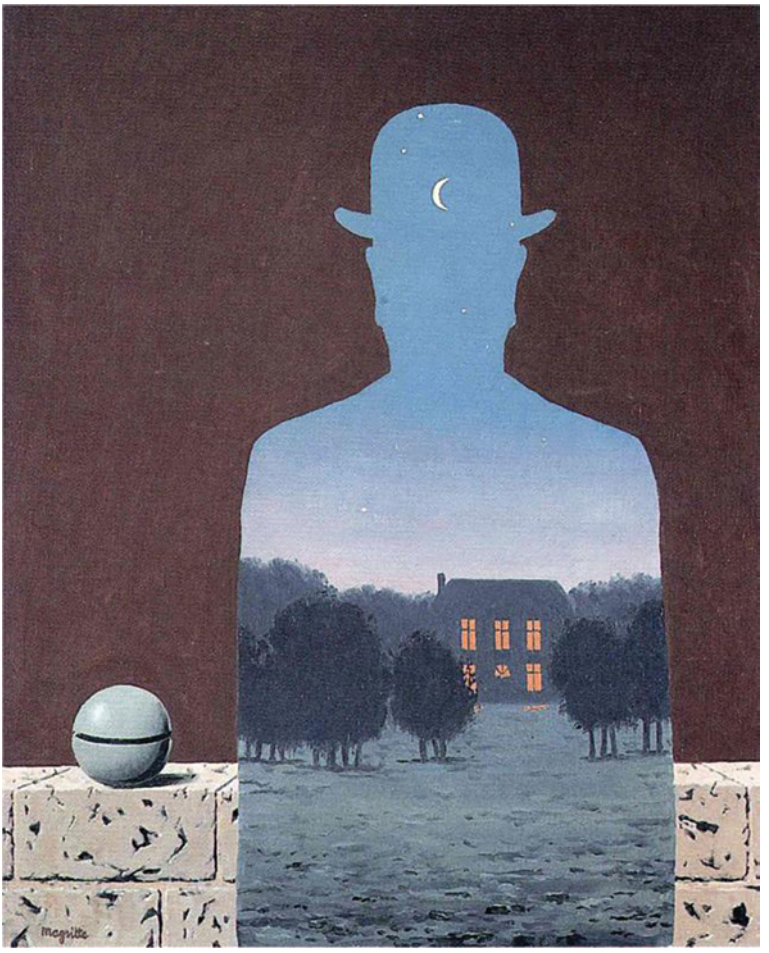

Figure 1. Magritte's *The Happy Donor* (1966). Here the very same image can induce bistable percepts of a material or a scene. What causes the central region to look like a material vs. a scene is the assignment of figure or ground status to that region. In other words, here we have a situation where global image-level analysis is clearly not sufficient to explain either material or scene perception.

three-dimensional material objects and scenes. However, a radically different approach holds that aspects of material and scene perception occur without assigning depth relations to different parts of the image (i.e., without segmenting the input into figure and ground regions). These models instead treat images as textures, and visual discrimination and categorization as image processing problems over global or local statistics (e.g., Motoyoshi *et al.*, 2007; Oliva and Torralba, 2001; Orlandi, 2014; Shrivastava *et al.*, 2015). Although image statistics models are good at mimicking some aspects of human

perception, a recurring motif in Magritte's paintings shows that they cannot fully explain either material or scene perception.

In the motif, a scene image contains a silhouette, which contains another scene image (see Fig. 1). The silhouette is bistable, appearing either as a figural region whose positive space is covered, or filled, with the interior scene texture, or as a ground region providing a window onto a more distant scene. This 'reversible figure–ground motif' demonstrates that the perception of an image region as a material vs. a scene cannot be explained by its image statistics, since the image statistics stay the same as our percept changes. Rather, we must appeal to classical concepts of perceptual organization (segmentation and figure–ground assignment) to explain a fundamental aspect of material and scene perception — namely, whether a region appears as a material or a scene in the first place.

2. The Reversible Figure–Ground Motif

In our discussion we will focus mainly on a single work, *The Happy Donor* (1966), although one can find many distinct and interesting variants (as well as outright repetitions) of the motif in Magritte's oeuvre. The painting depicts the silhouette of a person in a bowler hat against a burgundy background (see Fig. 1). At the bottom of the canvas is a small wall of stone bricks, on top of which rests a silver bell. The interior of the silhouette depicts an evening scene: a house with a lit interior rests among some trees in a wooded countryside, with the moon visible in the sky. Other highly similar compositions from Magritte's later period, which use the same silhouette, include *A Friend of Order* (1964) and *The King's Museum* (1966).

The Happy Donor, like all versions of the motif, has two key components: a contour (often constituting the silhouette of a familiar object), and two scene images in the regions on either side of the contour. The composition invites two different percepts: (1) The silhouette appears as a 'figure' that is covered or filled with the interior scene image. When this percept is generated, the interior scene image looks relatively flat. It may appear as a painted board, or as a drapery over a 3D volume (which is subtly implied by the region's convex shape), or as that volume's spatially extended filling. In all these cases, assigning this region the status of figure causes us to perceive it as some kind of material. (2) If one focuses on the details of the interior scene, the silhouette appears instead as a window onto the countryside. In this 'ground' interpretation, the interior scene image's depth cues become more salient, and instead of appearing confined to a person-sized space, the scene appears to extend in all directions behind the picture plane. The figure and ground percepts are bistable, in a manner similar to familiar illustrations of figure–ground reversal, such as Rubin's Vase (Rubin, 1915/1958). However, different observers may

find one or the other percept more natural, and switching the interpretation of the image may depend on where one attends. For example, we find that attending to occlusion cues where the silhouette meets the stone wall facilitates the figure percept, while attending to the house tends to promote the ground percept.

3. Explaining Our Perception of the Motif

Periodically vision scientists have drawn attention to similar motifs from Magritte's work. For example, Kanizsa (1985) and Pinna (2007) both provide brief discussions of *The Blank Check* (1965), in which the image of a person on horseback is intertwined with those of trees in a forest, to illustrate properties of amodal completion. However, these discussions have largely glossed over what we take to be the most interesting feature of the figure–ground motif — namely its bistability, and the radical consequences that this has for whether a given image region appears to be a material or a scene. We next sketch some factors that contribute to the figure vs. ground percepts in works like *The Happy Donor*.

The silhouettes used to generate the motif tend to have many of the classic properties that produce a figure percept, such as convexity, small area, enclosure, and symmetry (Harrower, 1936; Rubin, 1915/1958; for a review see Wagemans *et al.*, 2012). They are also typically of familiar objects, which further contributes to the figure interpretation (Peterson and Gibson, 1994). For example, the pigeon-shaped region in Fig. 2A has all of these properties and is readily seen as figure against the blue background. In Fig. 2B, with the addition of some new visual cues within the silhouette, a ground interpretation also becomes possible. That is, while we may see a pigeon-shaped figure filled or painted with a cloudy sky pattern, occlusion cues on the inside of the contour allow this region to also be seen as ground, and it looks highly bistable — compare to *The Return* (1940). In Fig. 2C, which introduces a scene image with stronger depth cues and more pronounced T-junctions within the contour, the ground percept is more dominant — compare to *The Plagiarism* (1940) and *The Flash* (1959). At the same time, the figure percept is still available if when viewing one focuses not on the central tree, but rather on the familiar bird shape of the contour itself.

A ground interpretation seems to be more likely when the interior scene image contains a central object that attracts attention and increases the processing of depth cues within this region. This effect may be an exception to the rule that 'accentuating' (i.e., drawing attention to) regions of an image generally causes them to be assigned figure status. For example, in Fig. 3A, the accentuating red dots make us more likely to see the central black region as figure (Pinna *et al.*, 2018). In Fig. 3B, the same exact contours are depicted,

Figure 2. Seeing the interior region as figure depends on whether it is filled with (A) a solid color, (B) a natural scene image with limited depth cues (a cloudy sky), or (C) a natural scene image with pronounced depth cues (a winter landscape). Note: readers may find the differences between these panels more salient when viewing them at a larger scale (e.g., increasing viewing zoom to 200%).

Figure 3. (A) The presence of red accentuating dots on Rubin's Vase makes us more likely to see the black region as figure (i.e., as a vase). (B) A variant using natural scene images. Here the Moon and its reflection, although positioned similarly to the red dots, seem to instead draw us into the scene and bias us toward seeing this region as ground.

but here the moon and its reflection in the water at night seem to 'draw us in' to the central region, resulting in a stronger bias to see this region as ground.

In the above examples, assigning figure status to a natural scene image causes us to experience it as having new material properties. For example, a picture of a countryside scene might appear as a flat board with paint on it, or as the skin/filling of a figural volume, which bulges toward us when we attend to subtle metric depth cues at its contour (Burge *et al.*, 2010). Figure 4 contains another example in which assigning figurehood to a region causes us to imbue it with new material properties. In this case, the top beach scene appears to 'drip' down over the lower beach region — compare to *The Muscles of the Sky* (1928) and *The Art of Conversation* (1950), variations III and IV.

Looking more closely at the top region of Fig. 4, it is clear that many factors contribute to its perceived material. First, of course, there is the assignment of figure/ground status — this region appears as a viscous liquid only when it is perceptually organized as figure. Also important is the shape of the contour, which explains, for example, why we see the upper region as viscous rather than runny. In fact, figure–ground assignment and contour shape are not independent factors. Contour shape influences figure–ground assignment (e.g., convex regions tend to be seen as figure), and figure–ground assignment in turn determines which region 'owns' the contour (e.g., which of the two regions in Fig. 4 is seen as drippy, and which is seen as occluded). Additionally, there are cues to material composition within the region itself, such

Figure 4. Image depicting the impact of contour shape and figure–ground assignment on the perceived material properties of a scene image.

as its luminance histogram, which signals the presence of a beach and water (Motoyoshi *et al.*, 2007). And finally, there are interactions between shape and texture — what is sometimes referred to as a material's 'habit' (Adelson, 2001). Most significant for the present discussion is how depth cues provided by the contour influence the perceived spatial layout of the scene image.

As we have alluded to, in some instances of Magritte's motif a scene image that is organized as figure may appear to bulge outward. Such impressions may reflect warping of the scene's perceived spatial layout based on the convex shape of its bounding contour, which itself provides cues to depth. For example, Fig. 5 depicts one forest image within another forest image. The smaller/more interior forest image appears as a figural region, and due to its convex shape we see it filling an egg-shaped object that is reflective or transparent — even though the interior region does not contain the spectral distortions typical of these materials (Fleming *et al.*, 2004, 2011a). Thus, both the perceived spatial layout and the material properties of a scene image depend on figure–ground assignment, and on metric depth cues in its bounding contour.

Figure 5. Two superimposed images of the woods to the south of Brussels. The central scene image is perceived as occupying a 3D volume evoked by the convex contour. Here we see the importance of ordinal and metric depth cues in determining the scene's perceived spatial layout, as well as its appearance as a material (in this case, a reflective or transparent material such as glass, or water).

In Magritte's bistable motif, the assignment of figure vs. ground status to an image region (and hence its appearance as a material vs. a scene) may often depend on whether we attend to metric depth cues in the region's convex contour (which promote the perception of a figural volume), or rather to depth cues within the image (which promote its being perceived as a ground region). Moreover, actively switching between these material and scene interpretations is likely an important part of the pleasure of viewing this motif in Magritte's work (Jakesch, Leder and Forster, 2013; Muth and Carbon, 2013; Muth, Hesslinger and Carbon, 2018; Van de Cruys and Wagemans, 2011).

4. The Motif and Image Statistics

Some vision scientists have noted that materials and scenes can both be thought of as species of image textures, and usefully analyzed using techniques from computer vision that extract global or local statistical properties of natural images. Within both of these domains, coarse image statistics turn out to be all one needs to perform certain categorization tasks — e.g., deciding

whether an image of a material is glossy or not (e.g., Motoyoshi *et al.*, 2007; Fleming *et al.*, 2011b), or deciding whether an image of a scene is of a beach or a forest (e.g., Oliva and Schyns, 2000; Oliva and Torralba, 2006; Torralba and Oliva, 2003). Such results have been taken to support a radical view of visual processing more generally, according to which image segmentation is not a necessary precursor to many aspects of seeing (Orlandi, 2014). However, the reversible figure–ground motif makes it clear that, while image statistics may in some cases be sufficient for categorizing an image as one of several types of materials, or for categorizing it as one of several types of scenes, they are wholly insufficient to explain whether we discriminate parts of the visual world as scenes or materials *in the first place*. After all, in paintings like *The Happy Donor*, no change in image statistics determines whether we switch from seeing a region as a scene to seeing it as a material. Here, explaining what we see requires an appeal to non-image-based visual representations — i.e., to classical notions of segmentation, perceptual organization, and distal representation.

The conclusion that we draw from the reversible figure–ground motif complements other recent studies, which have also highlighted the insufficiency of global image statistics for certain types of visual judgments (Anderson and Kim, 2009; Kim *et al.*, 2016; Olkkonen and Brainard, 2011). For example, changing the bounding contour of a grating pattern (while preserving its internal image statistics) can determine whether it is perceived as matte or metallic (Marlow and Anderson, 2015; Marlow *et al.*, 2015). And we see similar effects in Magritte's paintings, where scene images take on material appearances when they are organized as figure, with their particular material qualities (e.g., the degree to which they may look 'viscous' or 'stretched') determined largely by metric depth cues in the region's contour. Consider the viscous appearance of the upper region in Fig. 4, or the glassy appearance in Fig. 5 — in these examples shape cues make all the difference.

To understand either material or scene perception, we need to first explain how images are organized into figure and ground regions. One promising approach holds that the visual system performs this segmentation based on statistical regularities between flat images and states of the world (Brunswik and Kamiya, 1953). For example, studies of natural images have found that smaller, more convex, and lower image regions, which all tend to be perceived as figure, *really do* tend to correspond to ordinally closer real-world surfaces (Fowlkes *et al.*, 2007; see also Burge *et al.*, 2010 for a similar demonstration involving metric depth cues in natural images). Given these findings, it is notable that a silhouette such as the pigeon in Fig. 2C, which has most of the classical cues to figurehood, is still more readily seen as ground when it is filled with a scene image that contains strong depth cues. However, this perceptual outcome could reflect another empirical regularity: when a scene image is bounded within another scene image, this may most often reflect a

ground region viewed through an aperture such as a doorway or window. And accordingly, we see that doorways and windows feature prominently in many versions of Magritte's motif — e.g., *Portrait of Germaine Nellens* (1962).

5. Conclusion

For Magritte, painting was a way of depicting thoughts, each canvas an intellectual experiment (Magritte, 1979/2016; Paquet, 2000). These experiments were carried out to solve 'problems' presented by phenomenal objects. In the case of the window, this resulted in *The Human Condition* (1933/1935), which depicts a painting in front of a window opening onto the same landscape that appears in the painting. This results in a disconcerting conflict between one perceptual interpretation, of the image in the painting and the image in the window corresponding to different depth planes, and another interpretation, of their reflecting a continuous distal surface. As observed by Magritte (1979/2016, p. 65), "For the viewer, the tree was simultaneously inside the room in the picture and outside in the real landscape". Perhaps we can learn something from Magritte's persistent focus on the potential 'treachery of images', and from his recurring insight that the appearance of an image is largely determined by how we as viewers parse it.

Acknowledgements

This project was supported by a long-term structural funding from the Flemish Government (METH/14/02) to Johan Wagemans. This project has received funding from the FWO and European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501, via a FWO [PEGASUS]^2 Marie Skłodowska-Curie fellowship (12T9217N) to J.B.R.

References

Adelson, E. H. (2001). On seeing stuff: The perception of materials by humans and machines, *Proc. SPIE* **4299**, 1–12.

Anderson, B. L. and Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness, *J. Vis.* **9**, 10. doi:10.1167/9.11.10.

Brunswick, E. and Kamiya, J. (1953). Ecological cue-validity of 'proximity' and of other Gestalt factors, *Am. J. Psychol.* **66**, 20–32.

Burge, J., Fowlkes, C. C. and Banks, M. S. (2010). Natural-scene statistics predict how the figure–ground cue of convexity affects human depth perception, *J. Neurosci.* **30**, 7269–7280.

Fleming, R. W., Torralba, A. and Adelson, E. H. (2004). Specular reflections and the perception of shape, *J. Vis.* **4**, 798–820.

Fleming, R. W., Holtmann-Rice, D. and Bülthoff, H. H. (2011a). Estimation of 3D shape from image orientations, *Proc. Natl Acad. Sci. USA* **108**, 20438–20443.

Fleming, R. W., Jäkel, F. and Maloney, L. T. (2011b). Visual perception of thick transparent materials, *Psychol. Sci.* **22**, 812–820.

Fowlkes, C. C., Martin, D. R. and Malik, J. (2007). Local figure–ground cues are valid for natural images, *J. Vis.* **7**, 2. doi:10.1167/7.8.2.

Harrower, M. R. (1936). Some factors determining figure-ground segmentation, *Br. J. Psychol.* **26**, 407–424.

Jakesch, M., Leder, H. and Forster, M. (2013). Image ambiguity and fluency. *PLoS One* **8**, e74084. doi:10.1371/journal.pone.0074084.

Kanizsa, G. (1985). Seeing and thinking, *Acta Psychol.* **59**, 23–33.

Kim, J., Tan, K. and Chowdhury, N. S. (2016). Image statistics and the fine lines of material perception, *i-Perception* **7**, 2041669516686457. doi:10.1177/2041669516658047.

Magritte, R. (2016). *Selected Writings* (K. Rooney and E. Plattner, Eds; J. Levy, Transl.). Alma Books, Richmond, UK. (Original work published by Alma Books Paris, France,1979.)

Marlow, P. J. and Anderson, B. L. (2015). Material properties derived from three-dimensional shape representations, *Vis. Res.* **115**, 199–208.

Marlow, P. J., Todorović, D. and Anderson, B. L. (2015). Coupled computations of three-dimensional shape and material, *Curr. Biol*, **25**, R221–R222.

Motoyoshi, I., Nishida, S., Sharan, L. and Adelson, E. H. (2007). Image statistics and the perception of surface qualities, *Nature* **447**(7141), 206–209.

Muth, C. and Carbon, C. C. (2013). The aesthetic aha: On the pleasure of having insights into Gestalt, *Acta Psychol. (Amst.)* **144**, 25–30.

Muth, C., Hesslinger, V. M. and Carbon, C. C. (2018). Variants of Semantic Instability (SeIns) in the arts. A classification study based on experiential reports. *Psychology of Aesthetics, Creativity, and the Arts* **12**, 11–23.

Oliva, A. and Schyns, P. G. (2000). Diagnostic colors mediate scene recognition, *Cogn. Psychol.* **41**, 176–210.

Oliva, A. and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope, *Int. J. Comput. Vis.* **42**, 145–175.

Oliva, A. and Torralba, A. (2006). Building the gist of a scene: the role of global image features in recognition, *Prog. Brain Res.* **155**, 23–36.

Olkkonen, M. and Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance, *i-Perception* **2**, 1014–1034.

Orlandi, N. (2014) *The Innocent Eye: Why Vision is not a Cognitive Process*. Oxford University Press, New York, NY, USA.

Paquet, M. (2000). *René Magritte 1898–1967: La Pensée Visible*. Benedikt Taschen, Cologne, Germany.

Peterson, M. A. and Gibson, B. S. (1994). Object recognition contributions to figure–ground organization: Operations on outlines and subjective contours, *Percept. Psychophys.* **56**, 551–564.

Pinna, B. (2007). Art as a scientific object: Toward a visual science of art, *Spat. Vis.* **20**, 493–508.

Pinna, B., Reeves, A., Koenderink, J., van Doorn, A. and Deiana, K. (2018). A new principle of figure-ground segregation: The accentuation, *Vision Res.* **143**, 9–25.

Rubin, E. (1958). Figure and ground, in: *Readings in Perception* D. C. Beardslee and M. Wertheimer (Eds), pp. 194–203. Van Nostrand, Princeton, NJ, USA. (Original work published 1915).

Shrivastava, P., Bhoyar, K. K. and Zadgaonkar, A. S. (2015). A novel approach to scene classification using K-means clustering, *Int. J. Comput. Appl.* **125**, 33–39.

Torralba, A. and Oliva, A. (2003). Statistics of natural image categories, *Network: Comput. Neural Syst.* **14**, 391–412.

Van de Cruys, S. and Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art, *i-Perception* **2**, 1035–1062.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M. and von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization, *Psychol. Bull.* **138**, 1172–1217.