Stephen Grossberg & Baingio Pinna

Neural Dynamics of Gestalt Principles of Perceptual Organization: From Grouping to Shape and Meaning

1. Introduction

Towards Explaining the Pinna "Wolfgang Metzger Award 2009" Percepts

Baingio Pinna's article "New Gestalt principles of perceptual organization: An extension from grouping to shape and meaning" won the 2009 Metzger Award of the Society for Gestalt Theory and its Applications. In this article, Pinna presented a variety of novel images that probe the way in which local and global factors interact to generate the percepts that we consciously see. As Pinna noted at the beginning of his article:

> "Perceptual organization is a key issue of Vision Science challenging all its main approaches, be they computational, psychophysical, neurophysiological or phenomenological. It concerns the problem of why we perceive a world articulated in objects such as people, cities, houses, cars and trees, and not in differences of luminances, edges and bars, or in Wertheimer's (1923) words:
>
> *I stand at the window and see a house, trees, sky. Theoretically I might say there were 327 brightnesses and nuances of color. Do I have "327"? No. I have sky, house, and trees. It is impossible to achieve "327" as such. And yet even though such droll calculation were possible and implied, say, for the house 120, the trees 90, the sky 117 -- I should at least have this arrangement and division of the total, and not, say, 127 and 100 and 100; or 150 and 177."*

Pinna's article goes on to propose intuitive explanations of these percepts in terms of Gestalt grouping principles. These percepts and explanations raise the questions: What are the organizational principles and neural mechanisms that the brain uses to generate these percepts, and what is the evolutionary advantage of mechanisms with these particular properties? The current article proposes neural explanations for these percepts, and clarifies how such neural explanations may unify and extend our understanding of how these percepts arise.

The current article will not attempt to technically describe these principles and mechanisms. It will content itself with an intuitive description that can provide a theoretical bridge between what is known about how the brain sees and the percepts that Pinna has described. Sufficient information will be given to

GESTALT THEORY
© 2012 (ISSN 0170-057 X)
Vol. 34, No.3/4, 399-482

understand that the percepts probe fundamental design principles of how the brain sees, which partially accounts for why we find these percepts so fascinating, as well as to cross the bridge for those who wish to study the technical details. Section 2 describes the neural modeling concepts and mechanisms that will be used to explain key properties of the Pinna (2009) percepts. Subsequent sections use these ideas to explain the Pinna percepts.

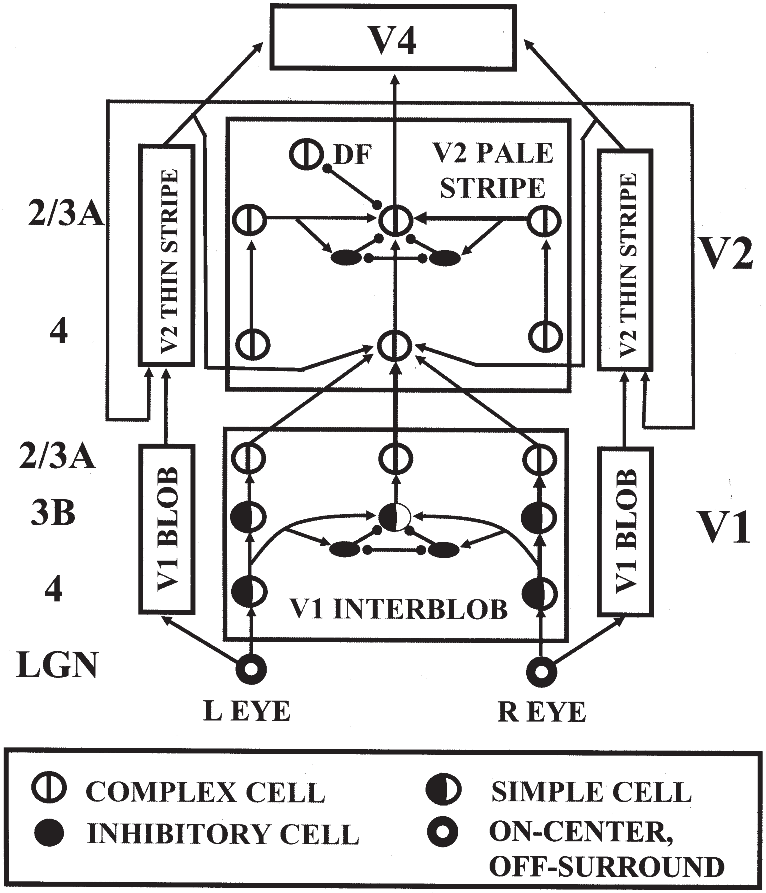


Fig. 1 3D LAMINART model circuit diagram. The model consists of a (V1 Interblob) - (V2 Pale Stripe)-V4 boundary stream which computes 3D perceptual groupings, and a (V1 Blob) - (V2 Thin Stripe)-V4 surface stream which computes 3D surface representations of lightness, color, and depth. The two processing streams interact to overcome their complementary deficiencies (Grossberg, 2000) and create consistent 3D boundary and surface percepts. [Reproduced with permission from Cao and Grossberg (2005).]

2. Neural Modeling Concepts and Mechanisms

2.1. FACADE and 3D LAMINART Models of 3D Vision and Figure-Ground Perception

Such explanations are possible at present because of neural modeling work about how the brain sees that has been led for the past forty years by Stephen Grossberg and his colleagues. This work has proposed unified quantitative explanations and predictions of many perceptual and brain data about how the brain sees. The relevant models are the FACADE (Form-And-Color-And-Depth) theory of 3D vision and figure-ground perception, and its extension in the 3D LAMINART theory (Figure 1) of how FACADE mechanisms may be embedded within the laminar circuits of visual cortex, leading to a great expansion of the theory's explanatory and predictive range on both the perceptual and neural levels. FACADE and 3D LAMINART focus on the mechanisms whereby the brain forms 3D boundaries and surfaces. The percepts described by Pinna (2009) also call for an analysis of shape and meaning.

In order to discuss these additional features, other models are also needed, and how they interact with 3D boundaries and surfaces. Notable among them are the Adaptive Resonance Theory, or ART, model of how the brain learns recognition categories, attends the critical features that are used to define these categories, and experiences conscious percepts; and the ARTSCAN model of how the brain coordinates spatial and object attention while it learns view-invariant and position-invariant object categories during free scanning of a scene with eye movements, maintains a stable representation of percepts under eye movements, directs attention and eye movements to salient object features, and represents the shape and position of objects in a scene.

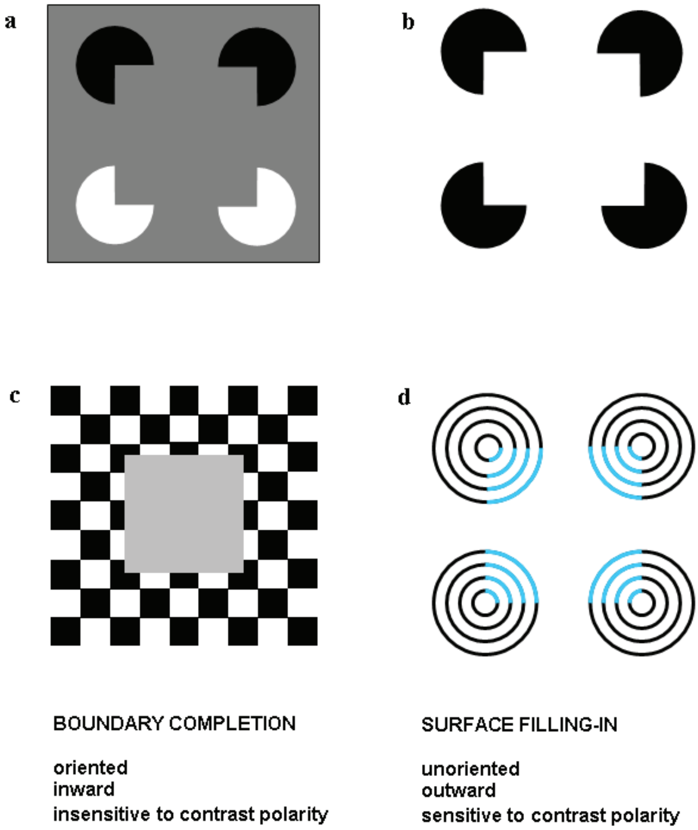


Fig. 2 (a) opposite-contrast Kanizsa square shows that both opposite-contrast polarity and same-contrast polarity collinear edges can group together, and that both sorts of groupings are part of the same boundary completion process. Because two pac men are darker than the background gray, and the other two are lighter than the background gray, they induce lightening and darkening effects that cancel out within the Kanizsa square, thereby creating an invisible, or amodal, square percept that is recognized but not seen. (b) same-contrast Kanizsa square is visible because all four black pac men induce brightness signals within the square that create a brighter square after surface filling-in. (c) pooling of opposite contrast at every position along the square borders illustrates how the brain can build an object boundary around a textured background and thus why "all boundaries are invisible." (d) neon color spreading vividly illustrates the computationally complementary properties of boundary completion and surface filling-in that are summarized at the bottom of the figure.

2.2. 3D Boundaries and Surfaces

Grossberg and colleagues (Cohen & Grossberg, 1984; Grossberg, 1984, 1987, 1994; Grossberg & Mingolla, 1985a, 1985b; Grossberg & Todorovic, 1988) predicted that the functional units of vision are 3D *boundaries* and *surfaces*. In particular, boundaries may be completed by a process called *bipole grouping* whereby cortical grouping cells (e.g., in layer 2/3 pale stripes of cortical area V2) respond when (almost) colinear and (almost) like-oriented cells cooperate to activate intervening cells. Bipole grouping propagates *inwardly* between like *oriented* cells and can occur between cells with opposite contrast polarities, and is in this sense *insensitive* to contrast polarity. In contrast, surfaces may be completed by a process of surface filling-in that spreads *outwardly* in an *unoriented* way and, because visibility is a surface percept, is *sensitive* to contrast polarity. These three pairs of properties illustrate the prediction that boundaries and surfaces are computed using *complementary* laws (Figure 2d). How such complementary laws can support consistent visual percepts is one of the challenges posed by the Pinna (2009) data.

2.3. Analog Coherence of Perceptual Groupings

The 3D LAMINART model (e.g., Grossberg, Mingolla, & Ross, 1997; Grossberg, 1999; Grossberg & Raizada, 2000) explained how the organization of the neocortex into layers enables a fundamental property of boundary groupings and, more generally, groupings using long-range cooperative connections in many neocortical areas; namely, how the strength of a boundary grouping can increase with the amount of perceptual evidence for it; e.g., with the number, contrast, colinearity, etc. of boundary inducers. This property is called *analog coherence*. It is *analog* because of the gradual increase in boundary strength with evidence. It is *coherent* because winning boundaries are chosen and bound together by circuits that include positive feedback loops which maintain the winning boundaries, while negative feedback loops suppress the losing boundaries. Often these analog coherent boundaries are represented by synchronous oscillations (Grossberg & Grunewald, 1997; Grossberg & Somers, 1991; Yazdanbakhsh & Grossberg, 2004).

2.4. A Pre-Attentive Grouping is its own Attentional Prime

The extension of the FACADE model to the 3D LAMINART model did more than explicate the functional significance of identified cells within the laminar circuits of visual cortex. It also greatly expanded the number of perceptual phenomena that the model could explain and predict. This expansion of explanatory range can be traced to the fact that the 3D LAMINART model proposed a solution of a fundamental problem that reconciles demands of perception with those of recognition.

We are all familiar with many illusory contours, such as the Kanizsa square (Figures 2a and 2b), in which an illusory boundary can form over positions that receive no contrastive bottom-up inputs. Indeed, bipole grouping networks simulate how illusory contours may form (e.g., Fang & Grossberg, 2009; Gove, Grossberg, & Mingolla, 1995). However, as noted below, attentional matching in Adaptive Resonance Theory, or ART, circuits can typically only modulate, but not fully activate, the cells in their top-down learned prototypes. Moreover, theorems proved about ART show that this modulatory property is necessary to ensure stable memories of learned recognition categories (e.g., Carpenter & Grossberg, 1987). In other words, if cells can fire without bottom-up inuts, then memory stability can be compromised. How, then, do illusory contours form over positions which receive no bottom-up inputs without destabilizing the prior visually-induced learning of the bipole receptive fields that enabled the illusory contours to form?

This paradox can also be stated as a problem about cortical development. If top-down attention is needed to dynamically stabilize learned memories, then how do perceptual circuits, such as bipole grouping circuits, develop before the circuits at higher cortical levels that are the sources of top-down attentional feedback can themselves develop? In other words, if a given cortical region cannot develop stably enough to feed higher cortical levels with reliable signals, then how can those higher cortical levels provide reliable top-down signals with which to stabilize their development? How is this infinite regress prevented?

The circuits of the LAMINART model (Figure 3) embody a solution, which is often described by the statement that "a pre-attentive grouping is its own attentional prime" (Grossberg, 1999). This sentence summarizes the following considerations: One major route for top-down attention from a higher cortical area is this: *Inter*cortical output signals from layer 6 of a higher cortical area feed back to layer 6 of a lower cortical area, which then activates layer 4 of the latter area via a moculatory on-center, off-surround network (Figure 3a). These interactions create a top-down, modulatory on-center, off-surround network, as ART requires attention to do. However, in LAMINART, *intra*cortical, but interlaminar, connections from layer 2/3 of a cortical area—the layer where groupings form—can also send feedback signals to layer 6, which then sends modulatory on-center, off-surround signals to layer 4 (Figure 3b). In other words, the grouping cells also activate the modulatory on-center, off-surround part of the attentional network, and can do so when this network is developing. The off-surround can inhibit cells that do not contribute to the winning grouping. The remaining active cells can then learn to fine-tune their connections to better support the grouping process via the usual "cells that fire together wire together" kinds of laws. The fact that a "pre-attentive grouping is its own attentional prime" allows each cortical area to stably develop on its own, before its later development and learning are additionally modulated via intercortical top-down attentive signals.

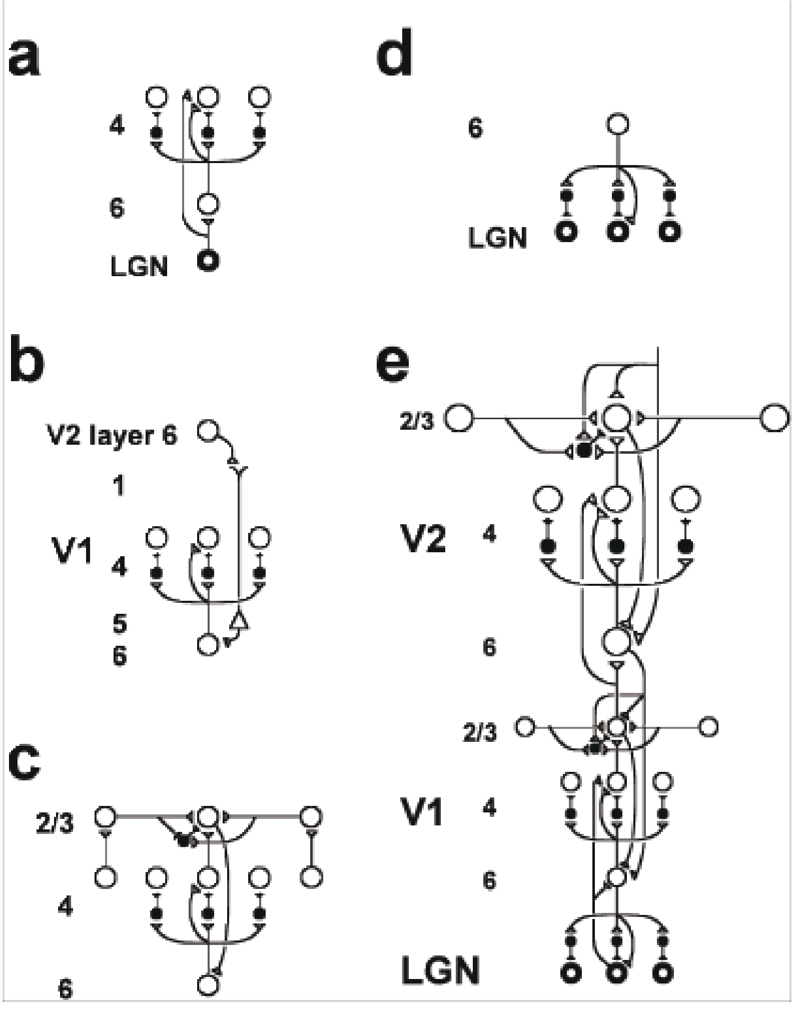


Fig. 3 How known cortical connections join the layer 6→4 and layer 2/3 circuits to form an entire V1/V2 laminar model. Inhibitory interneurons are shown filled-in black. (a) The LGN provides bottom-up activation to layer 4 via two routes. First, it makes a strong connection directly into layer 4. Second, LGN axons send collaterals into layer 6, and thereby also activate layer 4 via the 6 → 4 on-center off-surround path. The combined effect of the bottom-up LGN pathways is to stimulate layer 4 via an on-center off-surround, which provides divisive contrast normalization (Grossberg, 1973, 1980; Heeger, 1992) of layer 4 cell responses. (b) Folded feedback carries attentional signals from higher cortex into layer 4 of V1, via the modulatory 6→4 path. Corticocortical feedback axons tend preferentially to originate in layer 6 of the higher area and to terminate in layer 1 of the lower cortex (Salin and Bullier, 1995, p.110), where they can excite the apical dendrites of layer 5 pyramidal cells whose axons send collaterals into layer 6. The triangle in the figure represents such a layer 5 pyramidal cell. Several other routes through which feedback can pass into V1 layer 6 exist (see Raizada and Grossberg (2001) for a review). Having arrived in layer 6, the feedback is then "folded" back up into the feedforward stream by passing through the 6 → 4 on-center off-surround

path (Bullier et al., 1996). (c) Connecting the 6 → 4 on-center off-surround to the layer 2/3 grouping circuit: like-oriented layer 4 simple cells with opposite contrast polarities compete (not shown) before generating half-wave rectified outputs that converge onto layer 2/3 complex cells in the column above them. Just like attentional signals from higher cortex, as shown in (b), groupings that form within layer 2/3 also send activation into the folded feedback path, to enhance their own positions in layer 4 beneath them via the 6 → 4 on-center, and to suppress input to other groupings via the 6 → 4 off-surround. There exist direct layer 2/3 à 6 connections in macaque V1, as well as indirect routes via layer 5. (d) Top-down corticogeniculate feedback from V1 layer 6 to LGN also has an on-center off-surround anatomy, similar to the 6 → 4 path. The on-center feedback selectively enhances LGN cells that are consistent with the activation that they cause (Sillito et al., 1994), and the off-surround contributes to length-sensitive (endstopped) responses that facilitate grouping perpendicular to line ends. (e) The entire V1/V2 circuit: V2 repeats the laminar pattern of V1 circuitry, but at a larger spatial scale. In particular, the horizontal layer 2/3 connections have a longer range in V2, allowing above-threshold perceptual groupings between more widely spaced inducing stimuli to form (Amir, Harel, & Malach, 1993). V1 layer 2/3 projects up to V2 layers 6 and 4, just as LGN projects to layers 6 and 4 of V1. Higher cortical areas send feedback into V2 which ultimately reaches layer 6, just as V2 feedback acts on layer 6 of V1 (Sandell & Schiller, 1982). Feedback paths from higher cortical areas straight into V1 (not shown) can complement and enhance feedback from V2 into V1. Top-down attention can also modulate layer 2/3 pyramidal cells directly by activating both the pyramidal cells and inhibitory interneurons in that layer. The inhibition tends to balance the excitation, leading to a modulatory effect. These top-down attentional pathways tend to synapse in layer 1, as shown in Figure 2b. Their synapses on apical dendrites in layer 1 are not shown, for simplicity. [Reprinted with permission from Raizada and Grossberg (2001).]

2.5. Attention can Flow Along and Enhance the Strength of a Boundary Grouping

Because a "pre-attentive grouping is its own attentional prime", attention at one location of a grouping can flow along, and enhance the strength of, the entire grouping, as Roelfsema, Lamme, & Spekreijse (1998) have shown with neurophysiological recordings in V1. Grossberg & Raizada (2000) have simulated these data using the LAMINART model (Figures 3b, 3c, and 3e). This property will be important below in showing how an attentional spotlight focused at one position can strengthen, and thereby cause the popout of, an entire perceptual grouping that goes through that position.

2.6. FACADE Principles and Mechanisms

Figure 4 is a macrocircuit of the main processing stages of FACADE theory, which are embodied by laminar circuits in the 3D LAMINART model (Figure 1). These stages are summarized to clarify some of the concepts that help to explain various percepts in Pinna (2009).

Monocular processing of left-eye and right-eye inputs by the retina and lateral geniculate nucleus (LGN) discounts the illuminant and generates parallel signals to the boundary and surface cortical processing streams. These signals are output to model cortical simple cells in the Monocular Boundary stage of cortical area V1, and to the Monocular Surface stage of V1, which contains filling-in domains (FIDOs). Model simple cells have oriented receptive fields that are sensitive to a particular contrast polarity, and come in multiple sizes. Simple cells that respond

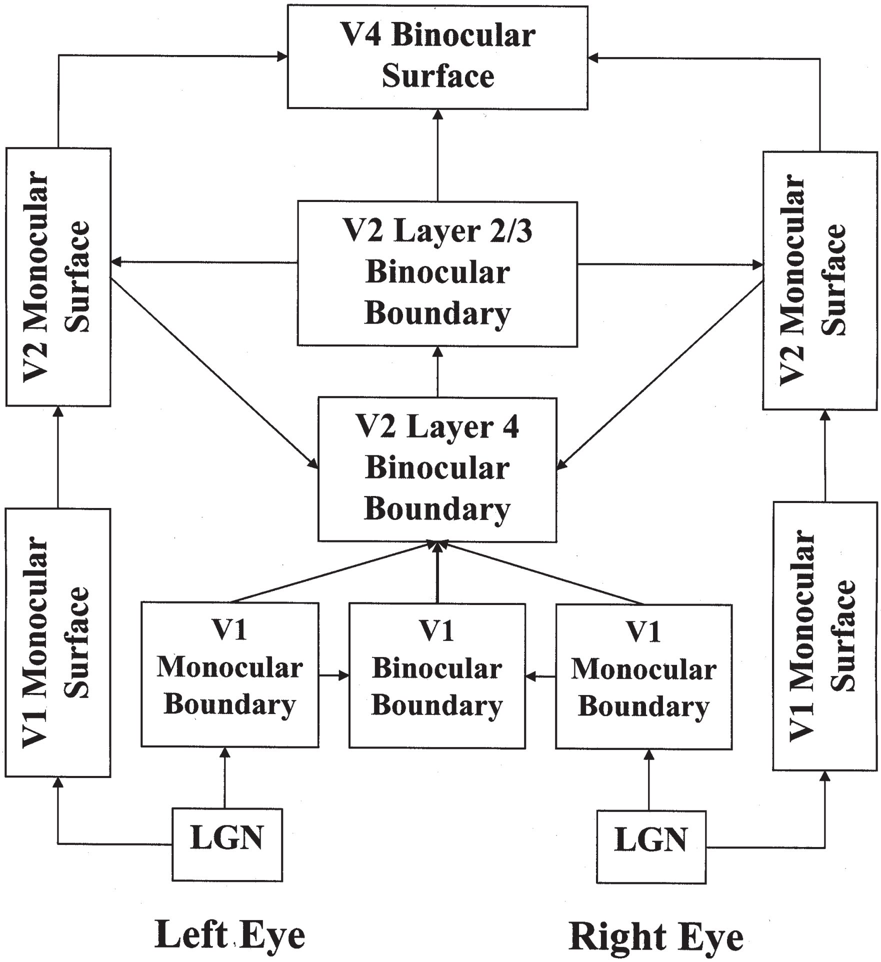


Fig. 4 A block diagram of the enhanced 3D LAMINART model. [Reproduced with permission from Cao and Grossberg (2005).]

to the left and right eyes and the same contrast polarity are combined to generate disparity-sensitive binocular simple cells at the Binocular Boundary stage of V1. Binocular simple cells that are sensitive to opposite contrast polarities and multiple color selectivities are combined to form complex cells at the V1 Binocular Boundary stage. Complex cells with larger receptive fields can binocularly fuse a broader range of disparities than can cells with smaller receptive fields, thereby realizing a "size-disparity correlation" (Smallman & MacLeod, 1994). Competition across disparity at each position and among cells of a given size scale sharpens complex cell disparity tuning (Fahle & Westheimer, 1995). Spatial competition (end-stopping) and orientational competition convert complex cell responses into spatially and orientationally sharper responses at hypercomplex cells. This combination of spatial and orientational competition can also create boundaries at the ends of lines, called *end cuts*, which help to complete boundaries at line ends, and serve as inducers for illusory contours that emerge from arrays of line ends, such as the illusory circle that is induced in the Ehrenstein illusion (Ehrenstein, 1925, 1930, 1941, 1954; Gove, Grossberg, & Mingolla, 1995; Grossberg & Mingolla, 1985b).

2.7. From Filters of Multiple Scales to Boundaries at Multiple Depths

How are responses from multiple receptive field *sizes* combined to generate boundary representations of relative *depths* from the observer? This problem arises because each receptive field size can respond to a range of binocular disparities, with larger scales able to bridge a larger range of disparities. FACADE theory proposes that hypercomplex cells in cortical area V1 activate bipole cells in the pale stripes of cortical area V2. These bipole cells are in the Binocular Boundary stage of cortical area V2. Both monocular boundary and binocular boundary signals from cortical area V1 are combined in layer 4 of cortical area V2 before the bipole cells are activated in layer 2/3 of V2. These bipole cells carry out long-range grouping and boundary completion via horizontal connections in layer 2/3. Bipole grouping collects together outputs from cells of all sizes that are sensitive to a given depth range. The bipole cells then send excitatory feedback signals to cells that represent the same position and orientation, and inhibitory feedback signals to cells at nearby positions and orientations (not shown in diagram). The feedback binds together, or groups, cells of multiple *sizes* into a boundary representation, or copy, that is sensitive to a prescribed range of *depths*. Multiple boundary copies are formed, each corresponding to different (but possibly overlapping) depth ranges.

2.8. Bipole Grouping Initiates Figure-Ground Separation

Bipole grouping also plays a keys role in figure-ground separation. Each bipole cell has an oriented receptive field with two branches (Figure 5). Each branch receives inputs from a range of almost colinear orientations and positions. When the bipole cell does not receive direct bottom-up activation of its cell body, it can fire only if both receptive field branches are simultaneously active. This assures that the cells do not complete boundaries beyond a line end unless there is another line end which provides evidence for such a linkage. The bipole cell thus behaves like a statistical AND gate. Such cells were first used by Cohen & Grossberg (1984), Grossberg (1984), and Grossberg & Mingolla (1985b) to model data about perceptual grouping and filling-in. Cells with similar properties in cortical area V2 were first reported by von der Heydt, Peterhans, & Baumgartner (1984). Bipole cell properties are also consistent with many subsequent psychophysical

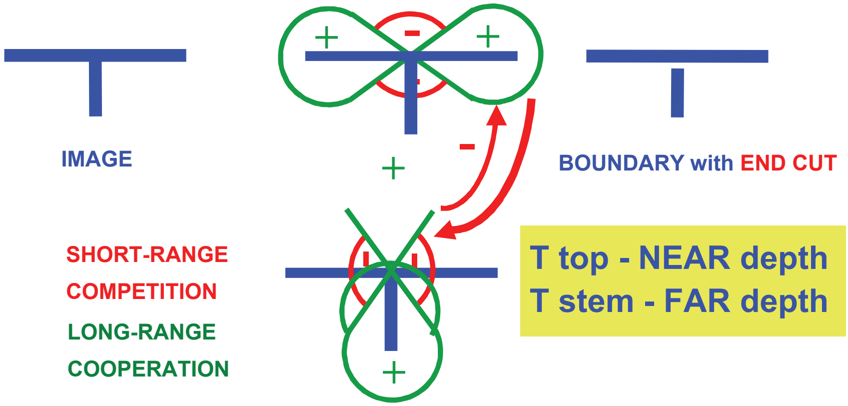


Fig. 5 How bipole grouping initiates figure-ground separation. The left-hand column shows a T-junction that is part of an image. The middle column shows how bipole long-range cooperative grouping and shorter-range competition favor the top of the T and through orientational and spatial competition create a gap at the top of the stem of the T. Such an end gap allows filling-in to spread outside the occluded figure, which initiates figure-ground separation in the manner described in the text. See Grossberg (1994, 1997) for details.

data; e.g., Field et al. (1993) and Shipley & Kellman (1992). Grossberg & Seitz (2003) and Grossberg & Williamson (2001) have simulated how bipole properties may develop in laminar cortical models.

How does the bipole grouping process contribute to figure-ground separation? Long-range excitatory bipole signals combine with shorter-range inhibitory signals to make the system sensitive to T-junctions (Figure 5), without the use of explicit T-junction operators. In particular, horizontally-oriented bipole cells that are located where the top of the T joins its stem receive excitatory inputs to both of their receptive field branches. Vertically-oriented bipole cells that process the stem of the T where it joins the top receive excitatory support only in the one branch that is activated by the stem. Because of this excitatory imbalance, inhibition of the stem by the top can cause a gap in the stem boundary (Figure 5), termed an *end gap*. During filling-in, boundaries contain the filling-in process. Where end gaps occur, brightness or color can flow out of that region. FACADE theory predicts how this escape of color or brightness via filling-in initiates figure-ground separation; see Grossberg (1994, 1997), Grossberg & McLoughlin (1997), and Kelly & Grossberg (2000) for examples.

2.9. Depth-Selective Surface Capture

How do multiple depth-selective boundary copies capture brightness and color signals within depth-selective FCS surface representations? This happens in at least two stages. The first Monocular Surface stage contains *monocular Filling-In DOmains*, or FIDOs, and may exist in V2 thin stripes. Each monocular

FIDO is broken into three pairs of opponent FIDOs (black/white, red/green, and blue/yellow) that receive achromatic and chromatic signals from a single eye. A pair of monocular FIDOs, one for each eye, corresponds to each depth-selective boundary copy, and receives its strongest boundary-gating signals from this boundary. Each monocular FIDO may also receive weaker boundary signals from boundary copies that represent depths near to that of its primary boundary copy. In this way, a finite set of FIDOs can represent a continuous change in perceived depth. See Grossberg & Swaminathan (2004, Figure 23) for a simulation of this property in a 3D slanted percept.

Surface capture occurs when boundary-gating signals interact with illuminant-discounted surface signals. The illuminant-discounted monocular surface signals are input to *all* the monocular FIDOs (Figure 4). Only some FIDOs will selectively fill-in these signals, and thereby lift monocular FIDO signals into depth-selective surface representations for filling-in. The binocular boundary signals from V2 layer 2/3 determine which V2 monocular FIDOs will fill-in. These boundary signals selectively capture surface inputs that are spatially coincident and orientationally aligned with them. Other surface inputs are suppressed. These properties emerge when double-opponent inputs, boundary-gating signals, and surface filling-in processes interact. See Grossberg (1987, Sections 25-26) and Grossberg (1994, Section 45) for details.

Because these filled-in surfaces are activated by depth-selective boundaries, they inherit the depths of their boundaries. 3D surfaces may hereby represent depth as well as brightness and color. This link between depth, brightness, and color helps to explain many percepts, including the percept of "proximity-luminance covariation," or why brighter surfaces tend to look closer; e.g., Egusa (1983).

Not every filling-in event can generate a visible surface. Because activity spreads until it hits a boundary, only surfaces that are surrounded by a *connected* boundary are effectively filled-in. Otherwise, the spreading activity can dissipate across the FIDO. This property helps to explain data ranging from neon color spreading to how T-junctions influence 3-D figure-ground perception (Grossberg, 1994).

An analysis of how these boundary and surface processes react to 3D images shows that too many boundary and surface fragments are formed as a result of the size-disparity correlation (Julesz & Schumer, 1981; Kulikowski, 1978; Richards & Kaye, 1974; Schor et al, 1984; Tyler, 1975). In particular, larger scales can fuse a larger range of disparities than can smaller scales. How are the surface depths that we perceive selected from this range of possibilities across all scales? FACADE's proposed answer to this question follows from its answer to the more fundamental question: How is perceptual *consistency* derived from boundary-surface *complementarity*? FACADE predicts how such *complementary consistency* may be achieved by feedback between the boundary and surface streams that is predicted to occur between the pale stripes and thin strips of cortical area V2. This mutual feedback also helps to explain why cells in the blob and interblob cortical streams may share so many receptive field properties even though they carry out such different tasks. In particular, boundary cells, which summate inputs from both contrast polarities, can also be modulated by feedback from surface cells, which are sensitive to just one contrast polarity.

2.10. Surface Contour Feedback Ensures Complementary Consistency

Boundary-surface consistency derives from a contrast-sensitive process that detects the contours of successfully filled-in regions within the monocular FIDOs. Only successfully filled-in regions can activate such a contour-sensitive process, because other regions either do not fill-in at all, or their filling-in dissipates across space. These filled-in contours activate FCS-to-BCS *surface contour* feedback signals (pathways in Figure 4 from V2 monocular surfaces to binocular boundaries in V2 layer 4) that strengthen boundaries at their own positions and depths, while inhibiting redundant boundaries at farther depths. This inhibition from near-to-far is called *boundary pruning*. It illustrates a perceptual principle called the "asymmetry between near and far." This principle shows itself in many data, including 3D neon color spreading (Nakayama, Shimojo, & Ramachandran, 1990). Grossberg (1994) discusses how to explain such data, and Grossberg & Yazdanbakhsh (2005) simulate them.

How does boundary pruning influence figure-ground separation? Boundary pruning spares the closest surface representation that successfully fills-in a region, and inhibits redundant copies of occluding object boundaries that would otherwise form at farther depths. When these redundant occluding boundaries are removed, the remaining boundaries of partially occluded objects can be completed behind them within boundary copies that represent farther depths. Occluding surfaces hereby form in front of occluded surfaces, while occluded representations are completed behind them.

2.11. Perception-Recognition Dilemma:
Perception of Opaque Occluding Surfaces vs. Recognition of Occluded Surfaces

Visible brightness percepts are not represented within the monocular FIDOs. Model V2 representations of binocular boundaries and monocular filled-in surfaces are predicted to be *amodal*, or perceptually invisible. These representations are also predicted to directly activate object recognition mechanisms in inferotemporal cortex and beyond. Because boundary pruning in V2 enables boundaries of occluded objects to be completed, these completed boundaries are easier to recognize. The monocular FIDO surface representations in V2 can then fill-in an occluded object surface representation within these completed object boundaries, even behind an opaque occluding object. Because these V2 representations are amodal, we can hereby *know* the color of occluded regions without *seeing* them. Said in another way, if we could see them, then all occluding surfaces would look transparent, since a visible representation of the completed occluded surface behind them could also be seen.

How, then, do we *see* opaque occluding surfaces? How does the visual cortex generate representations of occluding and occluded objects that can be easily recognized, yet also allow us to consciously see, and reach for, only the unoccluded parts of objects? In so doing, the visual cortex solves a fundamental problem about how we can *recognize* occluded objects without requiring us to *see* all occluding objects as transparent. This solution helps to explain many paradoxical visual percepts.

FACADE theory proposes that visible unoccluded surfaces are represented at a subsequent processing stage, the binocular FIDO stage (Figure 4), that is predicted to occur in cortical area V4. It is here that *modal*, or visible, surface representations occur, and we see only unoccluded parts of occluded objects, except when transparent percepts are generated by special circumstances, as simulated by Grossberg & Yazdanbakhsh (2005).

To realize this goal, binocular FIDOs use a different combination of boundary and surface representations than is found at the monocular FIDOs (Figure 4). The surface representations at the monocular FIDOs are depth-selective, but they do not combine brightness and color signals from both eyes. Binocular combination of brightness and color signals takes place at the binocular FIDOs, where signals from the V2 monocular surface representations from both eyes are binocularly matched at the binocular FIDOs (Figure 4). These matched signals are redundantly represented on multiple FIDOs. The redundant binocular signals are pruned by inhibitory contrast-sensitive signals from the monocular FIDOs (not shown in Figure 4). As in the case of boundary pruning, these *surface pruning* signals arise from surface regions that successfully fill-in within the monocular FIDOs. These signals inhibit the FCS signals at their own positions and farther depths. As a result, occluding objects do not redundantly fill-in surface representations at multiple depths. Surface pruning is another example of the asymmetry between near and far.

As in the monocular surface FIDOs, surface signals to the binocular surface FIDOs can initiate filling-in only where they are spatially coincident and orientationally aligned with boundaries. The pathways in Figure 4 from V2 layer 2/3 binocular boundaries to the binocular surface FIDOs carry out depth-selective surface capture of the binocularly-matched surface signals that survive surface pruning. In all, binocular FIDOs fill in surface signals that: (a) survive within-depth binocular matching and across-depth inhibition; (b) are spatially coincident and orientationally aligned with boundaries; and (c) are surrounded by a connected boundary or, in the case of shaded surfaces, a plexus of boundaries that is called a *boundary web*.

One further property completes this summary: At the binocular FIDOs, boundaries are added to FIDOs that represent their own and farther depths. This asymmetry between near and far is called *boundary enrichment*. Enriched boundaries prevent occluding objects from looking transparent by blocking filling-in behind them of surface lightness and color from their occluded objects. The total filled-in surface representation across all binocular FIDOs represents the visible percept. It is called a *FACADE representation* because it multiplexes the properties of Form-And-Color-And-Depth that give FACADE theory its name.

2.12. ART Models of Brain Learning, Recognition, and Conscious Perception

As the FACADE model was being developed, a parallel stream of modeling was aimed at understanding the neural mechanisms that support conscious experiences, including conscious visual percepts. Such explanations are developed in Adaptive Resonance Theory, or ART, which was introduced in Grossberg (1976a, 1976b). ART predicts that "all conscious states are resonant states". The name 3D LAMINART theory denotes a synthesis of ideas about how 3D vision works in laminar visual cortical circuits that also embody ART mechanisms.

ART is a cognitive and neural theory of how the brain autonomously learns to categorize, recognize, and predict objects and events in a changing world. The problem of learning makes the unity of conscious experience hard to understand, if only because humans are able to rapidly learn enormous amounts of new information, on their own, throughout life. How do humans integrate all this information into unified conscious experiences that cohere into a sense of self? To a remarkable degree, humans can rapidly learn new facts without being forced to just as rapidly forget what they already know. As a result, we can confidently go out into the world without fearing that, in learning to recognize a new friend's face, we will suddenly forget the faces of our family and friends. This is sometimes called the problem of *catastrophic forgetting*.

ART avoids catastrophic forgetting by proposing a solution of the *stability-plasticity dilemma* (Grossberg, 1980), or how brains can learn quickly without also catastrophically forgetting already learned memories just as quickly. In so doing, ART clarifies key brain processes from which conscious experiences emerge. It describes a functional link between processes of Consciousness, Learning, Expectation, Attention, Resonance, and Synchrony (the CLEARS processes). ART predicted that all brain representations that solve the stability-plasticity dilemma use variations of CLEARS mechanisms (Grossberg, 1978, 1980, 2007a). Through these CLEARS relationships, ART clarifies why many animals are intentional beings who pay attention to salient objects, why “all conscious states are resonant states”, and how brains can learn both *many-to-one maps* (representations whereby many object views, positions, and sizes all activate the same invariant object category) and *one-to-many maps* (representations that enable us to expertly know many things about individual objects and events).

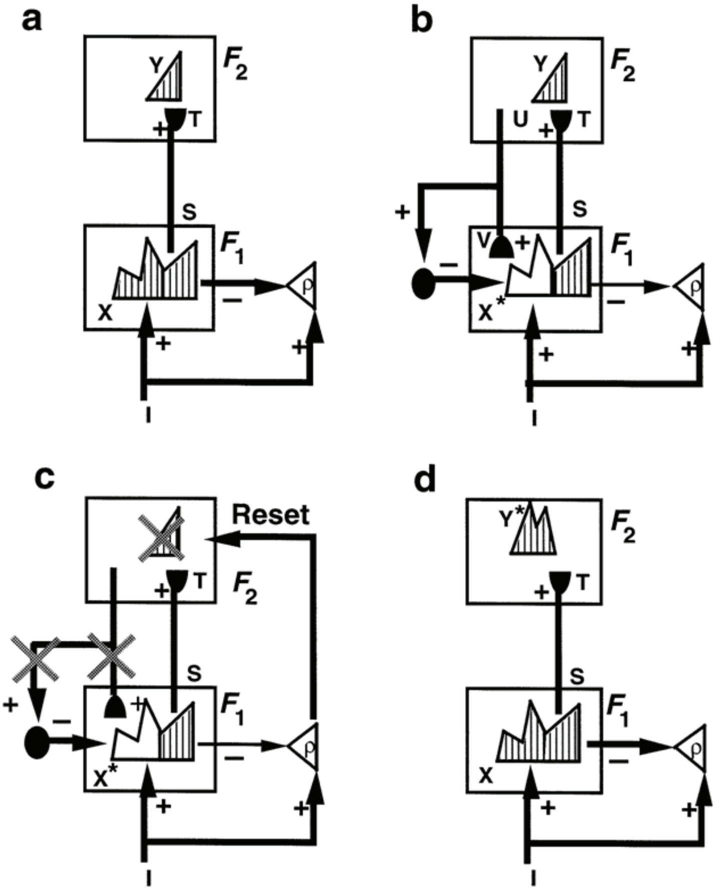


Fig. 6 Search for a recognition code within an ART learning circuit: (a) Input pattern I is instated across feature detectors at level *F1* as an activity pattern X, while it nonspecifically activates the orienting system *A* with gain ρ. X inhibits *A* and generates output pattern S. S is multiplied by learned adaptive weights to form the input pattern T. T activates category cells Y at level *F2*. (b) Y generates the top-down signals U which are multiplied by adaptive weights and added at *F1* cells to form a *prototype* V that encodes the learned expectation of active *F2* categories. If V mismatches I at F1, then a new STM activity pattern X* (the hatched pattern) is selected at *F1*. X* is active at I features that are confirmed by *V*. Mismatched features (white area) are inhibited. When X changes to X*, total inhibition decreases from *F1* to *A*. (c) If inhibition decreases sufficiently, *A* releases a nonspecific arousal burst to *F2*; that is, “novel events are arousing”. Arousal resets *F2* by inhibiting Y. (d) After Y is inhibited, X is reinstated and Y stays inhibited as X activates a different activity pattern Y*. Search for better *F2* category continues until a better matching or novel category is selected. When search ends, an attentive resonance triggers learning of the attended data. [Adapted with permission from Carpenter and Grossberg (1993).]

2.13. Attention, Matching, Biased Competition, Resonance, and Learning

ART accomplishes these properties by proposing how top-down expectations focus attention on salient combinations of cues, and characterizes how attention may operate via a form of self-normalizing "biased competition" (Desimone, 1998). Indeed, attentive matching in ART was predicted to be carried out by a top-down, modulatory on-center, off-surround network (Carpenter & Grossberg, 1987, 1991; Grossberg, 1980). In addition to anticipating the concept of "biased competition" by two decades, the exact form used to such attentive matching anticipated the currently popular "normalization model of attention" (Bhatt, Carpenter, & Grossberg, 2007; Reynolds & Heeger, 2009).

Such top-down attentive matching helps to solve the stability-plasticity dilemma (Figure 4): When a good enough match occurs, a synchronous resonant state emerges that focuses attention upon salient combinations of features, and incorporates these features, through learning, into the bottom-up adaptive filters that activate recognition categories, and into the top-down expectations that are activated, in turn, by these categories; hence the name *adaptive* resonance. All of the main ART predictions have been supported by psychological and neurobiological experiments since it was introduced in 1976; see Grossberg (1999, 2003), Grossberg & Versace (2008), and Raizada & Grossberg (2003) for some reviews. In particular, the prediction that synchronous resonances occur between multiple cortical and subcortical areas has become a very active topic of current investigation; e.g., Buschman & Miller (2007), Engel, Fries, & Singer (2001), Grossberg (2009), and Pollen (1999).

ART has undergone continual development to explain and predict increasingly large behavioral and neurobiological data bases, ranging from normal and abnormal aspects of human and animal perception and cognition, to the spiking and oscillatory dynamics of hierarchically organized laminar thalamocortical networks in multiple modalities. Indeed, some ART models explain and predict behavioral, anatomical, neurophysiological, biophysical, and even biochemical data, thereby providing a growing set of examples capable of partially solving the classical mind/body problem.

2.14. Habituative Gates in Perceptual and Recognition Processes

In all the FACADE, 3D LAMINART, and ART models of perception and recognition, there exist activity-dependent habituative gating processes at the synapses of various connections. These gating processes may be realized either at presynaptic transmitters or postsynaptic receptors. The laws for habituative gates were introduced in Grossberg (1968, 1969). Neurobiological data and supportive modeling in support of them were reported by Abbott et al. (1997) in visual cortex and by Tsodyks & Markram (1997) in somatosensory cortex, using the names synaptic depression and dynamic synapses. These gates processes play several related functional roles. During the development of cortical maps, they prevent perseverative activation of the same cells, and enable new inputs to learn how to activate new cells; e.g., Grossberg & Seitz (2003) and Olson & Grossberg (1998). During percepts of changing visual inputs, they limit persistent activation of the same cells, and enable percepts not to be pathologically smeared across a scene in response to moving objects; e.g., Francis & Grossberg (1996) and Francis et al. (1994). Habituative gates also play a key role in bistable percepts, since habituation of the pathways that support one percept can enable a competing percept to become dominant for a while; e.g., Grossberg and Swaminathan (2004), Grossberg & Yazdanbakhsh (2005), and Grossberg et al. (2008). During the learning of recognition categories, they enable the brain to reset categories whose top-down expectations mismatch bottom-up input patterns, and to search for better-matching categories; e.g., Carpenter & Grossberg (1987, 1991) and Grossberg (1980). Habituative gates also enable individual cells to contrast-normalize their responses to different levels of input intensity.

2.15. ARTSCAN: Invariant Object Learning, Spatial Attention, and Eye Movement Control

One of the developments of ART is to explain how we can learn to recognize *objects*, which is one of the primary goals of Gestalt psychology. This ARTSCAN model clarifies how spatial attention and object attention may be coordinated during the learning of view-invariant and positionally-invariant object categories. It is known both psychophysically and neurophysiologically that spatial attention can attract eye movements and can change the gain with which objects are perceived. In addition, spatial attention can flow along perceptual groupings, as was modeled in the LAMINART simulations of Grossberg & Raizada (2000). In explaining some of the Pinna (2009) percepts, we need to understand how the form of an image can attract spatial attention, how spatial attention can determine where the eyes look, and how attended perceptual groupings can be strengthened through contrast gain control.

To understand how this happens, several basic questions need to be answered: How can multiple views of an object that is seen in different positions and from different distances with respect to an observer all activate the same invariant object category at a sufficiently high processing level? How does the brain learn what an object is under both unsupervised and supervised learning conditions? How does the brain learn to bind multiple views of an object into a view-invariant and positionally-invariant object category while freely scanning a scene with eye movements?

To answer these questions, one also needs to solve the following problem: As eyes scan a scene, two successive eye movements may focus on different parts of the same object or on different objects. How does the brain avoid erroneously classifying views of different objects together, even before the brain knows what the object is? One cannot say that the brain does this by knowing that some views belong together whereas others do not, because this can happen even before the brain has a concept of what the object is. Indeed, such scanning eye movements may be used to learn the object concept in the first place.

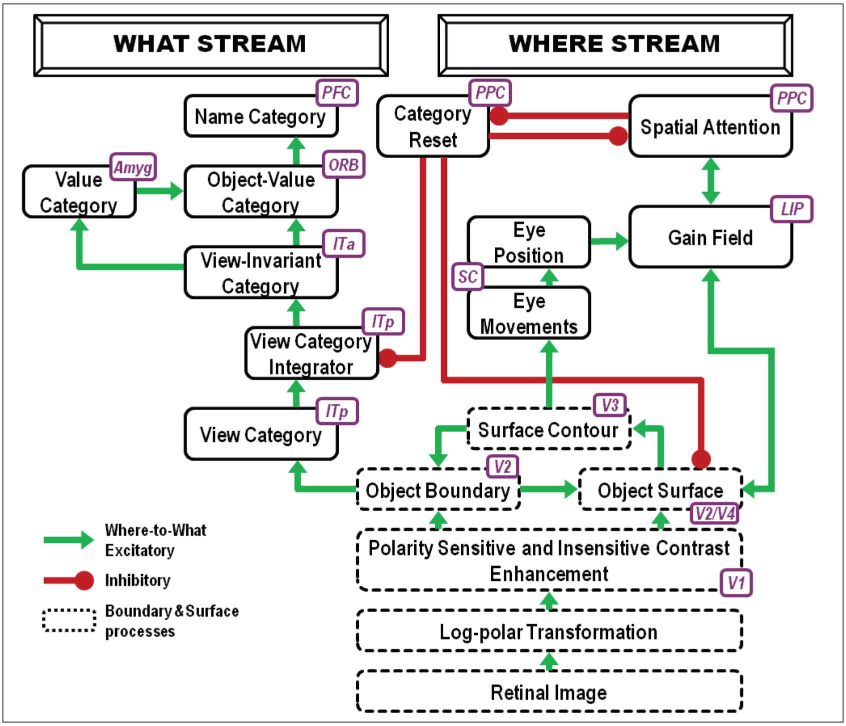


Fig. 7 Macrocircuit of the ARTSCAN model. The Boundary and Surface processes have dashed borders and send input to both the What and Where cortical processing streams. The small white tabs with rounded purple edges next to each box represent the anatomical region in which the process is thought to happen. Green arrows represent excitatory connections, and red connections with a round head represent inhibitory ones. V1: striate visual cortex; V2, V3, V4: prestiate visual cortical areas; LIP: lateral intra-parietal cortex; PPC: posterior parietal cortex; SC: superior colliculus; ITp: posterior part of inferotemporal cortex: ITa: anterior part of inferotemporal cortex; ORB: orbitofrontal cortex; Amyg: amygdala; PFC: prefrontal cortex. See text for details.

The ARTSCAN model (Figure 7) clarifies how the brain uses scanning saccadic eye movements to learn view-invariant and position-invariant object categories (Cao, Grossberg, & Markowitz, 2011; Fazl, Grossberg, & Mingolla, 2008; Foley, Grossberg, & Mingolla, 2012; Grossberg, 2007b, 2009; Grossberg, Markowitz, & Cao, 2011). The discussion about ART above considered only one form of object attention (Posner, 1980) in the What cortical stream, the kind that focuses attention upon the critical feature pattern of a category prototype. ARTSCAN explains how object attention works with spatial attention (Duncan, 1984) in the Where cortical stream to direct eye movements that explore object surfaces. ARTSCAN makes a major new prediction about how spatial and object attention are related; namely, spatial attention coordinates the learning of invariant object categories during free viewing conditions. ARTSCAN hereby clarifies how the view-specific categories that are learned with ART mechanisms may, at a higher cortical level, be bound together into view-invariant and positionally-invariant object categories.

This process begins when a view-specific category of a novel object is learned, and activates cells at a higher cortical level that will become a view-invariant object category as multiple view categories are associated with it. Indeed, as the eyes move around an object surface, multiple view-specific categories are learned of the object (e.g., in ITp; see Figure 7) and are associated with the emerging invariant object category (e.g., in ITa; see Figure 7). How does the brain know how to prevent the invariant object category from being reset while it is getting associated with multiple view-specific categories of a single object, each of which must be reset to enable the next view-specific category to be activated and learned?

2.16. Shrouds, Surface-Shroud Resonance, and Attention Shifts

ARTSCAN predicts how a *pre-attentively* formed surface representation activates an *attentional shroud* (Tyler & Kontsevich, 1995), or form-fitting distribution of spatial attention, even before the brain can recognize the surface as representing a particular object. This shroud persists within the Where Stream during active scanning of an object. The shroud protects the view-invariant category from getting reset, even while view-specific categories are reset, as the eyes explore an object. The shroud does this by inhibiting the ITa reset mechanism.

How does the shroud persist during active scanning of an object? A *surface-shroud resonance* arises due to feedback interactions between a surface representation (e.g., in area V4) and spatial attention (e.g., in posterior parietal cortex, or PPC), and focuses spatial attention upon the object to be learned (Figure 7). When the shroud collapses, the Category Reset stage is disinhibited, giving rise to a transient burst of inhibition that resets the active invariant object category. The collapse of the shroud also enables spatial attention to shift to another object and the eyes to move to explore that object's surface, whereupon new view-specific and view-invariant object categories can be activated. The cycle can then repeat itself.

2.17. Surface-Shroud Resonances and Conscious Seeing

ARTSCAN provides new insights into basic issues such as: What do we consciously see? How is seeing related to recognition? How does seeing or recognition of objects fail to occur during certain conditions? ARTSCAN provides a deceptively simple answer to the first question that extends and unifies two earlier predictions: FACADE theory predicted that all conscious visible qualia are surface percepts, and ART predicted that all conscious events are resonant events. But what sort of resonance supports conscious percepts of visible qualia?

ARTSCAN predicts that visible qualia are consciously seen as part of surface-shroud resonances; that is, we see the visual qualia of a surface when they are synchronized and amplified within a surface-shroud resonance. Once such a resonance begins due to V4-parietal feedback, it can propagate to other brain regions, including V2 and V1. This proposal clarifies, for example, why parts of a scene are not seen when a parietal lesion occurs. Although the visual cortex is intact, the usual trigger for seeing is eliminated by the parietal lesion. ARTSCAN also provides a simple explanation of other sorts of failures of recognition, such as when perceptual crowding occurs (Foley, Grossberg, & Mingolla, 2012): The cortical magnification factor, among other variables, can cause multiple object surfaces to share a single surface-shroud resonance. Since surface-shroud resonances create a link between conscious perception and recognition, multiple objects that share a single resonance cannot be individually recognized.

2.18. Parietal Reset, Attended Surface Contrast Gain, and Figure-Ground Separation

The ARTSCAN model has been used to explain and predict a variety of other data. A key ARTSCAN prediction is that a spatial attention shift (shroud collapse) causes a transient reset burst in parietal cortex that, in turn, causes a shift in categorization rules (new object category activation). In support of this prediction, Chiu & Yantis (2009) have reported evidence using fast event-related fMRI in humans of a domain-independent burst (i.e., reset of a shroud that represents any object) in parietal cortex that mediates between a shift of spatial attention (i.e., collapse of a shroud) and a shift in categorization rules (i.e., activation of new category). Positive feedback from a shroud to its surface is predicted to increase the contrast gain of the attended surface, as has been reported in both psychophysical experiments (Carrasco, Penpeci-Talgar, & Eckstein, 2000) and neurophysiological recordings from cortical area V4 (Reynolds, Chelazzi, & Desimone, 1999; Reynolds & Desimone, 2003; Reynolds, Pasternak, & Desimone, 2000). In addition, the surface-shroud resonance strengthens feedback signals between the attended surface and its generative boundaries, thereby facilitating figure-ground separation of distinct objects in a scene (Grossberg, 1994; Grossberg & Swaminathan, 2004; Grossberg & Yazdanbakhsh, 2005; Kelly & Grossberg, 2000).

2.19. Figure-Ground Separation, Surface Contours, and Complementary Consistency

How does figure-ground separation occur? How can evolution discover a process as seemingly specialized as figure-ground separation? Before the eyes can explore individual surfaces to learn invariant object categories with which to represent them, these surfaces need to be separated from one another by figure-ground separation.

FACADE theory proposes that figure-ground separation arose from the solution of a more fundamental design problem. As noted in Section 2.2, Grossberg and his colleagues predicted that 3D perceptual boundaries and surfaces form using computationally *complementary* rules (Figure 2). If, however, boundaries and surfaces form using complementary rules, then how is a *consistent* conscious percept derived from them? That is, how is *complementary consistency* assured? Grossberg (1994, 1997) showed that the mechanisms which may achieve complementary consistency also automatically trigger figure-ground separation (see Section 2.8). In other words, the mechanisms which enable consistent boundary-surface percepts to occur also make possible the perception of distinct objects. These articles also proposed how neural mechanisms that arose to carry out 3D vision could also interpret 2D images as 3D representations of the world. FACADE theory predicted how complementary consistency is achieved by appropriately designed feedback signals between boundaries and surfaces: Signals from boundaries to surfaces activate surface filling-in (they are filling-in *generators*) and prevent the flow across boundaries of filled-in lightnesses and colors (they are also filling-in *barriers*). As summarized in Section 2.10, feedback signals from surfaces to boundaries assure complementary consistency and trigger figure-ground separation. In particular, the feedback signals from a surface back to its generative boundaries strengthen consistent boundaries, inhibit irrelevant boundaries, and trigger figure-ground separation.

These feedback signals are called *surface contour* signals because they are generated by surfaces that fill-in within closed boundaries (Section 2.10). During 3D vision and figure-ground separation, not all boundaries are connected, due to the size-disparity correlation. Surface contour signals select and enhance the closed boundaries that succeed in generating the surfaces that may enter conscious perception, thereby assuring that a consistent set of boundaries and surfaces are formed, while also, as an automatic consequence, initiating the figure-ground separation of objects from one another.

Surface contours are generated by contrast-sensitive networks at the boundaries of successfully filled-in surfaces; that is, surfaces which are surrounded by a connected boundary and thus do not allow their lightness or color signals to spill out into the scenic background. Surface contour signals are generated at the bounding contours of successfully filled-in surfaces, because these are the positions where sufficiently large changes in filled-in contrast occur across space. When the surface contrast is enhanced by top-down spatial attention as part of a surface-shroud resonance, its surface contour signals (which are contrast-sensitive) become stronger, so the strength of the boundaries around the attended surface increase also. This facilitates selection of these boundaries during figure-ground separation as the "intrinsic" boundaries that bound perceptual figures.

2.20. Depth-Specific and Color-Specific Surface Contours

A little more information about surface contours will be needed to understand the Pinna (2009) data. Boundary and surface representations are depth-selective. As noted in Sections 2.7 and 2.9, multiple depth-selective boundary completion networks interact with multiple depth-selective and color-selective surface Filling-In DOmains, or FIDOs. Each such depth-selective boundary representation provides filling-in generators and barriers for the corresponding depth-selective FIDOs. Since the FIDOs represent achromatic and chromatic features using double-opponent cells, each boundary is topographically broadcast to the multiple achromatic and chromatic FIDOs at its depth.

Surface contour output signals are generated from each of these depth-selective and color-selective FIDOs back to their inducing boundaries. As a result, if an image contains a particular color or contrast across space, the surface contours from the corresponding FIDO can selectively strengthen the boundaries that abut this color or contrast.

2.21. Spatial Attention, Eye Movement Search, and Attentive Gain Control

The ARTSCAN model also predicts that surface contour signals are used to determine where the eyes look on an attended surface. In particular, corollary discharges of surface contour signals are predicted to be processed via cortical area V3A (Caplovitz & Tse, 2007; Nakamura & Colby, 2000) and to generate saccadic commands that are restricted to the attended surface (Theeuwes, Mathot, & Kingstone, 2010) until the shroud collapses and spatial attention shifts to enshroud another object.

Why is it plausible, mechanistically speaking, for surface contour signals to be a source of eye movement target locations, and for these commands to be chosen in cortical area V3A and beyond? It is not possible to generate eye movements that are restricted to a single object until that object is separated from other objects in a scene by figure-ground separation. If figure-ground separation begins in cortical area V2, then these eye movement commands need to be generated no earlier than cortical area V2. Surface contour signals are plausible candidates from which to derive eye movement target commands because they are stronger at contour discontinuities and other distinctive contour features that are typical end points of saccadic movements. ARTSCAN predicts how surface contour signals may be contrast-enhanced at a subsequent processing stage to select the largest signal as the next saccadic eye movement command. Cortical area V3A is known to be a region where vision and motor properties are both represented, indeed that "neurons within V3A…process continuously moving contour curvature as a trackable feature…not to solve the "ventral problem" of determining object shape but in order to solve the "dorsal problem" of what is going where" (Caplovitz & Tse, 2007, p. 1179).

In summary, a surface-shroud resonance means that spatial attention is being paid to the surface of a particular object; the resonance gain amplifies the surface, surface contours, and boundaries of the object; the positions of highest activity on the surface contours are contrast-enhanced and chosen to determine where the eyes look next on the surface; and they thereby help to determine the flow of spatial attention across the object.

With this background, below we briefly summarize the displays in the Pinna (2009) article and the percepts that they generate, and then outline explanations of them using FACADE, 3D LAMINART, ART, and ARTSCAN principles and mechanisms. In order to facilitate comparison of the Pinna (2009) displays with our neural explanations, the figure numbers in Pinna (2009) will be used when describing the displays in the figures below. Thus, the next section begins to number its figures again at 1.

3. Two Forms of Perceptual Organization: from Grouping to Shape

3.1. The Form of Grouping

For the mentioned Figures of Baingio Pinna (2009): *New Gestalt Principles of Perceptual Organization:An Extension from Grouping to Meaning, see Appendix resp. the online version of this paper* on http//:gth.krammerbuch.at

As noted in Pinna (2009), in Figures 1a and 1b, "two large square shapes made up of rows (a) or columns (b) of small squares" are perceived. In Figure 1c, the small squares do not show any preferential direction of the inner organization, i.e. neither rows nor columns, but "a lattice of small squares" is globally perceived. Under these conditions, the row or column organization can be induced through visual attention but not as clearly as in Figures 1a and 1b and it is easily reversible when attention is switched to one or the other result.

Pinna (2009) noted that

> "the inner local organization of the small squares in rows or columns is due to the Gestalt grouping principle of *similarity* (Wertheimer, 1923) stating that, all else being equal, the most similar elements (in color, brightness, size, empty/filled, shapes, etc.) are grouped together. The outer global organization in large square shapes is due to the principle of similarity of shape and to the principle of

> *exhaustiveness*, according to which, all else being equal, all the components of a stimulus pattern tend not to be left out but included as parts of a grouped whole."

The following 3D LAMINART neural principles and mechanisms can explain these percepts: As noted in Figure 1c, neither rows nor columns are preferred by the geometry of the figure. Thus, illusory contours can form in both horizontal and vertical orientations along the colinear sides of the squares (see the simulation in Figure 6A in Grossberg, Mingolla, & Ross, 1997). Other things being equal, these contours are of approximately equal strength. However, in Figure 1a, the black squares fill-in on a different FIDO than the white squares. They can therefore generate surface contours preferentially in the horizontal direction. These surface contours can feed back to the corresponding boundaries and strengthen them as well. Once the horizontal boundaries get stronger than the vertical boundaries, the competitive interactions that are part of the bipole grouping process can break the long-range vertical boundaries, leaving only the small-scale vertical boundaries of the squares themselves (see the simulation in Figure 6B in Grossberg, Mingolla, & Ross, 1997). The ability of the long-range horizontal boundaries to break the vertical boundaries derives from the LAMINART property of *analog coherence*; namely, boundaries with stronger inducers are stronger, and thus can break the weaker boundaries via competitive interactions. The same explanation holds for Figure 1b.

Why can attention cause similar grouping preferences? This follows from the ability of attention to flow along, and thereby strengthen, a grouping; cf. Figures 3b, 3c, and 3e in Section 2. In particular, if attention happens to fall on a long-range horizontal grouping, then it can amplify the strength of that entire grouping, thereby creating a percept of horizontal groupings while enabling it to break the long-range vertical groupings in much the same way that surface contour feedback can.

It should be emphasized that, since "all boundaries are invisible" (Section 2, Figures 2a and 2c), a percept of horizontal or vertical groupings is recognized, but not seen; it is "amodal" and does not usually generate consciously visible qualia.

3.2. The Form of Shape: The Rectangle Illusion

As Pinna (2009) noted, the shapes of the small and large squares are not perceived as isotropic (directional invariant) but rather exhibit a directional symmetry called "the rectangle illusion":

> "The inner organization in rows or columns orients and elongates the shape of both the small and large squares in the same direction as the one of the perceptual grouping." The spontaneous descriptions suggest that this form of grouping can influence the form of shape: squares "become" rectangles.

One factor that may contribute to this effect derives from the same mechanism that enables horizontal (vertical) groupings to be recognized in Figure 1a (Figure 1b). As noted above, in Figure 1a, the long-range horizontal boundaries are stronger than the vertical ones due to surface contour feedback from contrast-specific FIDOs. As a result, the long-range horizontal groupings can break the long-range vertical boundaries that try to form *between* pairs of squares. By the same mechanism, the stronger horizontal groupings can break the vertical boundaries *within* individual squares. Such *end gaps* in boundaries help to explain the percept of neon color spreading (Section 2, Figure 2d; Grossberg and Mingolla, 1985a; Grossberg, 1994), where color is visibly seen to spread out of end gaps in boundaries. In some percepts, these end gaps get repaired by subsequent processes. They nonetheless can have profound perceptual effects; e.g., see Grossberg (1994, Figures 3 - 5 and 26). In the present instance, end gaps enable some of the black color in Figure 1a to spread slightly in a horizontal direction within the boundary web that forms in response to the end gap, thereby causing a slight rectangular assymetry in the percept of the square.

This proposal offers a possible mechanistic explanation for Pinna's assertion that "the problems of grouping and shape appear different but phenomenally related....The form of shape is instead the result of a global perceptual process emerging parallel to or after the form of grouping and giving to the whole a unitary form along the boundary contours."

3.3. The Form of Figure-Ground Segregation

Pinna (2009) asserts that "if we compare figure-ground segregation with grouping, it is reasonable to assume that the former must precede the latter." This assertion needs to be qualified when one studies the dynamics of grouping and figure-ground separation because, as noted in Section 2.8, the laws of bipole grouping and complementary consistency can trigger figure-ground separation. Thus, figure-ground separation does not "precede" grouping, even though groupings and figure-ground separations may both be completed via their mutual interaction.

Pinna (2009) argues, in support of his claim, that "the dot, line or square elements on which grouping acts must be already segregated as a figure from the ground". However, two factors need to be kept in mind when evaluating this fact: (1) Both grouping and figure-ground separation seem to be completed within cortical areas V2 and V4; and (2) visible surface percepts, notably the visible *effects* of grouping and figure-ground separation, seem to be represented using interactions at higher cortical levels. In particular, they may be triggered by a surface-shroud resonance between cortical area V4 and the parietal cortex that may then propagate to lower cortical levels using ART top-down matching processes.

Pinna (2009) also notes that: "An exception to this rule is only to be observed in ambiguous pattern in which figure and ground reverse easily." However, bistable percepts such as the Necker cube (Grossberg & Swaminathan, 2004), bistable transparency (Grossberg & Yazdanbakhsh, 2005), and binocular rivalry (Grossberg, Yazdanbakhsh, Cao, & Swaminathan, 2008) have also been quantitatively simulated using the same FACADE and 3D LAMINART mechanisms that help to explain properties of grouping and figure-ground separation. These articles about reversible percepts also clarify how attention can select which of the possible percepts is seen as figure, how the selected figure may pop forward in depth when attention is paid to it, and how the surface qualia of an attended surface may appear more vivid, if only due to the top-down excitatory feedback from its surface-shroud resonance. In particular, in the case of bistable transparency, these mechanisms explain how the perceived brightness of a figure can change when attention is drawn to it, as Tse (2005) has reported. The percepts in Figure 2 can be explained by the same concepts. These percepts include:

> "**...** two overlapping shapes as illustrated in Figures 2b and 2c...The two triangles of Figure 2c can be easily reversed in their depth and amodal completion organization. In Figure 2d, the two overlapped shapes are not seen anymore but only "a regular six-pointed star is perceived".

Percepts such as these were explained in Grossberg (1994, 1997) and simulated in Kelly and Grossberg (2000) and Grossberg and Yazdanbakhsh (2005) using the following types of concepts. A random fluctuation in boundary strength, or an attentional spotlight on one boundary, can increase the strength of the boundary at that position. Suppose, for definiteness, that attention focuses on the northmost vertex in the middle of the star. Then attention can flow in the Northeast and Northwest directions along the adjacent boundaries, until it reaches the next star vertices, whence it flows in a southern direction along the two top vertical boundary fragments. When this attentional flow reaches the next vertices of the figure, it provides a momentary competitive advantage to these vertical boundaries in the boundary completion process using bipole cells. As a result, both vertical boundaries can be completed across the intervening parts of the star. This may not happen at the initial attentional focus because the Northeast and Northwest boundary strengthenings were balanced, so that neither boundary had a competitive advantage for bipole completion.

Then, by FACADE figure-ground processing that includes surface contour feedback, the surface shape within the completed boundaries is processed at a closer depth than that of the remaining triangular regions. In particular, the completed vertical boundaries become intrinsic boundaries that "belong" to the nearer shape. The remaining triangular boundaries and surfaces are processed at a further depth. The filling-in of the triangular regions at the further depth does not flow behind the nearer occluding shape because boundaries that represent a nearer depth are added to boundaries that represent a further depth, at the binocular FIDO processing stage, by a process of *boundary enrichment* that illustrates "the asymmetry of near and far" (Section 2.11; Grossberg, 1994).

The percept in Figure 2c may be generated if attention falls on one of the edges of the star. This provides a competitive advantage for colinear completion of its boundary. Attention can also flow around the nearest vertex to provide a competitive advantage for boundary completion to that boundary as well. A large triangular boundary is hereby generated. Because it is stronger than the abutting boundaries, it breaks them to create end gaps. This triggers figure-ground separation in which the selected triangle surface appears in front of, and transparent to, the remaining triangular surface.

Why can this percept be bistable? Activity-dependent babituative transmitter gating processes exist within the circuits that define the boundaries of the triangular figure that is in front (Section 2.14). When these gates habituate, these winning triangular boundaries can inhibit the losing triangular boundaries less, whence the losing boundaries can win the competition. When this happens, they can be seen in front and capture their filled-in surface lightness in a nearer depth plane. This perceptual switch can be explained using the FACADE theory mechanisms that were reviewed in Sections 2.8-2.10.

The percept in Figure 2d is more likely if attention is focused on the interior of the star-shaped region, so that no local boundary is favored.
Pinna (2009) notes that, according to Rubin (1915, 1921), the figures, perceived in each of the previous conditions, show the following properties:

> "(1) they appear closer to the observer than the background implying that figure-ground segregation is related to depth perception; (2) their color appears denser than the color of the ground, which appears empty and diaphanous; and (3) they assume the shape traced by the contour and therefore the contour belongs only to the figure, but not to the ground on the other side of the contour ("border-ownership", see Spillmann & Ehrenstein 2004)."

Properties (1) and (3) have already been explained. One factor that influences property (2) is the following one: The nearer figural shape also attracts a surface-shroud resonance, and the resonance's top-down excitatory signals when spatial attention is paid to the surface increase its perceived contrast.

Pinna (2009) interpreted the percepts in Figure 2 in terms of a symmetry principle and the *prägnanz* principle. The explanation above shows how focusing attention on different parts of the image can explain what may heuristically be attributed to influences of symmetry. Prägnanz refers to the uniqueness and stability of an object's representation. These verbal descriptors can now be replaced by precise mechanistic principles and mechanisms that have been used to simulate percepts of this kind.

3.4. The Illusion of the Distorted Seven-Pointed Star

The question to answer in this section is: can grouping and figure-ground segregation influence shape perception? As Pinna (2009) notes, in Figure 3° "a black regular seven-pointed star is perceived". However, after prolonged observation along the boundary contours of the shape, it is perceived as "an irregular seven-pointed shape that does not appear symmetric, as illustrated for example in the summation of Figure 3b, but crooked and unstable". The percepts in Figure 3b – 3g have the same type of explanation as the percepts in Figure 2. The main new feature is seen in Figure 3f, where the white dots stop the good continuation of the sides of the vertexes. The dot shapes block the bipole boundary completion processes, and thus the boundary-completion trigger of figure-ground separation, that generated the various percepts in Figures 2 and 3. Pinna (2009) describes this heuristically as follows: "This result underlines the role of the principle of good continuation that in these conditions wins against prägnanz and symmetry."

3.5. The Grouping Diamond Illusion

In Figure 4, the inner organization of elements influences the form of shape of both the elements and the whole. In Figure 4a, the small and the large composite squares are rotated by 35° and, in the words of Pinna (2009), "the similarity principle, pitted and winning against proximity, groups the components in columns along the diagonal of the whole square". In the percept induced by Figure 4b, other things being equal, there would be a tendency for bipole groupings to form along and between the oblique sides of the small black squares in both the Northeast and Northwest directions. Since the organization of squares across space is symmetric, these mutually perpendicular groupings will weaken each other between the individual squares via balanced mutual competition. In contrast, in the percept induced by Figure 4a, the contrast-specific surface contour feedback from FIDOs to the boundary grouping circuits strengthens the vertical organization of the color-specific boundary groupings. As a consequence "the inner and the whole squares appear like diamond shapes".

In response to Figure 4c, the percept is one of "small squares organized in tilted columns and creating a rotated large square. Both kind of squares (small and large) appear elongated in the same direction as the columns". This can be explained by how contrast-specific surface contour feedback in the Northwest direction strengthens the corresponding boundaries and enables them to successfully compete with the emergent boundaries in the Northeast direction. Pinna (2009) also reported that the inner structure of Figure 4a, due to the alternation of black and white columns, appears "waved like a sinusoid compared to the one perceived in Figure 4c where the waved alternation appears as having a square shape". This can be explained by the orientation-specificity of the surface contour feedback, which influences the good continuation properties of bipole grouping to do its best to provide a global "good continuation" of the local boundary contours, in much the same way that boundaries attempt to follow the curvature of a circle in completing the Ehrenstein illusion. See Gove, Grossberg, and Mingolla (1995) for a simulation of the Ehrenstein illusion. The other percepts in Figure 4 all have a similar explanation.

Pinna (2009) discussed these percepts "as the result of different kinds of perceptual organization: form of grouping and form of shape." The present mechanistic explanation distinguishes between the direct effects of bipole grouping between the squares (that is, the "form of grouping") vs. the effects of feedback from surface shape to bipole groupings via surface contour signals in selectively strengthening some groupings over others, or even, as in the case of the sinusoical groupings formed in the percept of Figure 4a, inducing different groupings that influence the global percept of surface shape ("form of shape").

3.6. The Illusion of the Scalene Triangle

The role of grouping in the perception of shape is also illustrated by using small triangles that create a large triangle. In Figure 5a, small isosceles triangles create "a large isosceles triangle pointing toward the top left-hand corner". This percept can be explained by the analog coherence of bipole grouping. In particular, the Northwest-pointing boundaries are longer than the Northeast boundaries and so, by analog coherence, are stronger and can win the competition for dominance where the boundaries cross. Hence the strongest boundary points to the upper left vertex. This strengthening also organizes the small triangles to point in that direction as well.

In Figure 5b,

> "... the grouping of the triangles is on the basis of similarity of lightness, synergistic with or parallel to the smallest side of each small triangle, thereby making 'the large elongated isosceles shape to appear more pronounced than the one of Figure 5a and more strongly determining the pointing of both large and small rectangles toward the top right-hand corner'".

Here, as in the percepts derived from Figure 4, surface-to-boundary surface contour feedback from color-specific FIDOs strengthens the boundaries along a like-contrast border. An array of parallel boundaries along the entire figure is hereby strengthened, thereby strengthening the Northeast-pointing axis of symmetry of the figure.

What does the "axis of symmetry of the figure" mean, mechanistically speaking? One factor that contributes to it is the structure of the attentional shroud that resonates with the object's surface representation. Due to the axial symmetry of the surface representation induced by Figure 5b, the shroud has a symmetric form facing towards the Northeast direction. Since the surface-shroud resonance defines the consciously seen percept, the percept inherits this orientational axis from it. Several modeling studies have illustrated how the orientation of spatial attention can also orient object recognition and orienting movements; e.g., Carpenter, Grossberg, & Lesher (1998) and Ivey, Bullock, & Grossberg (2011).

In Figure 5c,

> "... the grouping of isosceles triangles is now parallel to one of the two equal sides of each small triangle, therefore creating both locally and globally what we call 'the illusion of the scalene triangle': 'the three sides of both the small and large triangles appear unequal, i.e. scalene, pointing preferentially toward the bottom right-hand corner'".

This effect can be traced again to the role of surface contour feedback from the like-contrast surface regions to their inducing boundaries, combined with the analog coherence of boundary formation.

3.7. The Beveling Effect

Pinna (2009) studies variations of the perceptual property that "the background has not a boundary but it belongs unilaterally to the figure. Therefore the background rather appears as an empty space without a shape." As reviewed in Sections 2.8-2.10, this property can be traced to the way in which bipole grouping causes end gaps in boundaries, thereby initiating figure-ground separation that leads to the attachment of an intrinsic boundary to its surface shape at a closer perceived depth than the background. In Figure 6, Pinna (2009) notes how this property influences perception at the beveled edges of a figure:

> "The beveling appears in fact as something in between a figure and a background. Unlike a background, it is not 'nothing' but 'something'...it appears segregated as an object, it has a shape and like an object it is active in the grouping and shape organization with the other contiguous objects.... However, like a background, it appears as an empty space and manifests the properties of the background. It is a figure and a background at the same time....The results of Figure 6a showed that the beveling of the vertex of the large isosceles triangles of Figure 5...changes both the shape of the triangles into scalene and their pointing in the direction of the beveling."

This percept can be analyzed in much the same way as was the axis of symmetry of the percept induced by Figure 5b. In particular, when spatial attention is focused on the interior side of the beveled edge in Figure 6, the three black triangles there become figures and the white surrounding regions become background. In addition, as noted in Section 2.21, the surface contours generated during a surface-shroud resonance help to determine where the eyes will move, and therefore when attention will be focused. The three Southeast-pointing vertices of the small black triangles create strong surface contour signals that can induce eye movements to those positions. These three nearby target positions for eye movements can draw attention much more effectively than the single vertices of the outermost small triangles in the Northeast and Northwest directions. When the eyes move to the Southeast-pointing vertices, attention focuses on them and can spread along the continuous boundaries, thereby strengthening all the boundaries that point to the beveled vertices. The axis of symmetry of the global figure is hereby directly towards the beveled edges. With these additional factors in mind, the percepts in Figures 6b and 6c have a similar explanation, with Figure 6b being of particular interest because of the way that the like-contrast boundaries favor the Northeast orientation over the Northwest orientation.

3.8. The Concave/Convex Illusion

In Figure 7, Pinna (2009) presents a variant of the rod and frame illusion in which the orientation of obliquely oriented rods within vertically oriented frames affects the orientation of the surrounding frame in the opposite direction. Pinna (2009) notes that “This result reverses the well known effect of frames of reference, whose notion suggests a unidirectional influence of the surrounding including frame on the included elements.” Thus, “a large square with concave vertical and convex horizontal sides in Figure 7a and, conversely, a large square with convex vertical and concave horizontal sides in Figure 7b” are perceived, thereby creating “the concave/convex illusion”.

This percept may be understood as the result of basic mechanisms of bipole grouping. First, the orientations of the bars and the vertical edges of the bounding squares compete with each other at the FACADE stage of orientational competition (Section 2.6). This competition has the effect of causing an “angle expansion” between the two orientations that tilts the vertical sides of the small squares inwardly. Then the bipole grouping between the shifted orientations of the small squares leads to the concave boundary shape of the vertical sides of the large emergent square. Perpendicular end cut boundaries (Section 2.6) from the tilted vertical sides of the small square groupings help to create the convex horizontal sides of the global square grouping. The same explanation holds for the percept induced by Figure 7b.

The role of angle expansion is more directly illustrated by the percepts induced by the images in Figure 8. Pinna (2009) here notes that

> “by rotating the squares of Figure 7a, as illustrated in Figure 8, the concave/convex effect is ‘enhanced when the inset rectangles follow the diagonal of each square (Figure 8a) and weakened or annulled when they are parallel to the vertical orientation of the square sides (Figure 8b)’”.

Angle expansion occurs in response to Figure 8a but not in response to Figure 8b. Similar types of mechanisms help to explain the percepts that are induced by Figures 9 – 11 in Section 3.9 of Pinna (2009).

3.9. Undulation and Twist Illusions

In the Zöllner illusion, parallel lines are perceived as being tilted in a direction perpendicular to the intersecting oblique segments (Figure 12a): "the geometrically parallel rectangles are perceived converging and diverging". This illusion can also be explained by angle expansion due to orientational competition, followed by long-range bipole grouping.

Pinna (2009) developed the remaining images in Figure 12 to challenge "the perceptual enlargement of acute angles" by continuing the oblique segments of Figure 12a, so that each of them becomes the external component of a zigzag path. Figure 12b differs from Figure 12a in at least two ways, however. Unlike Figure 12a, where the oblique lines of two sides of a pair of contiguous rectangles both face up, or both face down, in Figure 12b, the lines on one side face up, while those on the other side face down. Thus, the primary mechanism for angle expansion using orientational competition is not present in the Pinna (2009) variation. In addition, each rectangle is filled with zigzag patterns that cross from one bounding edge to the other. Due to these differences, "the longer sides of the rectangle appear clearly undulated (Figure 12b)" and so this percept is called "the undulation illusion". Pinna (2009) also noted that

> "The undulation induction may derive from a stage of push-pull competition between overall like-orientation at nearby positions followed by a push-pull competition of perpendicular orientations at the same position (Grossberg & Mingolla, 1985). The same idea can also be expressed in terms of grouping of the inner pattern and the surrounding boundary contours, whose shape is influenced by the grouping of zigzags that induce an undulated directional symmetry along the boundaries."

In other words, adding the zigzags brings into play not only orientational competition, but also spatial competition among the extended zigzag boundary contours (Section 2.6). Unlike the Zöllner display, the zigzags in Figure 12b include both horizontal and vertical interior edges, which are mutually perpendicular, and thus exert maximal orientational competition upon each other. This configuration can create an unstable set of boundaries in Figure 12b when the spatial competition within orientations, followed by the orientational competition across orientations, interact in response to internal fluctuations in system noise or shifts in spatial attention. For example, a momentarily enhanced horizontal boundary within the interior of the figure can inhibit like-oriented horizontal cells in the surrounding region. Due to the push-pull organization of orientational competition (Grossberg, 1984; Francis & Grossberg, 1996; Francis, Grossberg, & Mingolla, 1994; Grossberg & Mingolla, 1985b), the further suppression of already inactive horizontally-oriented cells can enhance the activation, via disinhibition, of contiguous vertical boundary cells at the edge of the figure, thereby strengthening the orientational competition between this vertical edge and the figural boundary, leading to greater angle expansion. Moreover, the signals from these boundaries are habituatively gated (Section 2.14), and can thus be dominant at different times, as occurs during bistable percepts and rivalry. As a result, the vertical boundaries near the edge of the figure can be a source of orientational competition at some times, but not others, in a possibly cyclic fashion, depending upon how attention modulates the gain of particular boundaries, and thus their strength as well as their rate of habituation, through time. This alternation in boundary strengths can account for the undulatory percept in this illusion.

Pinna (2009) summaries the main phenomenal properties of the illusion as follows, which can all be analyzed in terms of the above mechanisms:

> "(i) Like Zöllner's tilt, the undulation illusion is enhanced by rotating the stimulus by 45° but it is clearly perceived when the rectangles are oriented vertically or horizontally (compare Figures 12b and 12c). (ii) The zigzags induce undulation at a long distance without intersecting the sides of the rectangle (Figure 12d). (iii) The perceived undulations of the two longer sides of the rectangles do not appear parallel, but tend to be oppositely curved with concave and convex alternations. (iv) By increasing the width of the rectangles the strength of the undulation increases accordingly (not illustrated). (v) The undulation is also induced when the intersecting segments are perpendicular to the target parallel lines (Figure 12e). Under this condition no Zöllner's tilt is expected, as a consequence the undulation illusion is a different new illusion. (vi) By increasing the width and the spatial frequency of the zigzags, the strength of the undulation decreases (not illustrated). (vii) By removing the target parallel lines, the illusory undulation is perceived either in the alignment of the zigzag terminators or, if the conditions are favorable, in the resulting line-induced illusory contours (Figure 12f). (ix) The undulation persists even without the interruptions of the zigzags in the spaces among the rectangles (not illustrated) and by replacing the zigzags with wavy lines (not illustrated)."

When a black bar is inserted in between each couple of adjacent rectangles, another effect emerges: "the straight bars appear twisted (Figure 12g)". If two parallel stripes are now inserted, "each of them appears twisted and both intertwine (Figure 12h)". Pinna (2009) called this "the twist illusion". Pinna (2009) commented that this percept could be understood using the spatial and orientational competitive stages of FACADE theory. Shifts in spatial attention and the temporal dynamics of habituative gating may also influence these percepts through time.

3.10. The Inverted Rod and Frame Illusion

The role of directional symmetry in influencing the shape of a square is illustrated by the percepts induced from Figures 13a - 13d, where "the square shapes within the four rows appear as diamonds (Figures 13a and 13d) or tilted squares (Figures 13b and 13c)" by virtue of the orientation of the inner rectangle or of the position of the outer circle, respectively. Pinna (2009) called this percept "the inverted rod and frame illusion". Pinna (2009) discussed this directional effect as being due to "the same way the zigzags determine the undulated directional symmetry of the rectangles in the undulation illusion".

However, the percepts that are induced by Figure 13 may arise from factors that may not be rate-limiting in the undulation illusion. In particular, these percepts require the distinction between seeing and knowing, or recognizing, whether there is a square or a diamond. It is perfectly possible to see and recognize all of the figures as squares, or as diamonds, by simply shifting attention, with no changes whatsoever in the images themselves. The positions of the inner rectangle and outer circle tend to reactively attract attention, but attention can be volitionally focused elsewhere as well. When attention is attracted to the rectangle, it enables an attentional shroud to spread over the entire figure that is symmetric with respect to the rectangle's orientation. Due to the intimate relationship between attentional shrouds, the orienting of attention for purposes of invariant recognition (Section 3.6), and recognition, the figure may therefore be recognized as a diamond if the rectangle is vertical or as a tilted square if the rectangle is obliquely oriented. The same kind of analysis holds for the influence of the circle on shroud formation, attentional orientation, and recognition, and the influence of the white squares on recognition in Figures 13e and 13f.

3.11. The Pointing Illusion

The pointing illusion that is illustrated by Figures 14a – 14d also has the same explanation. The hardest case to explain here is Figure 14a, since there is no internal rectangle to attract attention and to thus determine the symmetry of the resultant shroud. The symmetry in Figure 14a can be traced to the primacy of vertical and horizontal orientations in the brains of (at least) urban dwellers. This can be seen by tilting the head relative to the figure. The preferred axis of symmetry changes accordingly. Figures 14e – 14i have a similar explanation, keeping in mind, as in the analysis of the beveled edge illusion (Section 3.7 and Figure 6), how the pair of line ends in Figure 14e, the pair of corners in Figures 14f, 14g, and 14i, and the small black triangles in Figure 14h can preferentially attract focal attention and eye movements.

3.12. The Loss of Collinearity Effect

In Figure 15a, "circles and concentric dots appear collinear" as they really are. Giovannelli (1966) changed the arrangement of the surrounding circles, as in Figure 15b, so that "the dots do not appear collinear but slightly misaligned or lightly zigzagged" Pinna (2009) noted that "This effect was attributed (see also Kanizsa, 1972) to the conflict, i.e. to the predominance of the relative position of the wider frame of reference (the circles) over the included components (the dots)". Pinna (2009) suggested, instead, that the loss of collinearity is due to "the interaction between the two directional symmetries due to the circles and the dots and not to the effect induced by the wider frame of reference". Figure 15c illustrates that the frame of reference now appears misaligned due to the components, thus creating the opposite of the Giovannelli illusion, which Pinna (2009) called "the loss of collinearity effect": "the circles surrounding the dots appear slightly misaligned and alternately shifted up and down following the direction and the position of the dots". The other images in Figure 15 probe this effect in various ways. Pinna (2009) concludes that "Even if the form of grouping plays some role in eliciting these effects, it cannot fully explain them without invoking the directional symmetry and another kind of perceptual organization like the form of shape."

How shrouds form in response to the dots and the circles in Figures 15a – 15c helps to clarify what this other kind of perceptual organization may be, since the shapes of the shrouds that form in the parietal cortex influence the perception of the positions of the corresponding surfaces. For example, in response to the image in Figure 15c, placing a dot near the boundary of a circle can cause the shroud to bulge in that direction and thus to shift the perceived location of the circle. The relevance of spatial attentional factors in determining perceived position can be seen by overcoming its effect by volitionally focusing attention on specific parts of the image. For example, focusing on a dot in Figure 15c makes it easy to determine that the pair of continuous dots is colinear with it.

3.13. The Shape of Motion Due to Grouping

Figure 16 depends upon neural processes of motion processing and form-motion grouping that are not reviewed in this article. Some neural modeling articles with relevant dynamics include Baloch & Grossberg (1997), Berzhanskaya et al. (2007), Chey et al. (1997), and Grossberg et al. (2001).

3.14. The Watercolor Illusion and the Illusion of Shape

Watercolor illusion percepts such as those in Figure 17 were explained using FACADE theory principles and mechanisms in Pinna & Grossberg (2005).

3.15. Limits of the Forms of Grouping and Shape: The Perception of Absence

Figure 18 illustrates

> "… another kind of form that cannot be explained by the forms of grouping and shape. This new form is introduced by a variation of Figures 13e - 13f... Given the empty space in the top hand corner (Figure 18a) or on the top-left hand side (Figure 18b) of the whole figures, the directional asymmetry creates 'diamonds' in Figure 18a and 'squares' in Figure 18b...the gestalt grouping by symmetry does not instill polarity to the shape. Therefore, the results of the form of shape are different from the results of the form of grouping."

As noted in Section 3.11, this kind of percept can be understood by how the missing elements influence the symmetry of the shroud of each form, and how that determines how the surface is recognized.

Pinna (2009) discusses how the modified shapes are recognized in terms of "a new form of organization. This is the form of meaning". One factor that contributes to this recognition process is the following: Using ART-like processes, humans can have learned to recognize both the individual squares in the Figure 18 images and the emergent square and diamond shapes before being confronted with the modified shapes. When these shapes activate the learned categories that are most similar to the modified shapes, a matching process (Section 2.13) can begin on multiple spatial scales to compare the learned prototypes of the shapes and how they differ from the particular exemplars in Figure 18. In particular, a distributed categorical representation of object parts and wholes can be activated by the image of the object. An ART model capable of learning such a representation is called *distributed ARTMAP* (Carpenter, 1997). A full analysis of how these parts and wholes are compared with a particular exemplar object using a combination of perceptual, cognitive, and language understanding skills goes beyond the competence of currently available neural models.

4. From the Illusion of Shape to the Illusion of Meaning

More complex images that would require such a meaning analysis are illustrated in Figure 19. The image in Figure 19a is often perceived as "a beveled square". One can see how such a figure might activate the recognition category for a square, and then compare the prototype of a square with that of the image exemplar to focus attention upon the part that appears to be beveled. Prior experiences with beveling behaviors, or at least the meaning of beveling, would be needed to discover this interpretation.

Pinna (2009) notes that "In Figure 19b, neither the gestalt principles nor the form of shape can explain: 'a square perimeter whose top right corner is absent, missing, deleted or cut (89)'. In Figure 19c, 'the glass square is broken in a corner'. The square appears 'made of glass' [2]. As in the case of the beveled square, such interpretations would require read-out from a categorical representation of the object, followed by a part-whole analysis of the non-square components using previously learned categorical knowledge about the deviant shapes.

4.1. The Form of Meaning and its Levels of Perception

Where these shapes could arise from particular actions, they would depend for their interpretation on the sorts of learned interactions across the What and Where cortical streams whereby, say, a photograph of a baseball player holding a ball with a raised arm can be understood as a moment during throwing of the ball. The photograph captures one view of the throw. The entire throw may be understood in terms of how sequences of views of the throw through time are bound together by learning into action understand categories, using ARTSCAN mechanisms.

Grossberg and Vladusich (2010) have summarized and integrated neural models of some of the social cognition processes whereby a student can learn to recognize and imitate behaviors from a teacher who experiences actions in the world from a different perspective. This integrative model is called the CRIB (Circular Reactions for Imitative Behavior) model because it describes how perception-cognition-emotion-action cycles between a teacher and a learner, who experience the world from different perspectives, enable the learner to share joint attention with the teacher and thereby to learn to imitate various of the teacher's skills, including the use of tools. This kind of social cognitive meaning analysis is needed to fully understand the "meaning" interpretations of the other images in Figure 19.

4.2. The Amodal Wholeness

Pinna (2009) claims that

> "The ideal and contingent levels are not perceived in the same way: the former is perceived amodally, the latter modally. The complete *whole* (the square) is seen amodally (amodal completeness), while the incomplete *whole* (irregular shape) modally (modal incompleteness). Similarly, the complete *part* (the absence that at the ideal level is filled with what is missing) is perceived amodally, while the incomplete *part* (the absence as it is perceived) modally....There is a clear link between amodal wholeness and amodal completion (Michotte, 1951; Michotte, Thinès & Crabbé, 1964/1991). Amodal completion occurs when a portion of an object is hidden due to its occlusion behind another object. Under these conditions, the object perceived as occluded is seen as a unitary object, whose boundary contours amodally complete behind the overlapped modal object. The term 'amodal' refers to the fact that even though some object contours are not actually seen, the resulting object appears with a vivid sense of completeness and object unity."

By perceiving "a large square partially occluded by a small one", we perceive meanings similar to and with the same phenomenological structure as those illustrated in Figure 19.

There seems to be a problem, however, with this interpretation. For example, the incomplete horizontal and vertical edges of the shattered glass square in Figure 19c are not amodally completed to create a complete square, at least not by known rules of amodal completion. The same is true for almost all the images in Figure 19. However, such incomplete squares can activate, through bottom-up adaptive filters, the recognition category or categories that best represent the image, including categories of a complete square. Once activated, the category of a complete square, or of parts like representations of broken glass, can read-out, through its learned top-down expectations, a prototype of the complete square or other category, and these can be compared with the representation of the real image to generate a meaningful interpretation of it (Section 2.13). Thus, instead of depending upon horizontal amodal completion mechanisms, it may be easier to understand these meaning phenomena using bottom-up and top-down category learning and expectation processes, along with their embedding in social cognitive learning of events and actions.

4.3. Towards a Perceptual Language

Pinna (2009) notes that

> "The square appears to complete itself 'beyond' the absence that 'occludes' its wholeness thus eliciting its amodal completion. The absence is like a doing word or a 'happening', i.e. something that occurs to the square so that without it the square would have been complete. It follows that the happening is like a visual 'predicate' or like a perceptual 'verb' of the sentence expressing properties, existence, action or occurrence of the subject."

This view is consistent with a social cognitive understanding of the square as being the result of an action, just as a photo of a ballplayer holding a ball with a characteristic arm position signifies a throw of the ball caught in time.

4.4. The Illusion of Meaning

Pinna (2009) writes that "The three forms of organization – grouping, shape and meaning – are present all at the same time in every perceptual pattern. There is no perception without the co-presence of the three forms." In terms of neural mechanisms, these different processes are 3D boundary completion and surface filling-in, whose properties embody our usual understanding of Gestalt laws; surface-shroud resonances to generate consciously seen percepts, and the role of spatial attention in orienting recognition and movement; and ART category learning, recognition, and top-down attentive matching processes, combined with social cognitive processes of learning meaningful representations of events and actions from a teacher who experiences the world from a different perspective. Pinna (2009) presents the images in Figures 20 - 23 as illustrations of images where "the form of meaning is dominant over the others". These images provide a rich collection of challenges for understanding how a single image can be understood as a frame in a sequentially organized behavior, such as throwing a ball, using CRIB mechanisms to learn view-invariant recognition categories for action understanding and, more generally, the multi-modal perception-cognition-emotion-action circular reactions that embody its cognitive meaning.

Summary

This article summarized brain design principles and neural models which realize them to explain many visual percepts, notably those summarized in the 2009 Metzger prize-winning article of Baingio Pinna entitled "New Gestalt principles of perceptual organization: An extension from grouping to shape and meaning". The neural principles and models have been discovered by Stephen Grossberg and his colleagues. They include explanations of how the brain accomplishes 3-D vision and figure-ground perception using hierarchically organized interactions between 3-D boundaries and surfaces; how both opaque and transparent surfaces are generated; how brain processes of consciousness, learning, expectation, attention, resonance, and synchrony are organized and related; and how the brain learns invariant object categories during eye movement search. These models illustrate how the brain is organized into parallel processing streams that compute complementary properties, and how these streams interact to overcome their complementary deficiencies to generate conscious visual percepts. The article explains the basis for the prediction that we consciously see surface-shroud resonances, how we can recognize amodal percepts without seeing them, and how we can consciously recognize familiar percepts whether or not we can see them. These explanations are contained in the FACADE, 3D LAMINART, ART, and ARTSCAN neural models.

Keywords: Grouping, surface filling-in, figure-ground separation, attention, Adaptive Resonance Theory, FACADE, LAMINART, ARTSCAN

Zusammenfassung

Dieser Aufsatz fasst Grundlagen des Gehirnaufbaus und derjenigen neuronalen Modelle, die diese umsetzen, zusammen, um viele visuelle Empfindungen, insbesondere diejenigen, die in Baingio Pinnas 2009 mit dem Metzger-Preis ausgezeichneten Beitrag „New Gestalt Principles of Perceptual Organization: An Extension from Grouping to Meaning" zusammengefasst wurden, zu erklären. Diese neuronalen Prinzipien und Modelle wurden von Stephen Grossberg und seinen Kollegen entdeckt. Sie beinhalten Erklärungen dafür, wie das Gehirn dreidimensionales Sehen und Figur-Grund-Wahrnehmung bewerkstelligt, indem hierarchisch organisierte Interaktionen zwischen dreidimensionalen Begrenzungen und Oberflächen benützt werden; wie sowohl undurchsichtige als auch transparente Oberflächen generiert werden; wie Gehirnprozesse des Bewusstseins, des Lernens, des Erwartung, der Aufmerksamkeit, der Resonanz und der Synchronizität organisiert sind und in Beziehung stehen; und wie das Gehirn während der Suche mittels der Augenbewegungen unveränderliche Objektkategorien erlernt.

Diese Modelle veranschaulichen, wie das Gehirn in parallelen Verarbeitungsflüssen organisiert ist, die sich gegenseitig ergänzende Eigenschaften berechnen, und wie diese Flüsse zum Ausgleich komplementärer Mängelzustände miteinander interagieren, um bewusste visuelle Empfindungen zu erzeugen. Der Aufsatz erläutert die Grundlage für die Vorhersage, dass wir an der Oberfläche bedeckte Schwingungen bewusst sehen, wie wir, ohne sie zu sehen, amodale Empfindungen erkennen können, und wie wir gewohnte Empfindungen bewusst erkennen können, ob wir sie sehen oder nicht. Diese Erklärungen sind in den neuronalen Modellen FACADE, 3D LAMINART, ART und ARTSCAN enthalten.

Schlüsselwörter: Zusammenschluss, Oberflächenfüllung, Figur-Grund-Trennung, Aufmerksamkeit, Adaptive Resonance Theory, FACADE, LAMINART, ARTSCAN.

Financial Support: Stephen Grossberg was supported in part by the SyNAPSE program of DARPA (HR0011-09-C-0001). Baingio Pinna was supported by Finanziamento della Regione Autonoma della Sardegna, ai sensi della L.R. 7 agosto 2007, n. 7, Fondo d'Ateneo (ex 60%) and Alexander von Humboldt Foundation.

References

Abbott, L.F., Varela, K. Sen, K. & Nelson, S.B. (1997): Synaptic depression and cortical gain control. *Science*, **275**, 220-223.

Amir, Y., Harel, M. & Malach, R. (1993): Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *Journal of Comp. Neurology*, **334**, 19–46.

Baloch, A. A. & Grossberg, S. (1997): A neural model of high-level motion processing: Line motion and formotion dynamics. *Vision Research*, **37**, 3037-3059.

Berzhanskaya, J., Grossberg , S. & Mingolla, E. (2007): Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, **20**, 337-395.

Bhatt, R., Carpenter, G. & Grossberg, S. (2007): Texture segregation by visual cortex: Perceptual grouping, attention, and learning.*Vision Research*, **47,** 3173-3211 .

Bullier, J., Hupe, J. M., James, A. & Girard, P. (1996): Functional interactions between areas V1 and V2 in the monkey. *Journal of Physiology (Paris)*, **90**, 217–20.

Buschman, T. J. & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science,* **315**, 1860-1862.

Cao, Y. & Grossberg, S. (2005): A laminar cortical model of stereopsis and 3D surface perception: Closure and da Vinci stereopsis. *Spatial Vision*, **18**, 515-578.

Cao, Y. & Grossberg, S. (2012): Stereopsis and 3D surface perception by spiking neurons in laminar cortical circuits: A method of converting neural rate models into spiking models. *Neural Networks*, **26**, 75-98.

Cao, Y., Grossberg, S. & Markowitz, J. (2011): How does the brain rapidly learn and reorganize view- and positionally-invariant object representations in inferior temporal cortex? *Neural Networks*, **24**, 1050-1061.

Caplovitz, G. P. & Tse, P. U. (2007): V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex,* **17,** 1179-1189.

Carpenter, G.A. (1997): Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. *Neural Networks*, **10**, 1473-1494.

Carpenter, G.A. & Grossberg, S. (1987): A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54-115.

Carpenter, G.A. & Grossberg, S. (1991): *Pattern Recognition by Self-Organizing Neural Networks*. Cambridge, MA: MIT Press.

Carpenter, G.A. & Grossberg, S. (1993): Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neurosciences*, 16, 131-137.

Carpenter, G.A., Grossberg, S. & Lesher, G.W. (1998): The what-and-where filter: A spatial mapping neural network for object recognition and image understanding. *Computer Vision and Image Understanding*, **69,** 1-22.

Carrasco, M., Penpeci-Talgar, C. & Eckstein, M. (2000): Spatial covert attention increases contrast sensitivity across the CSF: support for signal enhancement. *Vision Research,* **40**, 1203-1215.

Chiu, Y. C. & Yantis, S. (2009): A domain-independent source of cognitive control for task sets: shifting spatial attention and switching categorization rules. *Journal of Neuroscience,* **29**, 3930-3938.

Cohen, M.A. & Grossberg, S. (1984): Neural dynamics of brightness perception: Features, boundaries, diffusion, and resonance. *Perception and Psychophysics*, **36**, 428-456.

Desimone, R. (1998): Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London,* **353**, 1245–1255.

Duncan, J. (1984): Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113*, 501-517.

Engel, A. K., Fries, P. & Singer, W. (2001): Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience,* **2**, 704-716.

Egusa, H. (1983): Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual fields. *Perception,* **12**, 167-175.

Ehrenstein, W. (1925): Versuche über die Beziehungen zwischen Bewegungs- und Gestaltwahrnehmung. *Zeitschrift für Psychologie,* **96**, 305-352.

Ehrenstein, W. (1930): Untersuchungen über Figur-Grund-Fragen. *Zeitschrift für Psychologie,* **117**, 339-412.

Ehrenstein, W. (1941): Über Abwandlungen der L. Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie,* **150**, 83-91.

Ehrenstein, W. (1954): *Probleme der ganzheitspsychologischen Wahrnehmungslehre.* 3rd. ed. Leipzig: J.A. Barth Verlag.

Fahle, M. & Westheimer, G. (1995): On the time-course of inhibition in the stereoscopic perception of rows of dots. *Vision Research,* **35**, 1393—1399.

Fang, L. & Grossberg, S. (2009): From stereogram to surface: How the brain sees the world in depth. *Spatial Vision*, **22**, 45-82.

Fazl, A., Grossberg, S. & Mingolla, E. (2009): View-invariant object category learning, recognition, and search: How spatial and object attention are coordinated using surface-based attentional shrouds. *Cognitive Psychology*, **58**, 1-48.

Field, D. J., Hayes, A. & Hess, R. F. (1993): Contour integration by the human visual system: Evidence for a local "association field". *Vision Research,* **33**, 173-193.

Foley, N.C., Grossberg, S. & Mingolla, E. (2012): Neural dynamics of object-based multifocal visual spatial attention and priming: Object cueing, useful-field-of-view, and crowding. *Cognitive Psychology*, **65**, 77-117.

Francis, G. & Grossberg, S. (1996): Cortical dynamics of boundary segmentation and reset: Persistence, afterimages, and residual traces. *Perception*, **35**, 543-567.

Francis, G., Grossberg, S. & Mingolla, E. (1994): Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, **34**, 1089-1104.

Gove, A., Grossberg, S. & Mingolla, E. (1995): Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience,* **12**, 1027-1052.

Grossberg, S. (1968): Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, **60**, 758-765.

Grossberg, S. (1969): On the production and release of chemical transmitters and related topics in cellular control *Journal of Theoretical Biology,* 22, 325-364.

Grossberg, S. (1976a): Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121-134.

Grossberg, S. (1976b): Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187-202.

Grossberg, S. (1978): A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology, Volume 5.* New York: Academic Press, pp. 233-374.

Grossberg, S. (1980): How does a brain build a cognitive code? *Psychological Review*, **87**, 1-51.

Grossberg, S. (1987): Cortical dynamics of three-dimensional form, color, and brightness perception, II: Binocular theory. *Perception and Psychophysics,* **41**, 117-158.

Grossberg, S. (1994): 3-D vision and figure-ground separation by visual cortex. *Perception and Psychophysics*, 1994, **55** , 48-120.

Grossberg, S. (1997): Cortical dynamics of three-dimensional figure-ground perception of two-dimensional figures. *Psychological Review*, **104** , 618-658.

Grossberg, S. (1999): How does the cerebral cortex work? Learning, attention and grouping by the laminar circuits of visual cortex. *Spatial Vision, **12,** 163-186.*

Grossberg, S. (2000): The complementary brain: Unifying brain dynamics and modularity. *Trends in Cognitive Sciences,* **4,** 233-246.

Grossberg, S. (2003): How does the cerebral cortex work? Development, learning, attention, and 3D vision by laminar circuits of visual cortex. *Behavioral and Cognitive Neuroscience Reviews*, **2**, 47-76.

Grossberg, S. (2007a): Consciousness CLEARS the mind. *Neural Networks*, **20,** 1040-1053.

Grossberg, S. (2007b): Towards a unified theory of neocortex: Laminar cortical circuits for vision and cognition. In P. Cisek, T. Drew & J. Kalaska (eds): *Computational Neuroscience: From Neurons to Theory and Back Again*. Amsterdam: Elvesier, pp. 79-104.

Grossberg, S. (2009): Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. *Philosophical Transactions of the Royal Society of London B Biological Sciences,* **364**, 1223-1234.

Grossberg, S. & Grunewald, A. (1997): Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience,* **9,** 117-132.

Grossberg, S. & Hong, S. (2006): A neural model of surface perception: Lightness, anchoring, and filling-in. *Spatial Vision*, **19**, 263-321.

Grossberg, S. & Howe, P.D.L. (2003): A laminar cortical model of stereopsis and three-dimensional surface perception. *Vision Research, **43**,* 801-829.

Grossberg, S., Kuhlmann, L. & Mingolla, E. (2007): A neural model of 3D shape-from-texture: Multiple-scale filtering, boundary grouping, and surface filling-in. *Vision Research*, **47**, 634-672.

Grossberg, S., Markowitz, J. & Cao, Y. (2011): On the road to invariant recognition: Explaining tradeoff and morph properties of cells in inferotemporal cortex using multiple-scale task-sensitive attentive learning. *Neural Networks*, **24**, 1036-1049.

Grossberg, S. & McLoughlin, N. (1997): Cortical dynamics of 3-D surface perception: Binocular and half-occluded scenic images. *Neural Networks*, 1997, **10** , 1583-1605.

Grossberg, S. & Mingolla, E. (1985a): Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, **92**, 173-211.

Grossberg, S. & Mingolla, E. (1985b): Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics,* **38**, 141–171.

Grossberg, S., Mingolla, E. & Ross, W.D. (1997): Visual brain and visual perception: How does the cortex do perceptual grouping? *Trends in Neurosciences,* **20**, 106-111.

Grossberg, S., Mingolla, E. & Viswanathan, L. (2001): Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research,* **41**, 2521-2553 .

Grossberg, S. & Raizada, R. (2000): Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research,* **40,** 1413-1432.

Grossberg, S. & Seitz, A. (2003): Laminar development of receptive fields, maps, and columns in visual cortex: The coordinating role of the subplate. *Cerebral Cortex*, **13**, 852-863.

Grossberg, S. & Somers, D. (1991): Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, **4**, 453-466.

Grossberg, S. & Swaminathan, G. (2004): A laminar cortical model for 3D perception of slanted and curved surfaces and of 2D images: development, attention and bistability. *Vision Research*, **44**, 1147-1187.

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

Grossberg, S. & Versace, M. (2008): Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. *Brain Research*, **1218**, 278-312.

Grossberg, S. & Vladusich, T. (2010): How do children learn to follow gaze, share joint attention, imitate their teachers, and use tools during social interactions? *Neural Networks*, **23**, 940-965.

Grossberg, S. & Williamson, J.R. (1999): A self-organizing neural system for learning to recognize textured scenes. *Vision Research,* **39,** 1385-1406.

Grossberg, S. & Yazdanbakhsh, A. (2005): Laminar cortical dynamics of 3D surface perception: Stratification, transparency, and neon color spreading. *Vision Research*, **45**, 1725-1743.

Grossberg, S., Yazdanbakhsh, A., Cao, Y. & Swaminathan, G. (2008): How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Research*, **48**, 2232-2250.

Ivey, R., Bullock, D. & Grossberg, S. (2011): A neuromorphic model of spatial lookahead planning. *Neural Networks*, **24**, 257-266.

Jastrow, J. (1891): A study of Zöllner's figures and other related illusions. *American Journal of Psychology,* **4**, 381-398.

Julesz, B. & Schumer, R.A. (1981): Early visual perception. *Annual Review of Psychology,* **32**, 572-627.

Kanizsa, G. (1972): "Errore del gestaltista" e altri errori da aspettativa. *Rivista di Psicologia,* **66**, 3-18.

Kelly, F.J. & Grossberg, S. (2000): Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception & Psychophysics,* **62,** 1596-1619.

Kulikowski, J. J. (1978): Limit of single vision in stereopsis depends on contour sharpness. *Nature,* **275**, 126-127.

Leveille, J., Versace, M. & Grossberg, S. (2010): Running as fast as it can: How spiking dynamics form object groupings in the laminar circuits of visual cortex. *Journal of Computational Neuroscience*, **28**, 323-346.

McLoughlin, N.P. & Grossberg, S. (1998): Cortical computation of stereo disparity. *Vision Research*, **38** , 91-99.

Michotte, A. (1951): Une nouvelle énigme de la psychologie de la perception: le «donné amodal» dans l'expérience. International Congress of Psychology, Stockholm.

Michotte, A., Thinès, G. & Crabbé, G. (1964): Les compléments amodaux des structures perceptuals. Studia psychologica, Louvain, Publications Universitaires. Repr. in G. Thinès, A. Costall & G. Butterworth (1991): *Michotte's experimental phenomenology of perception,* Hillsdale, NJ: Lawrence Erlbaum.

Morgan, M. (2002): Running rings around the brain. *The Guardian Thursday,* 24 January 2002.

Müller-Lyer, F. C. (1898): *Optische Urtheilstäuschungen.* Archiv für Anatomie und Physiologie, Physiologische Abteilung. **2**, 263-270.

Nakamura, K. & Colby, C. L. (2000): Visual, saccade-related, and cognitive activation of single neurons in monkey extrastriate area V3A. *Journal of Neurophysiology,* **84,** 677-692.

Nakayama, K., Shimojo, S. & Ramachandran, V.S. (1990): Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception,* 19, 497-513.

Olson, S. & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance. *Neural Networks*, **11**, 189-208.

Pinna, B. (2009): New Gestalt principles of perceptual organization: An extension from grouping to shape and meaning. *Gestalt Theory - An International Multidisciplinary Journal*, **32**, 11-78.

Pinna, B. & Grossberg, S. (2005): The watercolor illusion and neon color spreading: A unified analysis of new cases and neural mechanisms. *Journal of the Optical Society of America JOSA A,* **22**, 2207-2221.

Pollen, D. A. (1999): On the neural correlates of visual perception. *Cerebral Cortex,* **9**, 4–19.

Posner, M.I. (1980): Orienting of attention. *Quarterly Journal of Experimental Psychology,* **3***2*, 3-25.

Raizada, R. & Grossberg, S. (2001): Context-sensitive bindings by the laminar circuits of V1 and V2: A unified model of perceptual grouping, attention, and orientation contrast. *Visual Cognition,* 8, 431-466.

Raizada, R. & Grossberg, S. (2003): Towards a theory of the laminar architecture of cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, **13**, 100-113.

Reynolds, J., Chelazzi, L. & Desimone, R. (1999): Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience*, **19**, 1736–1753.

Reynolds, J. H. & Desimone, R. (2003): Interacting roles of attention and visual salience in V4. *Neuron,* **37**, 853-863.

Reynolds, J. H. & Heeger, D. J. (2009): The normalization model of attention. *Neuron,* **61**, 168-185.

Reynolds, J. H., Pasternak, T. & Desimone, R. (2000): Attention increase sensitivity of V4 neurons. *Neuron*, **26**, 703-714.

Richards, W. A. & Kaye, M. G. (1974): Local versus global stereopsis: Two mechanisms. *Vision Research,* **14**, 1345-1347.

Roelfsema, P. R., Lamme, V. A. F. & Spekreijse, H. (1998): Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, **395**, 376–381.

Rubin, E. (1915): *Synsoplevede Figurer*. Kobenhavn: Glydendalske.

Rubin, E. (1921): *Visuell wahrgenommene Figuren.* Kobenhavn: Gyldendalske Boghandel.

Salin, P.A. & Bullier, J. (1995): Corticocortical connections in the visual system: structure and function. *Physiological reviews,* **75**, 107-154.

Sandell, J. H. & Schiller, P. H. (1982): Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *Journal of Neurophysiology*, **48**, 38–48.

Schor, C. M., Wood, I. & Ogawa, J. (1984): Binocular sensory fusion is limited by spatial resolution. *Vision Research.* 24, 661-665.

Shipley, T. F. & Kellman, P. J. (1992): Perception of partly occluded objects and illusory figures: Evidence for an identify hypothesis. *Journal of Experimental Psychology: Human Perception & Performance,* **18**, 106-120.

Sillito, A. M., Jones, H. E., Gerstein, G. L. & West, D. C. (1994): Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, **369**, 479–482.

Smallman, H. S. & MacLeod, D. I. A. (1994): Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A,* **11**, 2169-2183.

Spillmann, L. & Ehrenstein, W.H. (2004): Gestalt factors in the visual neurosciences. In L. Chalupa & J. S. Werner (eds.): *The Visual Neurosciences.* Cambridge, MA, MIT Press, 1573-1589.

Theeuwes, J., Mathot, S. & Kingstone, A. (2010): Object-based eye movements: the eyes prefer to stay within the same object. *Attention, Perception & Psychophysics,* **72**, 597-601.

Tse, P. U. (2005): Voluntary attention modulates the brightness of transparent surfaces. *Vision Research*, **45**, 1095-1098.

Tsodyks, M. V. & Markram, H. (1997): The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences*, **94**, 719-723.

Tyler, C. S. (1975): Spatial organization of binocular disparity sensitivity. *Vision Research,* **15**, 583-590.

Tyler, C. W. & Kontsevich, L. L. (1995): Mechanisms of stereoscopic processing: stereoattention and surface perception in depth reconstruction. *Perception,* **24**, 127-153.

Von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984): Illusory contours and cortical neuron responses. *Science,* **224**, 1260-1262.

Wertheimer, M. (1923): Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung,* **4**, 301-350.

Yazdanbakhsh, A. & Grossberg, S. (2004): Fast synchronization of perceptual grouping in laminar visual cortical circuits. *Neural Networks*, **17**, 707-718.

Stephen Grossberg is Wang Professor of Cognitive and Neural Systems and Professor of Mathematics, Psychology, and Biomedical Engineering at Boston University. Grossberg is a major pioneer and current leader of theoretical research on how brains give rise to minds. In the 1960s, Grossberg introduced the paradigm and the core nonlinear differential equations that form the foundation of this field. His models and leadership have contributed to all branches of the field. (http://en.wikipedia.org/wiki/Stephen_Grossberg and http://cns.bu.edu/Profiles/Grossberg/GrossbergNNeditorial2010.pdf).

Address: Center for Adaptive Systems, Graduate Program in Cognitive and Neural Systems, Department of Mathematics, Boston University, 677 Beacon Street, Boston, MA 0221, USA.
E-mail: steve@bu.edu

Baingio Pinna has been Professor of Experimental Psychology and Visual Perception at the University of Sassari, Italy, since 1995. In 2001/02 he received a Research Fellowship from the Alexander Humboldt Foundation at Freiburg, Germany and was winner of a scientific productivity prize at the University of Sassari in 2007. His main research interests concern Gestalt psychology, visual illusions, psychophysics of perception of shape, motion, color and light, and vision science of art.

Address: Dipartimento di Scienze dei Linguaggi, University of Sassari, Italy.
E-mail: baingio@uniss.it

Appendix

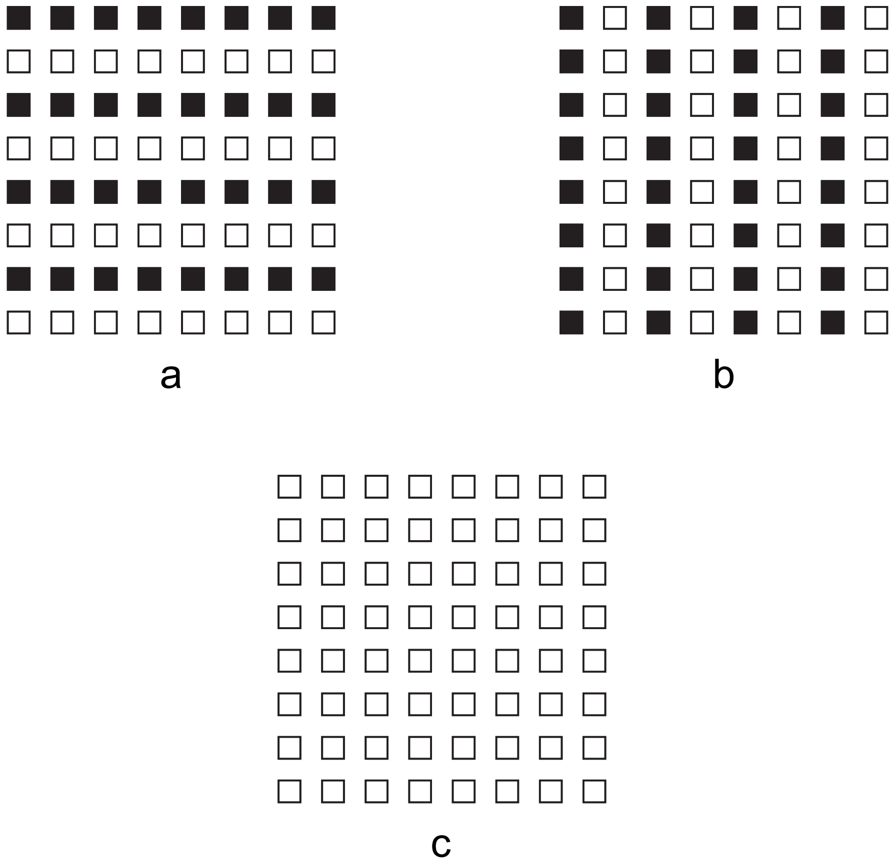


Figure 1. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

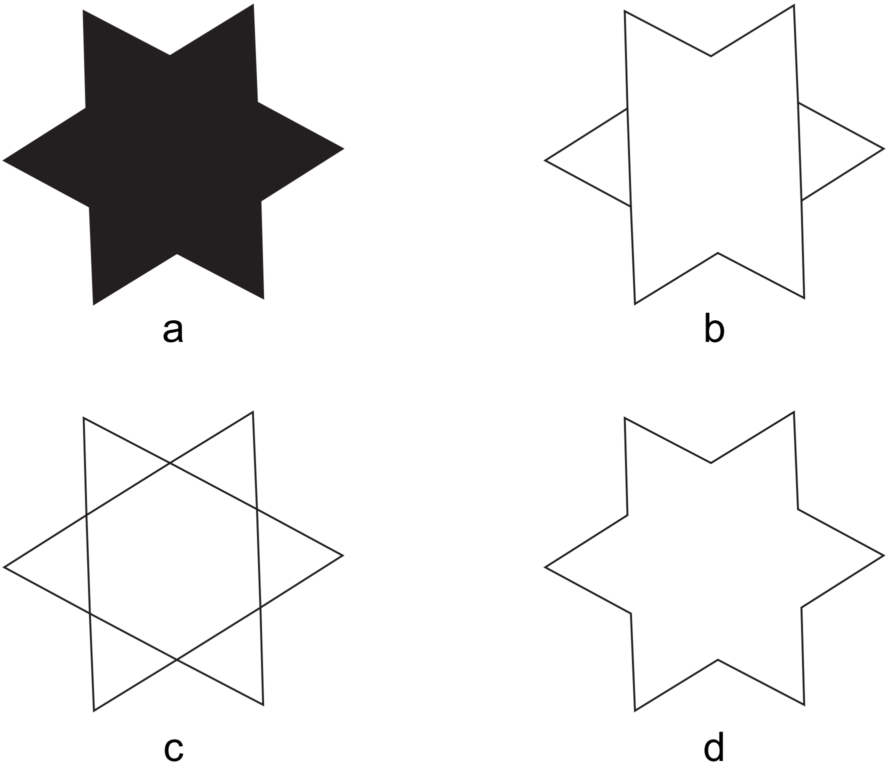


Figure 2. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

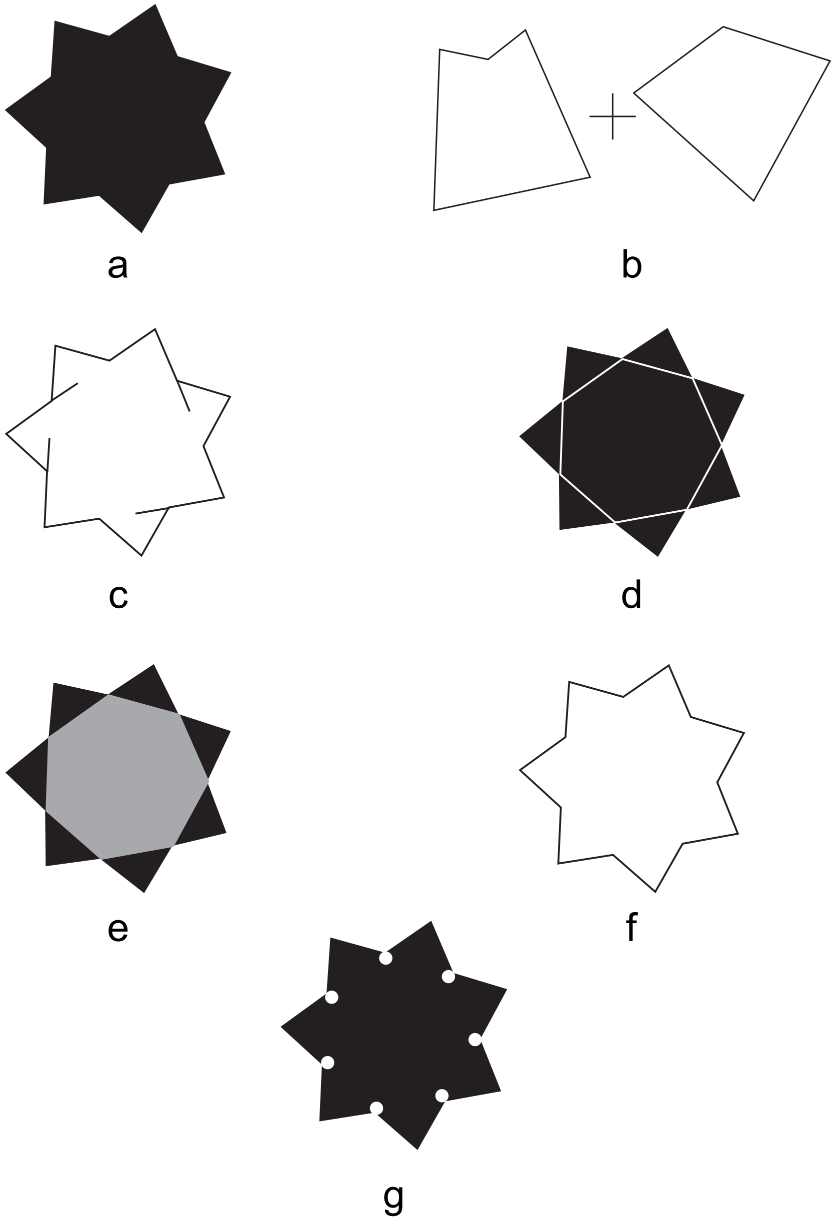


Figure 3. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

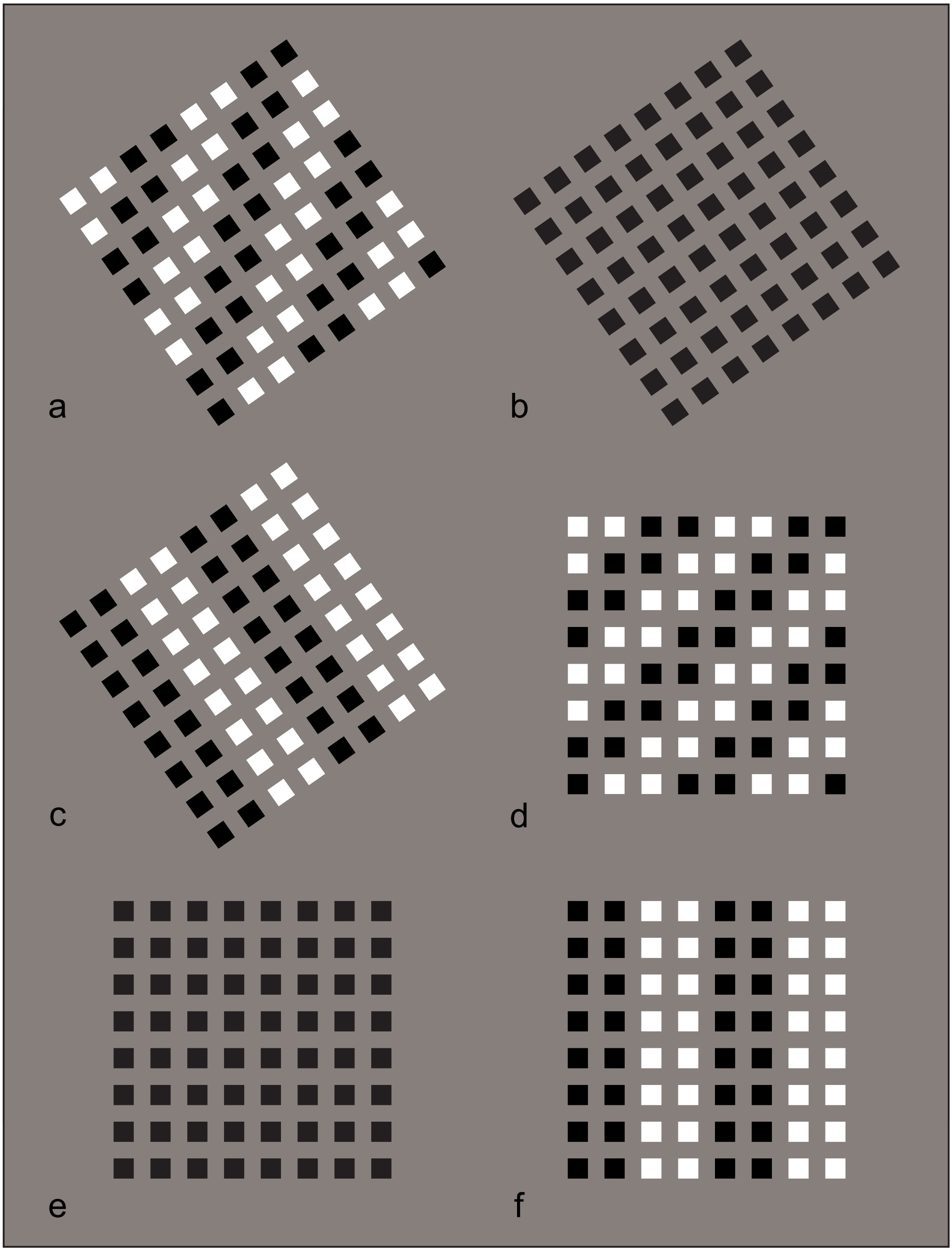


Figure 4. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

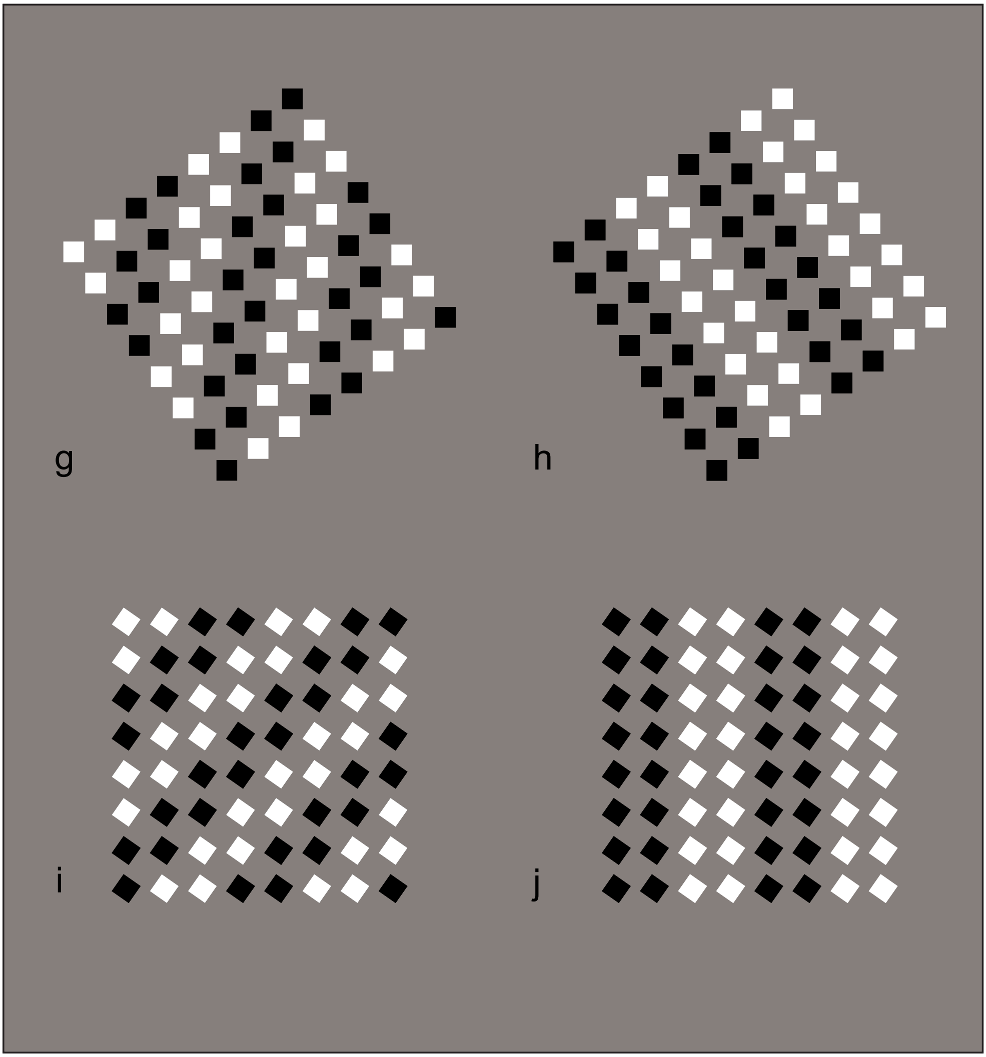


Figure 4. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

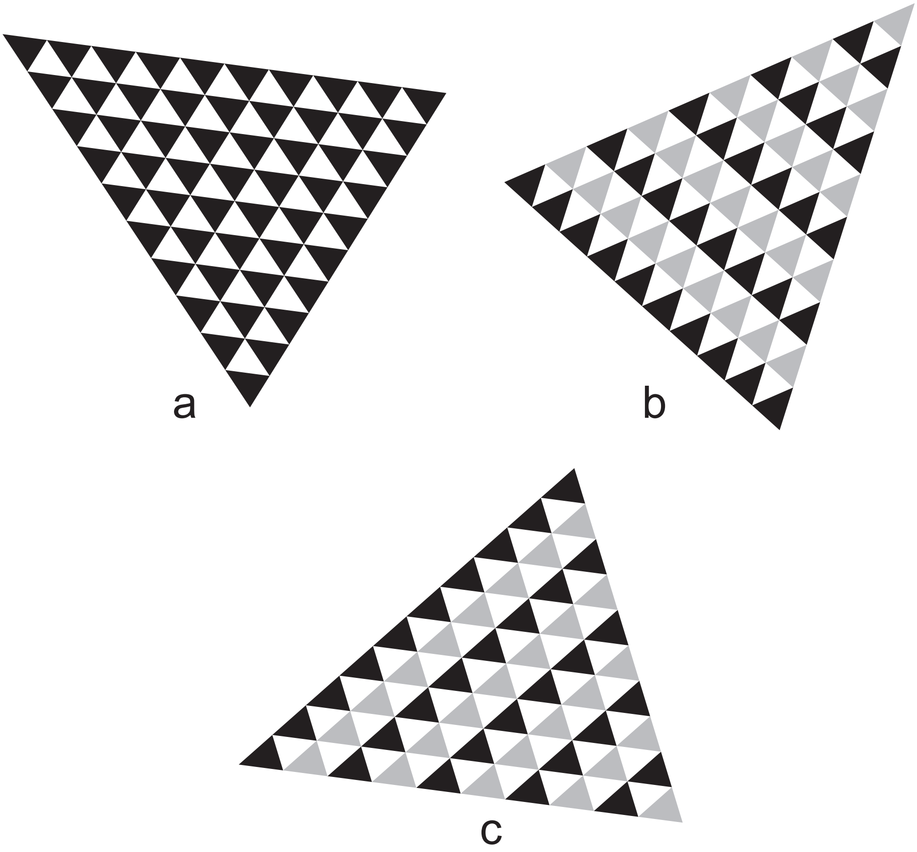


Figure 5. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

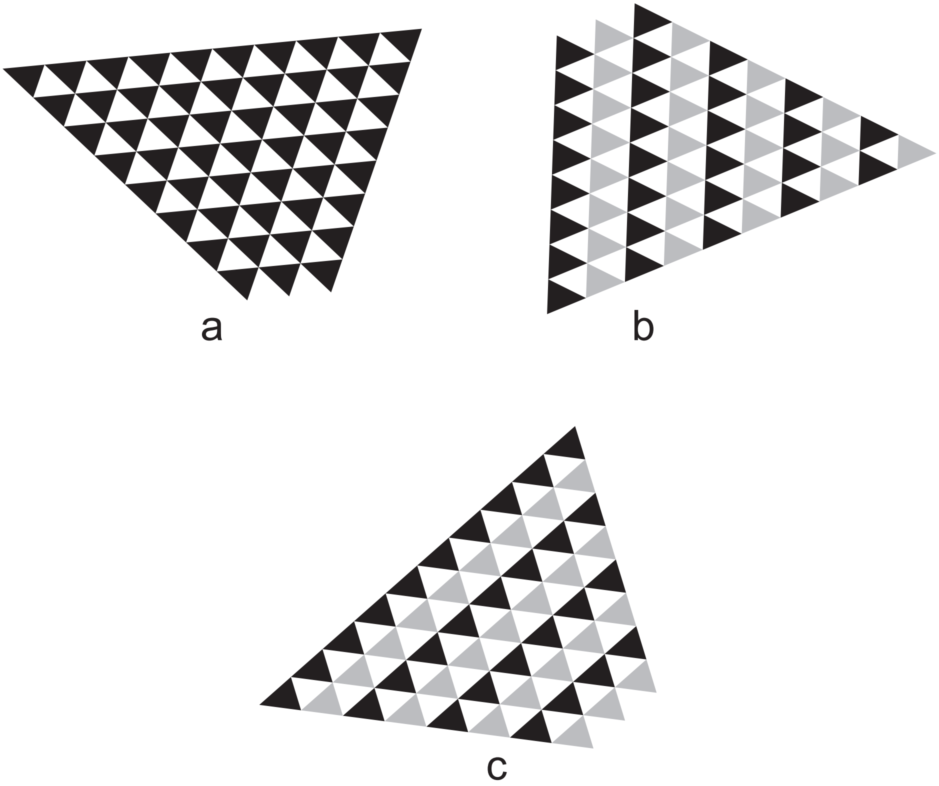
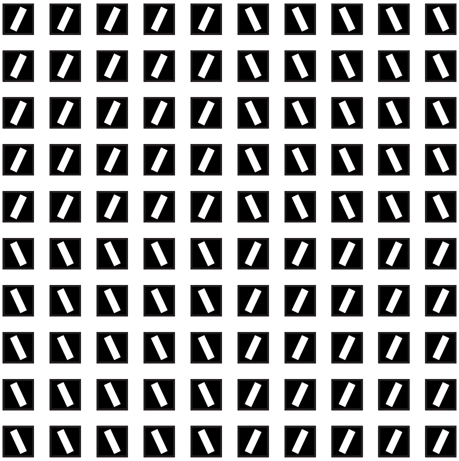
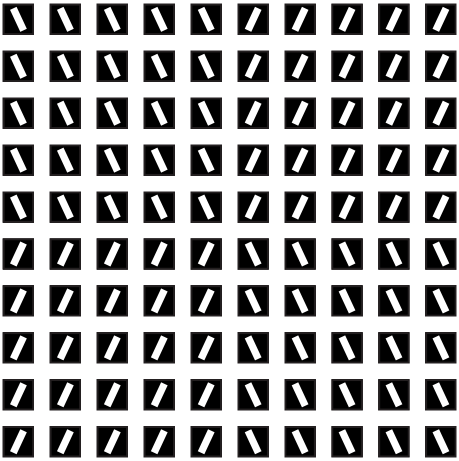


Figure 6. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



a



b

Figure 7. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

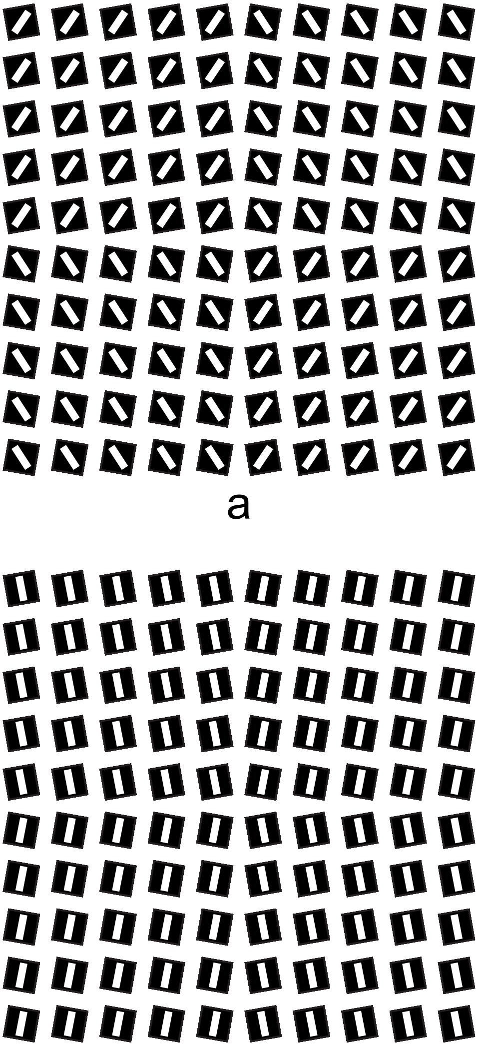


Figure 8. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

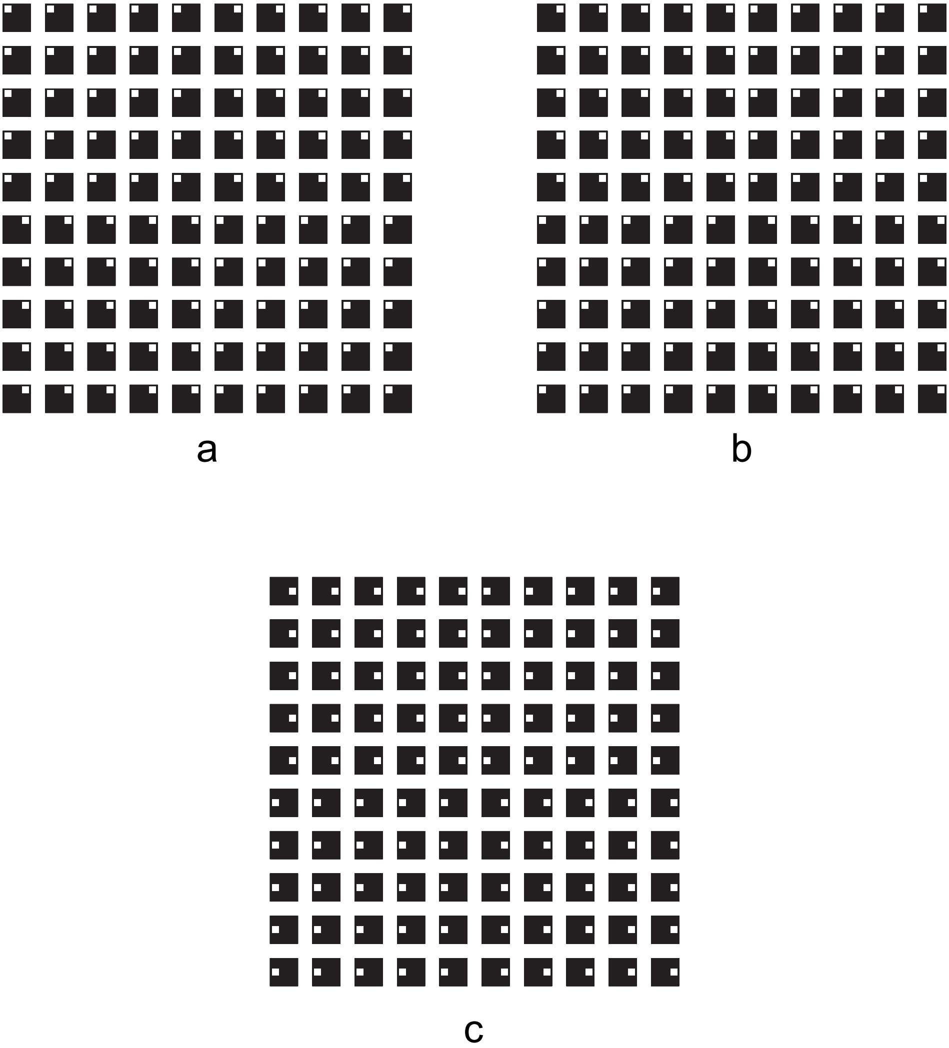
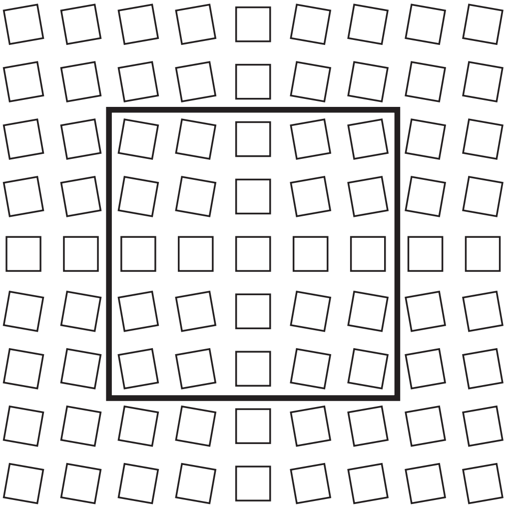


Figure 9. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



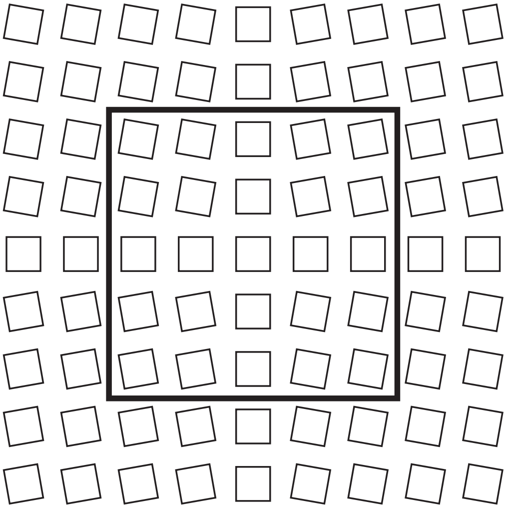


Figure 10. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

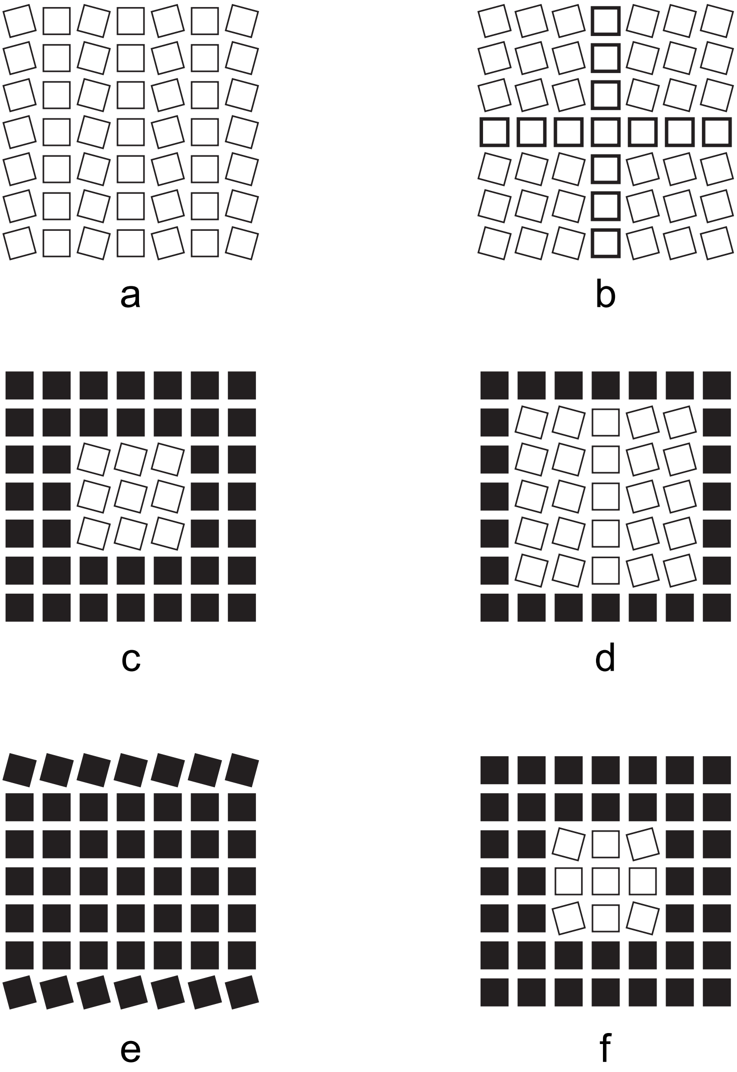


Figure 11. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

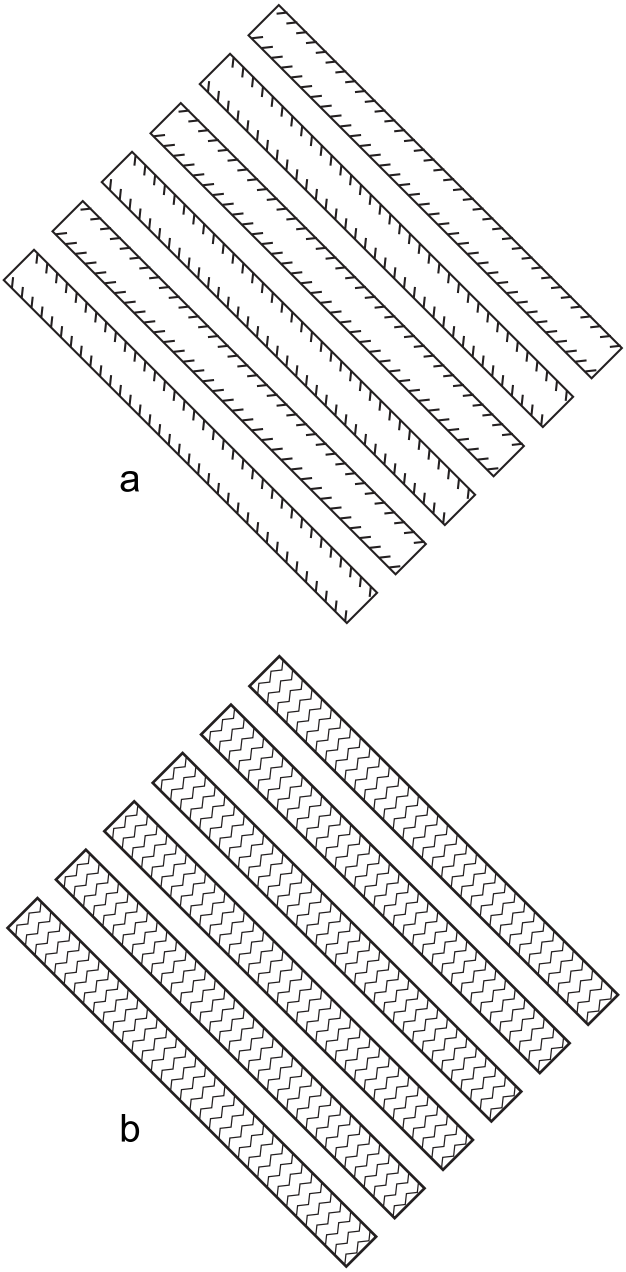


Figure 12. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

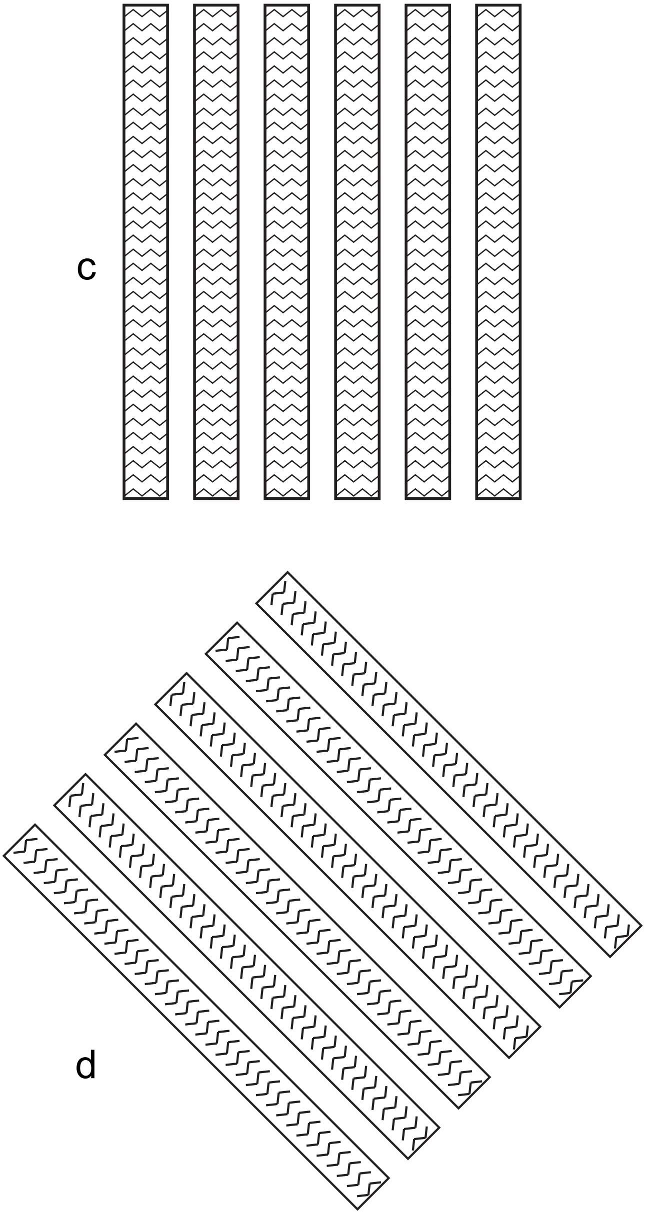


Figure 12. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

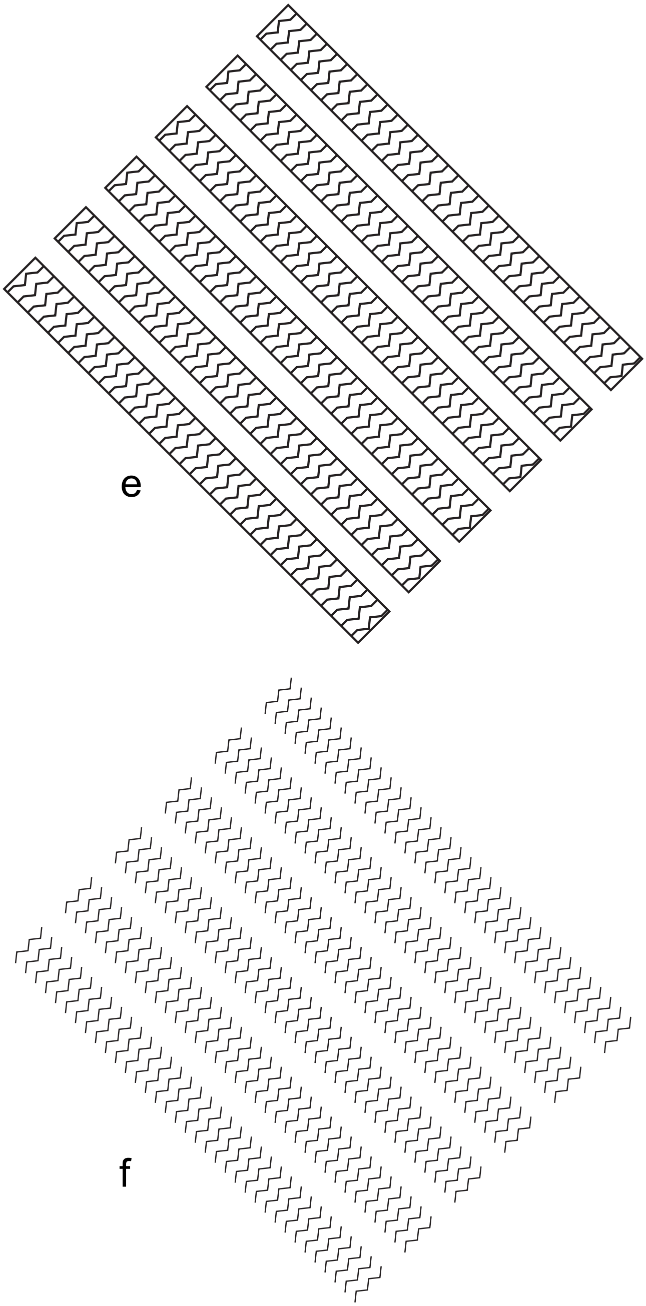


Figure 12. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

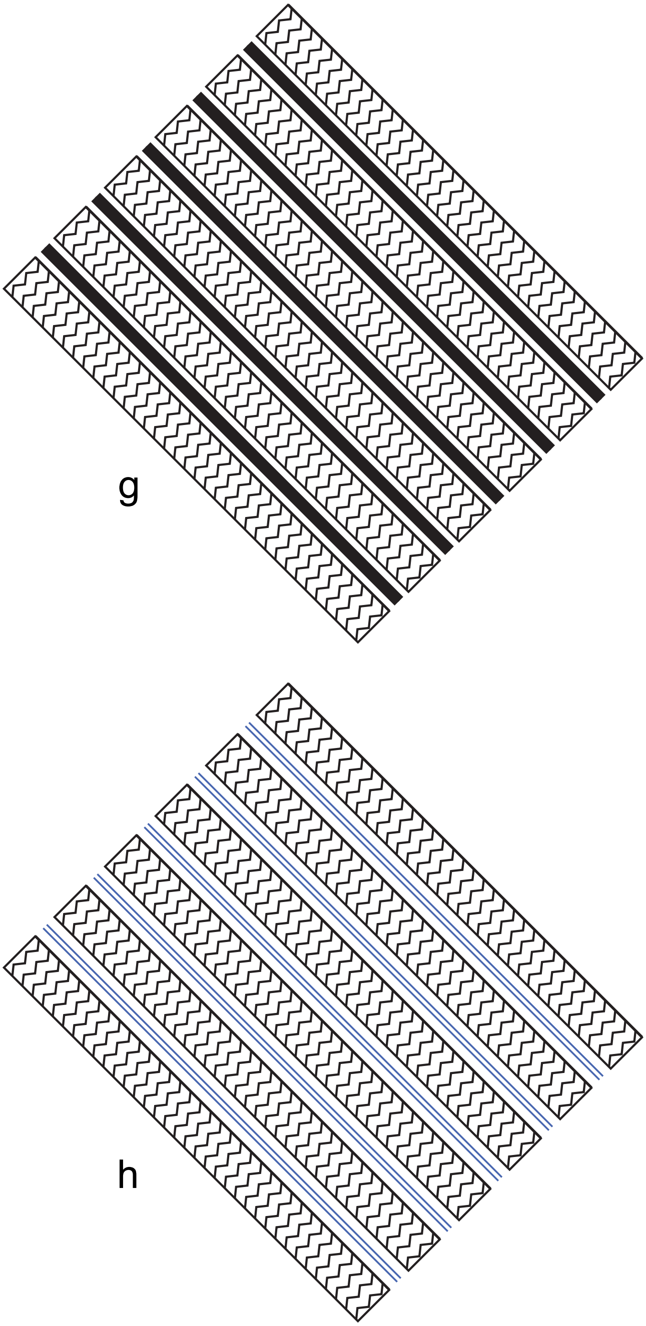


Figure 12. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

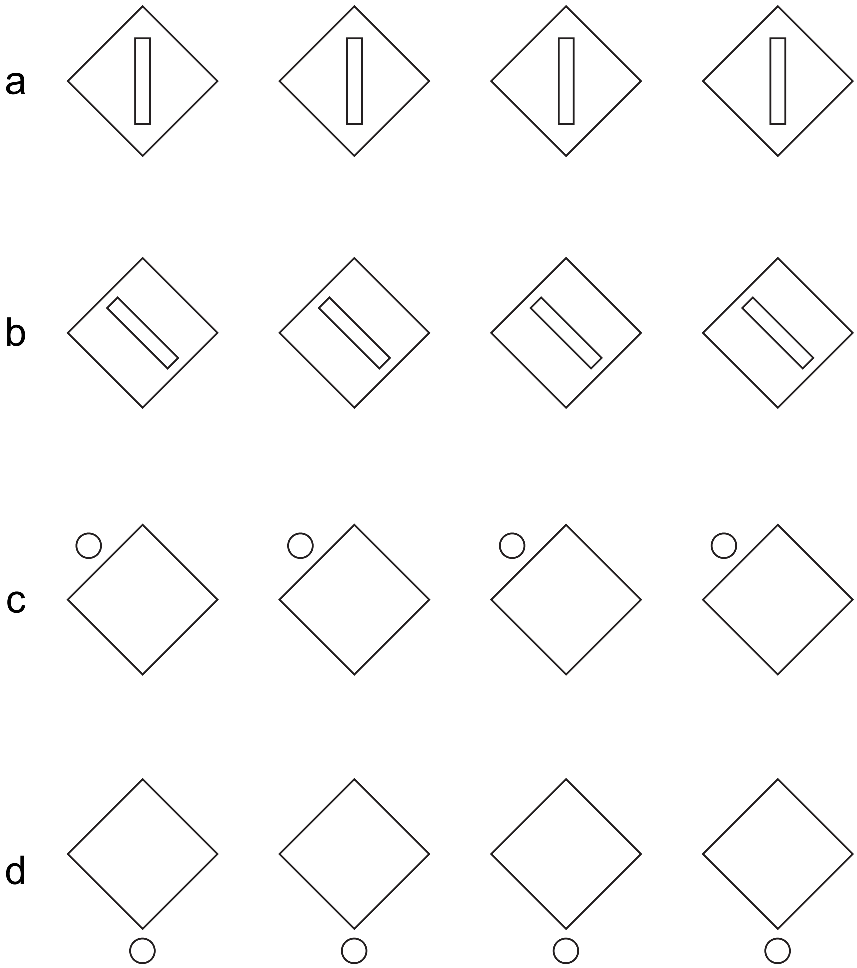


Figure 13. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

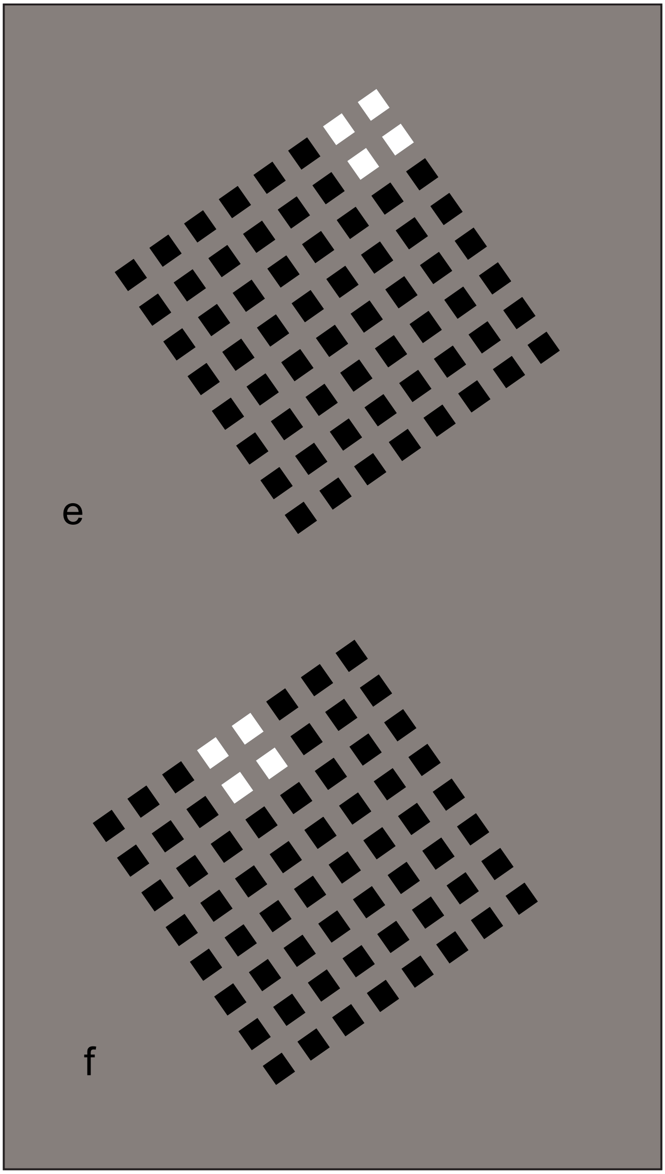


Figure 13. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

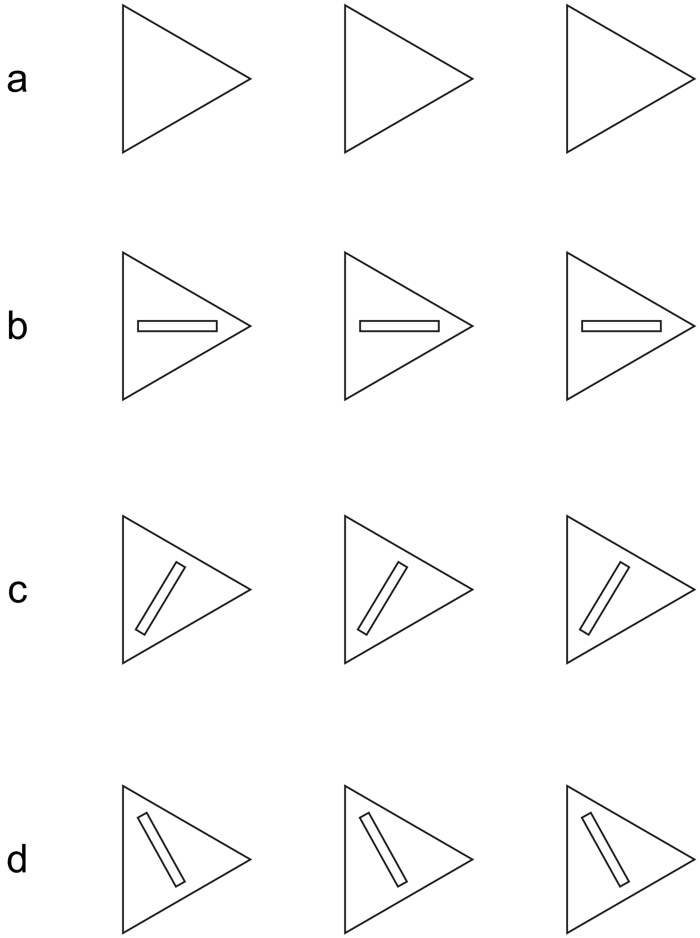


Figure 14. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

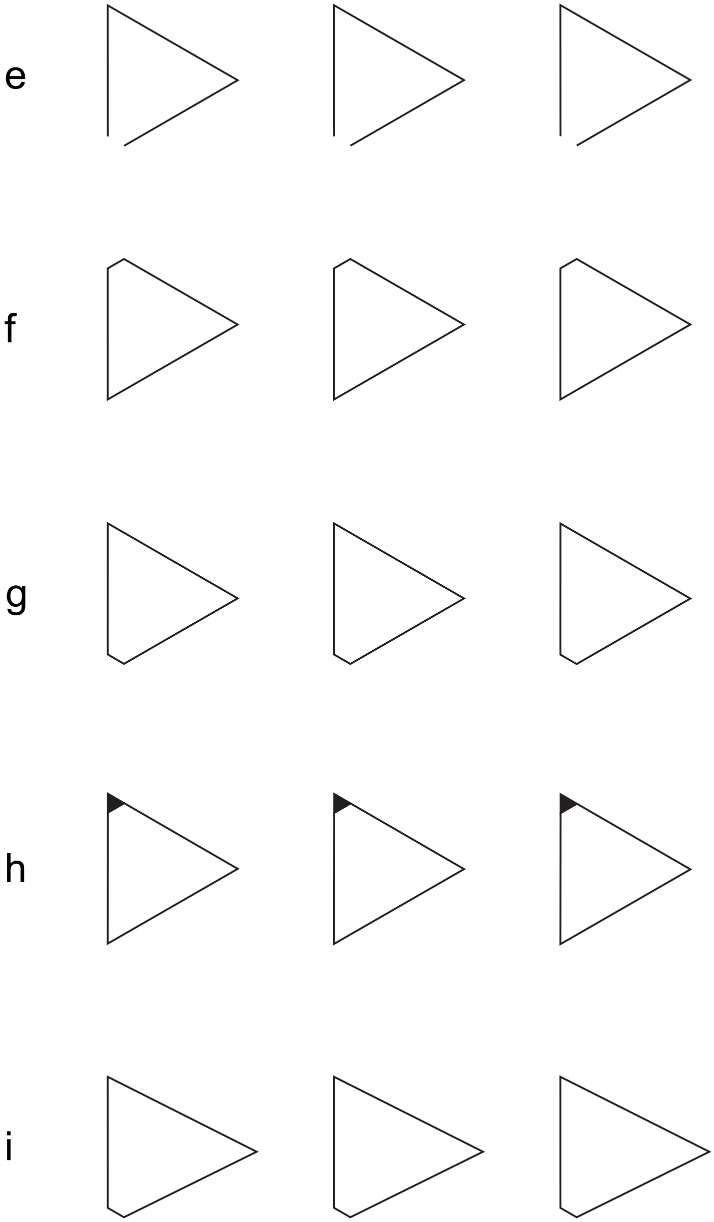


Figure 14. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

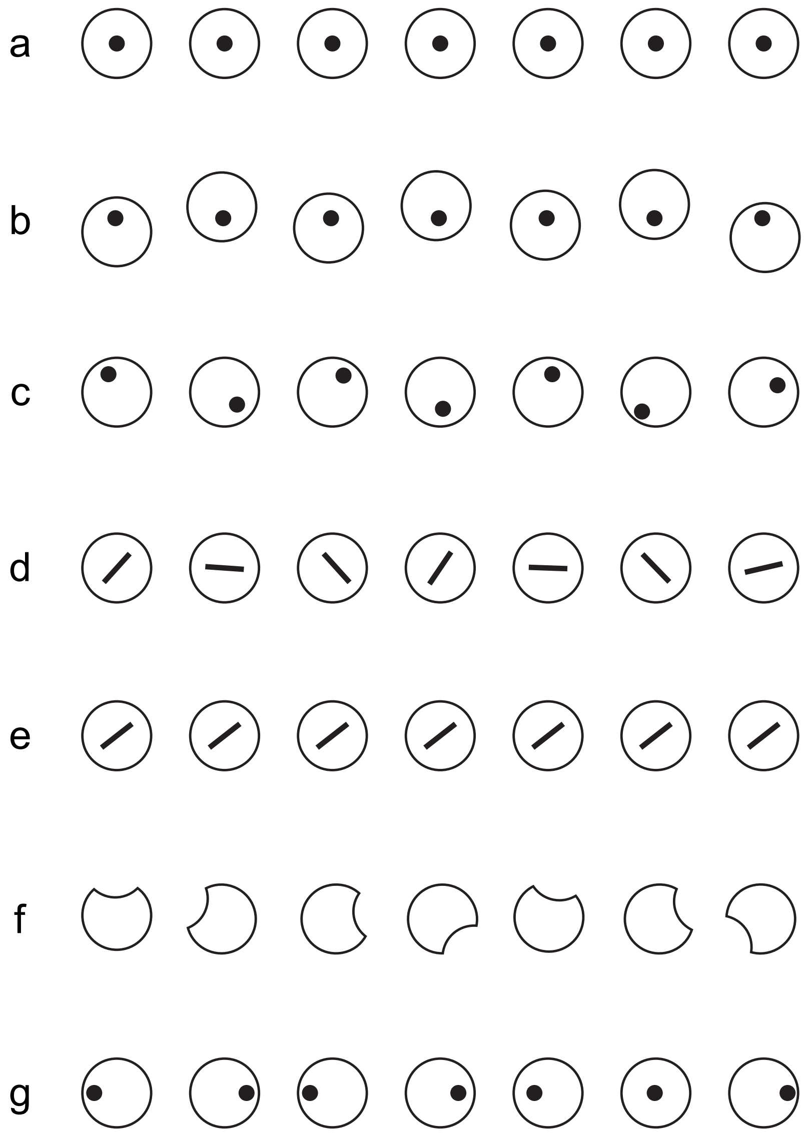
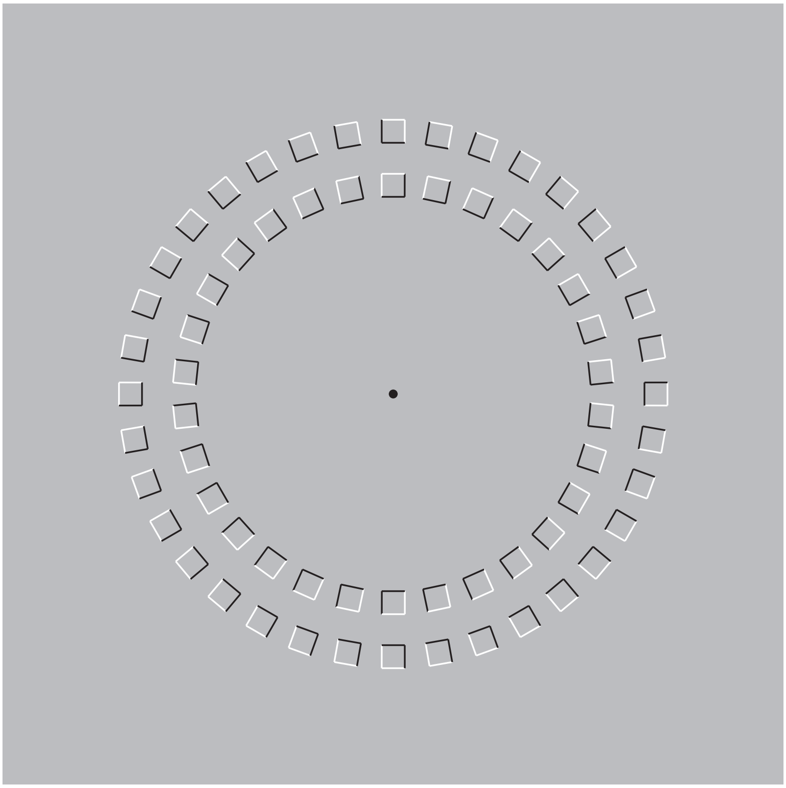
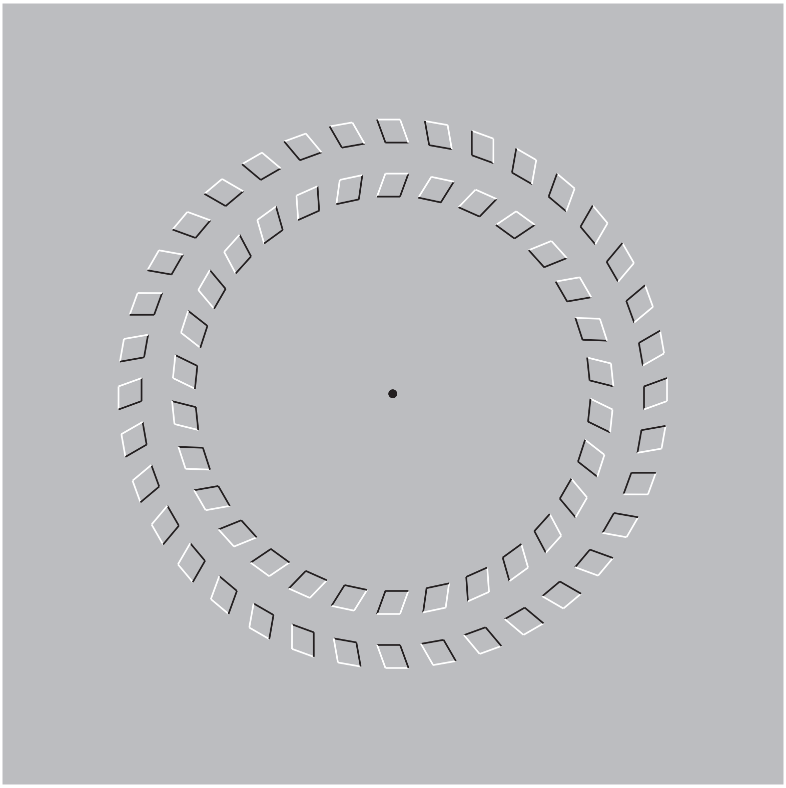


Figure 15. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



a

Figure 16. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



b

Figure 16. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

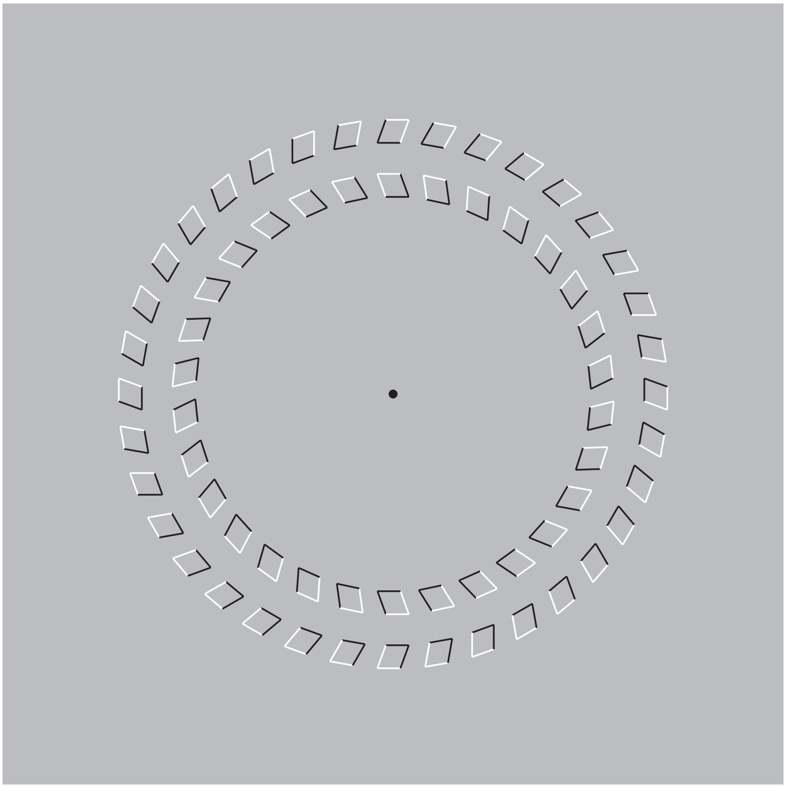
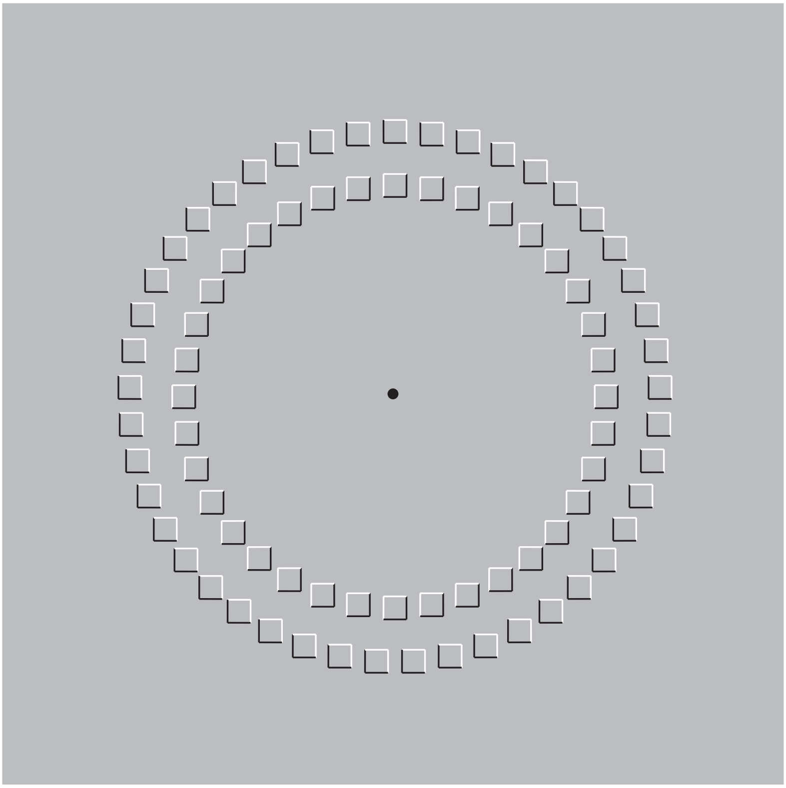


Figure 16. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



d

Figure 16. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

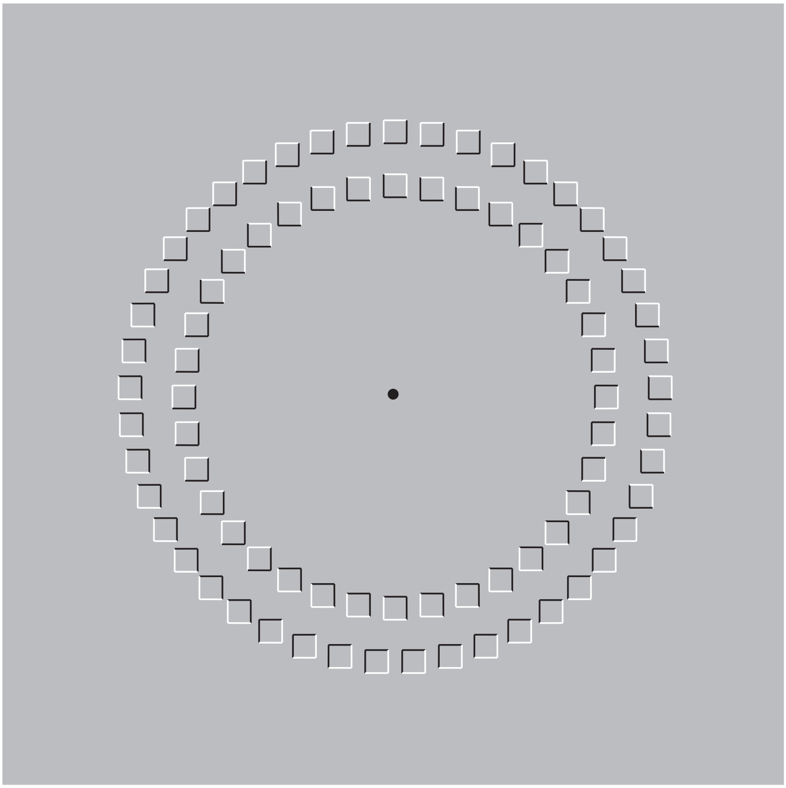


Figure 16. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

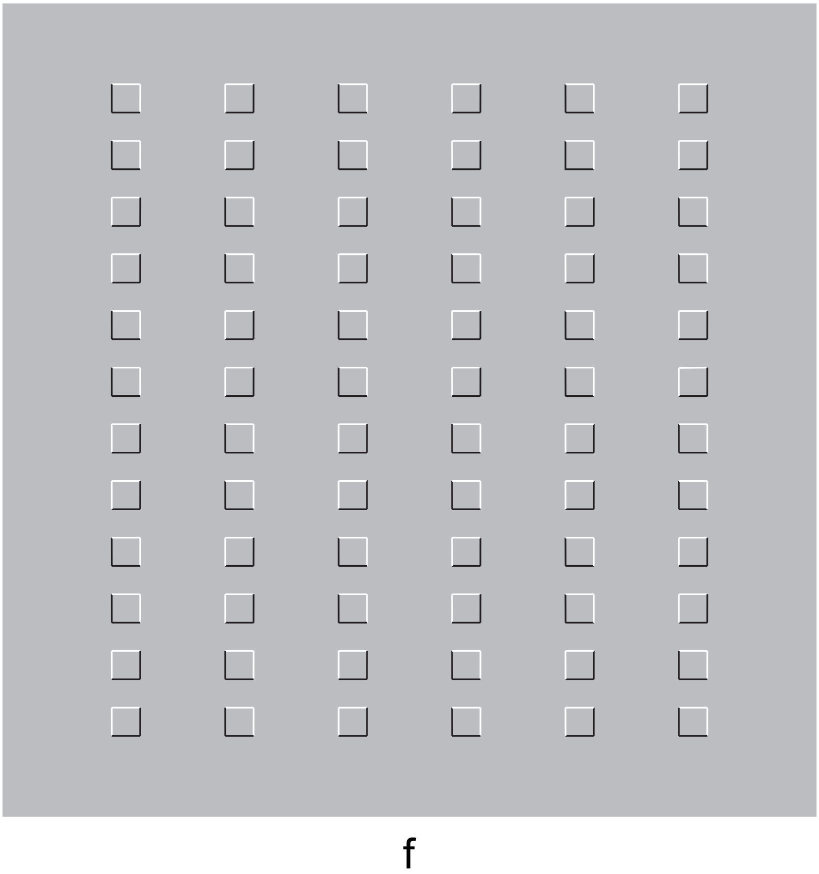


Figure 16. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

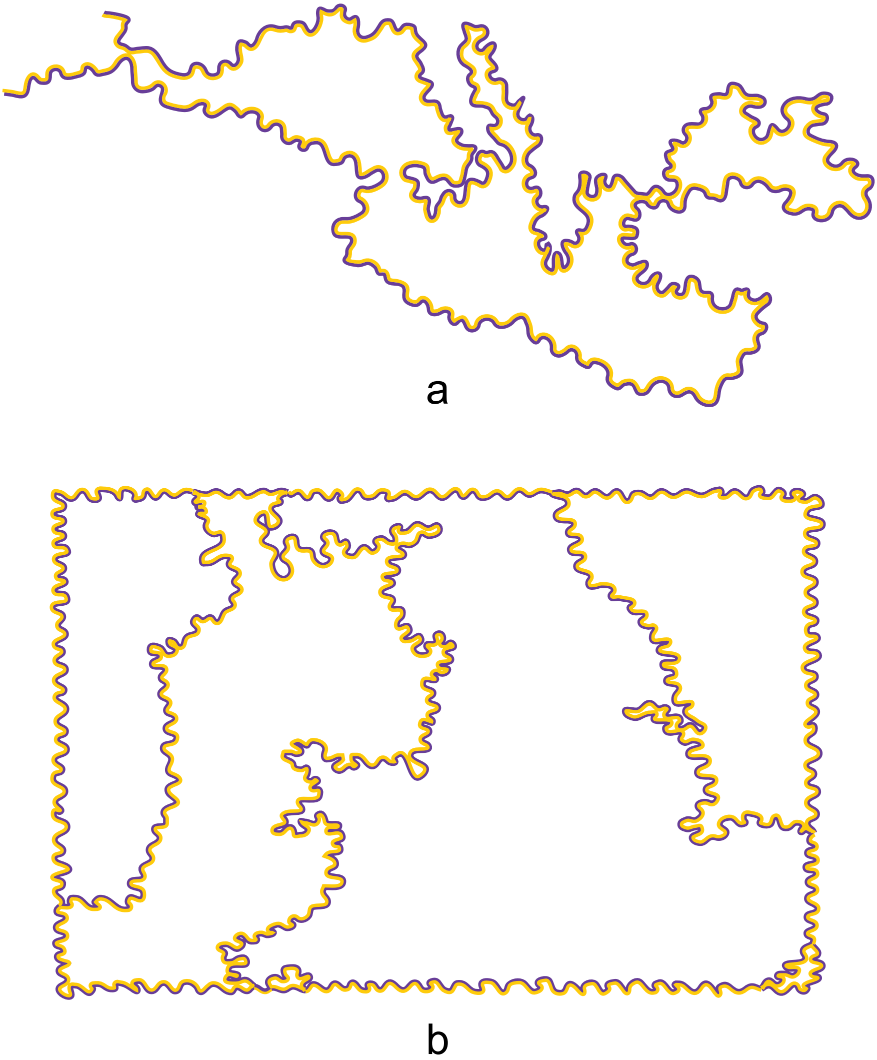


Figure 17. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

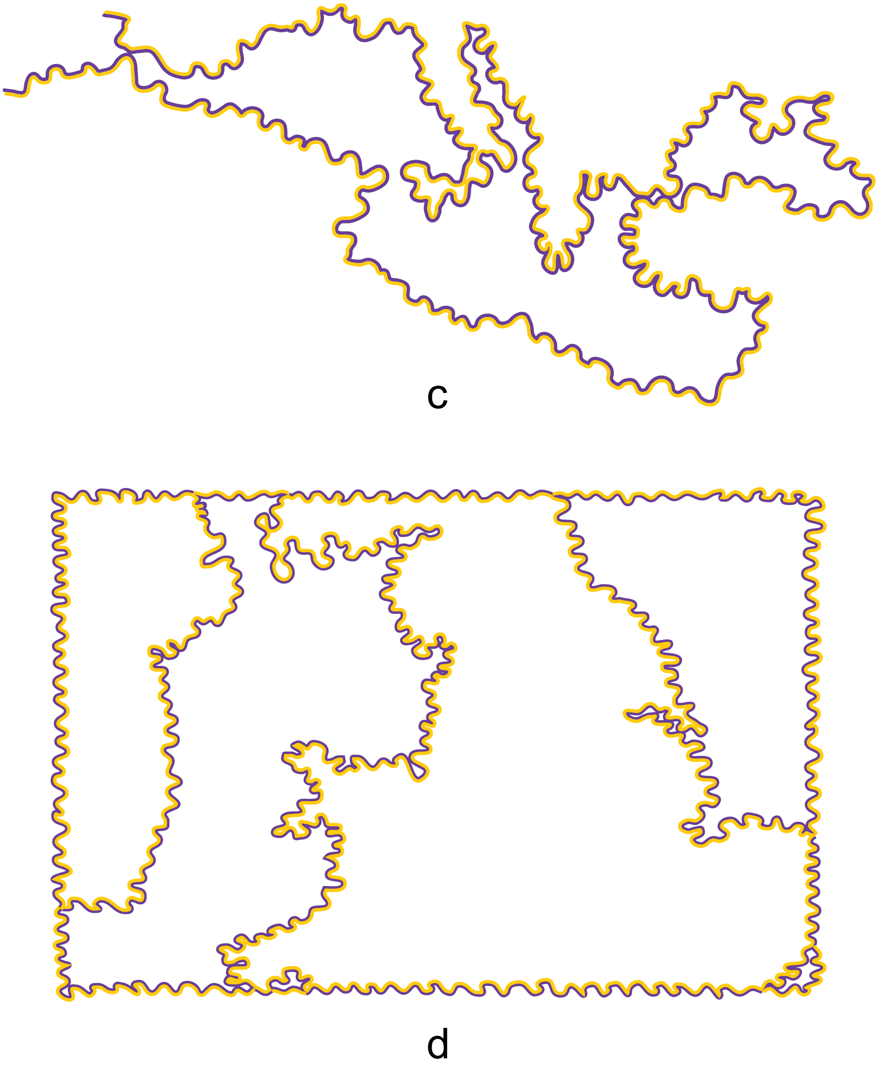
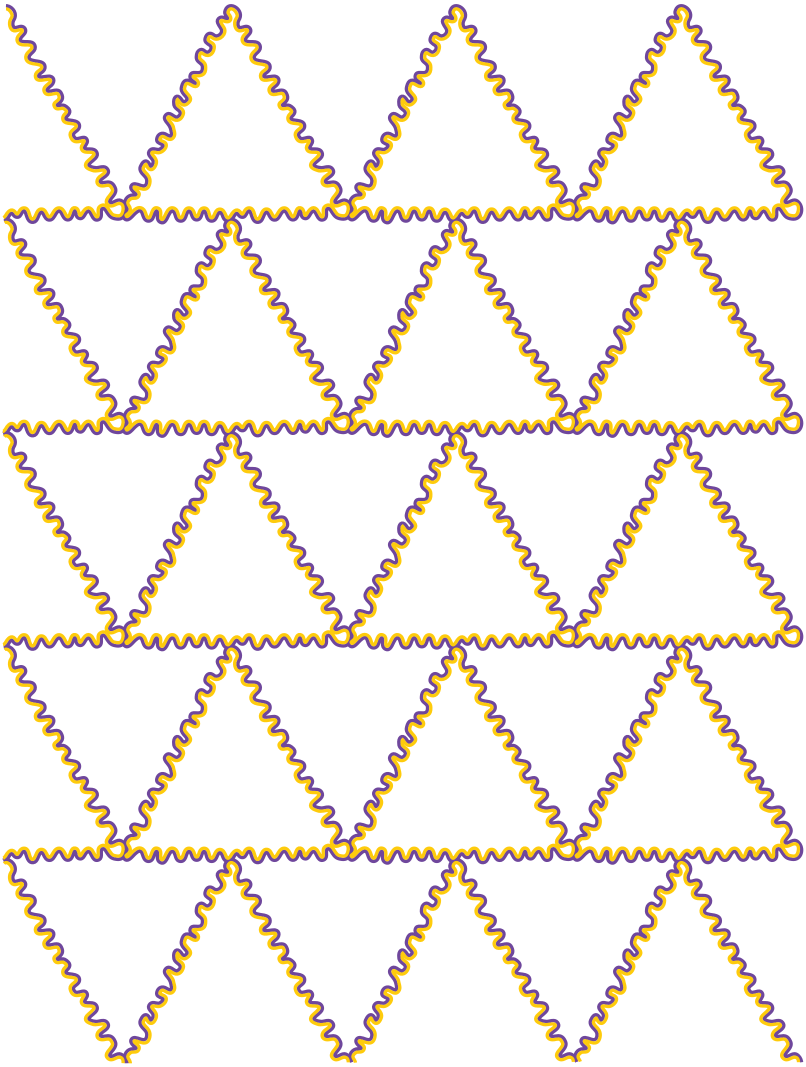


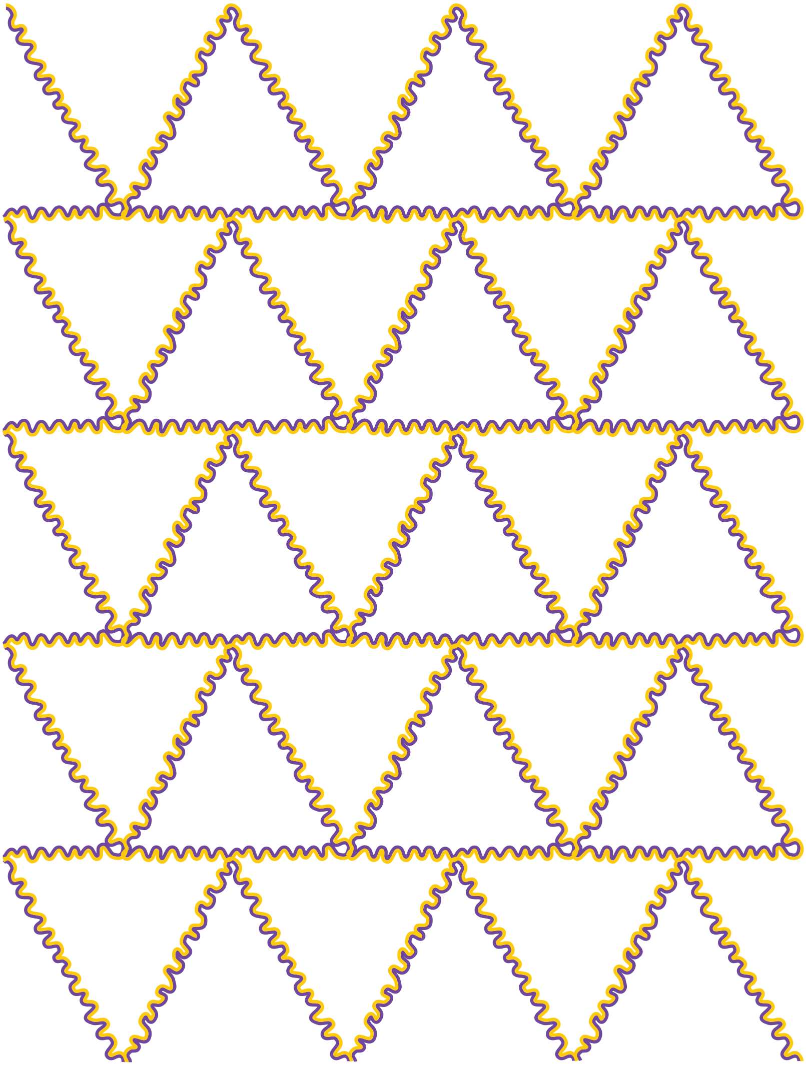
Figure 17. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



e

Figure 17. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



f

Figure 17. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

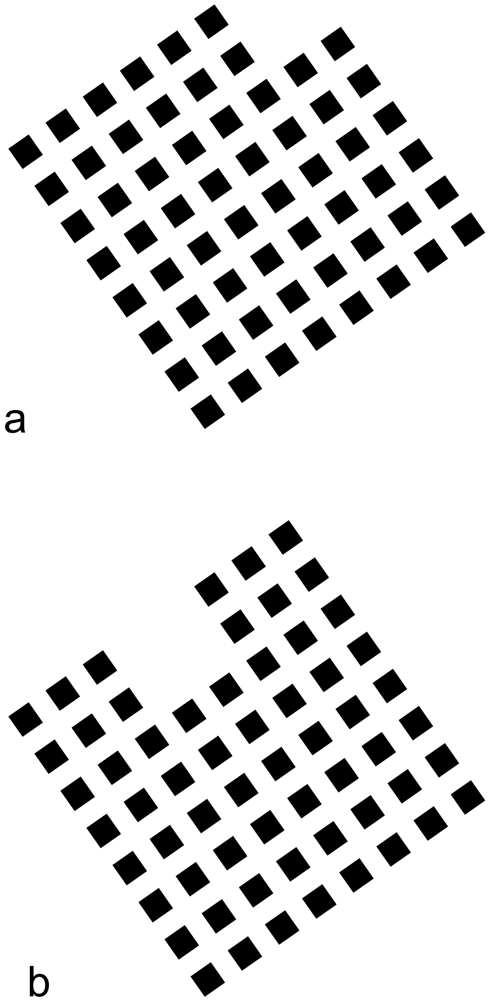


Figure 18. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

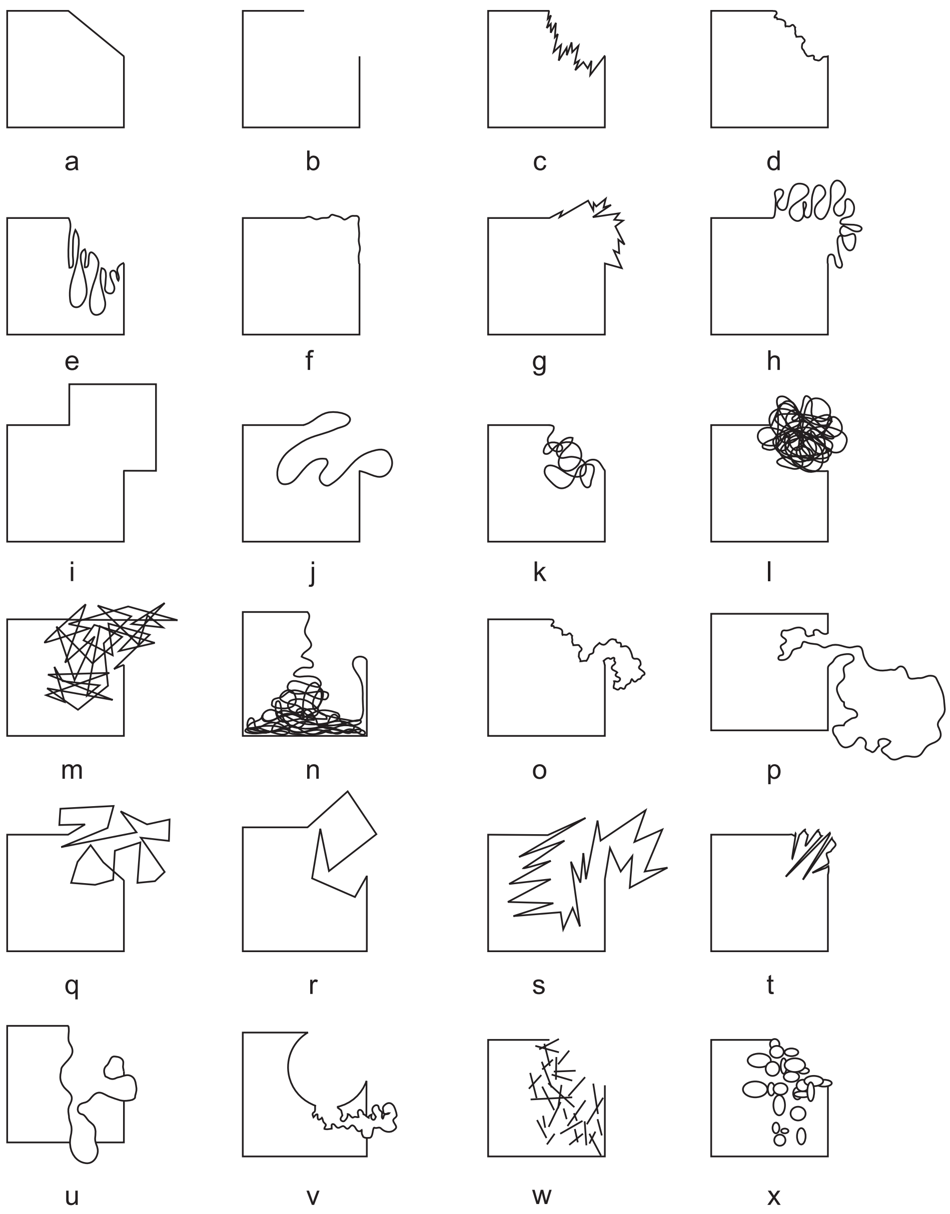


Figure 19. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

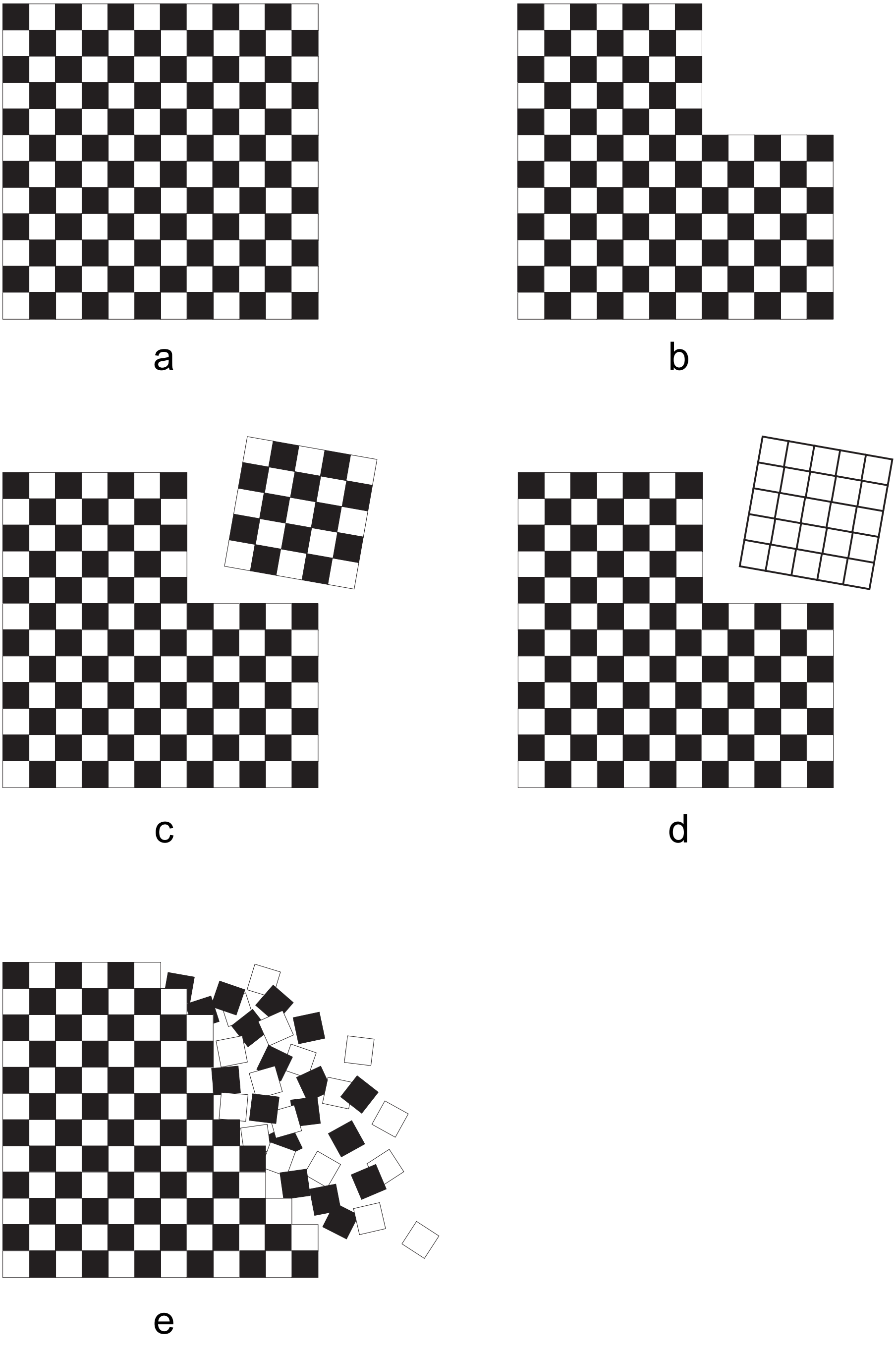
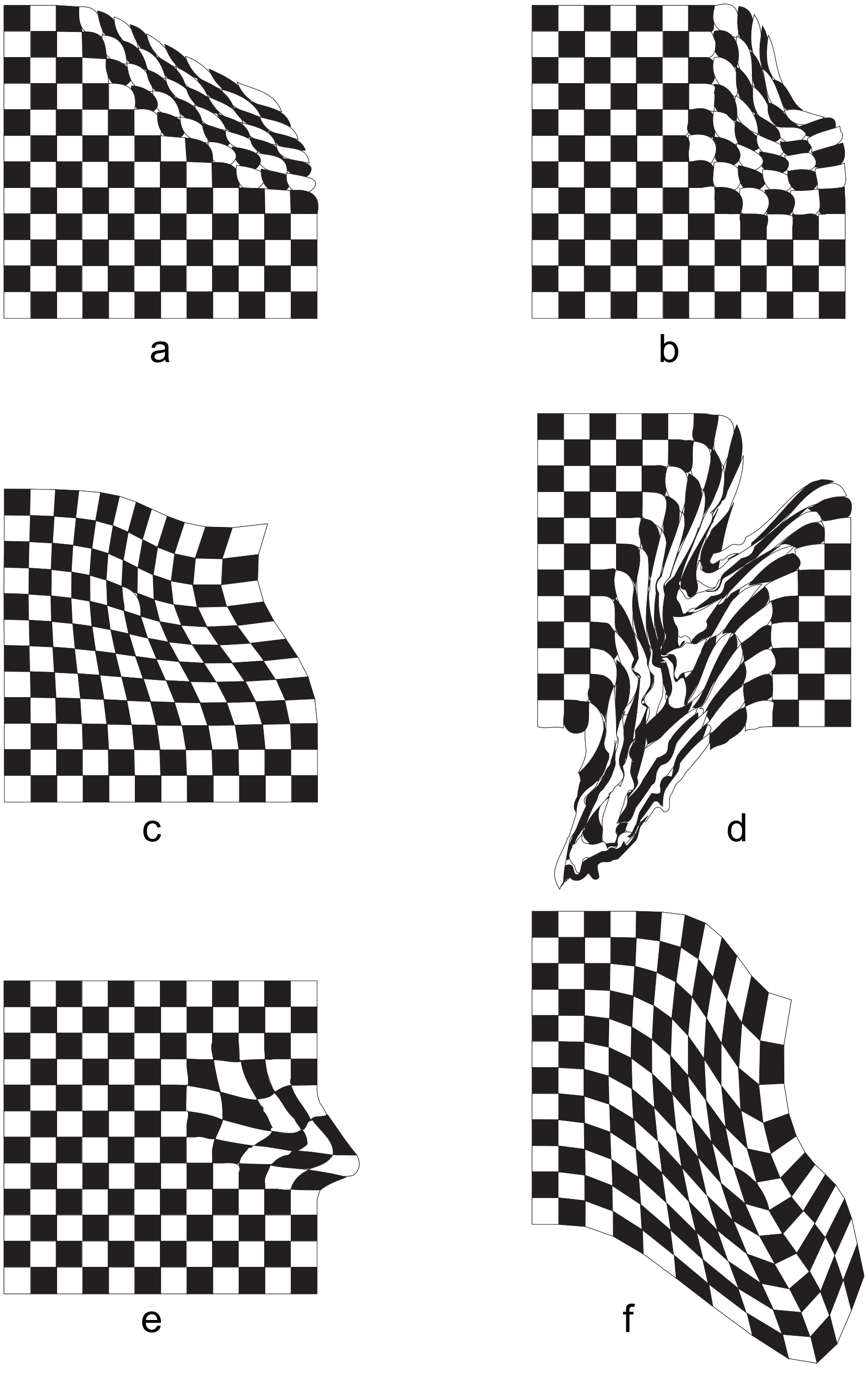


Figure 20. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning



a

b

c

d

e

f

Figure 21. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

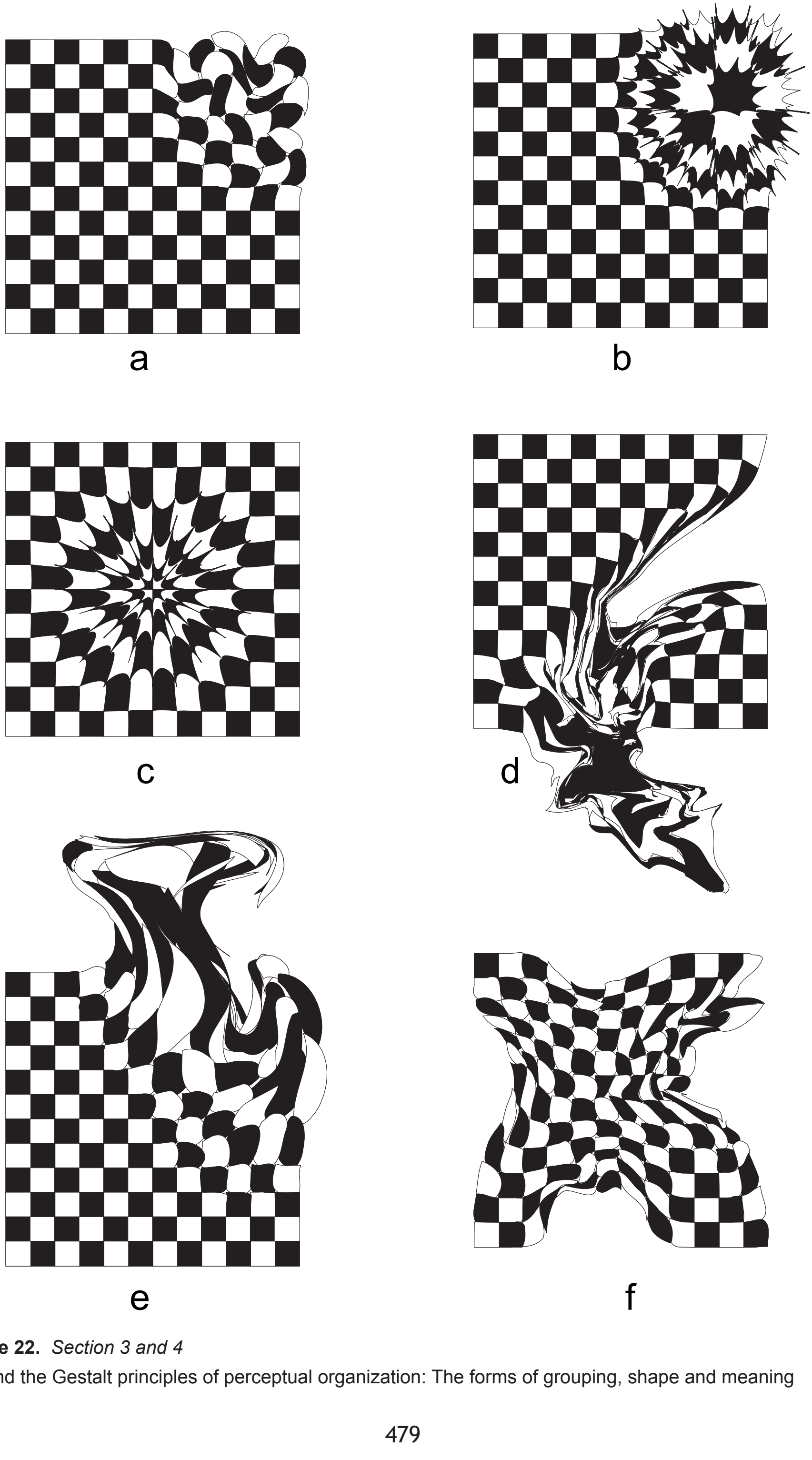


Figure 22. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

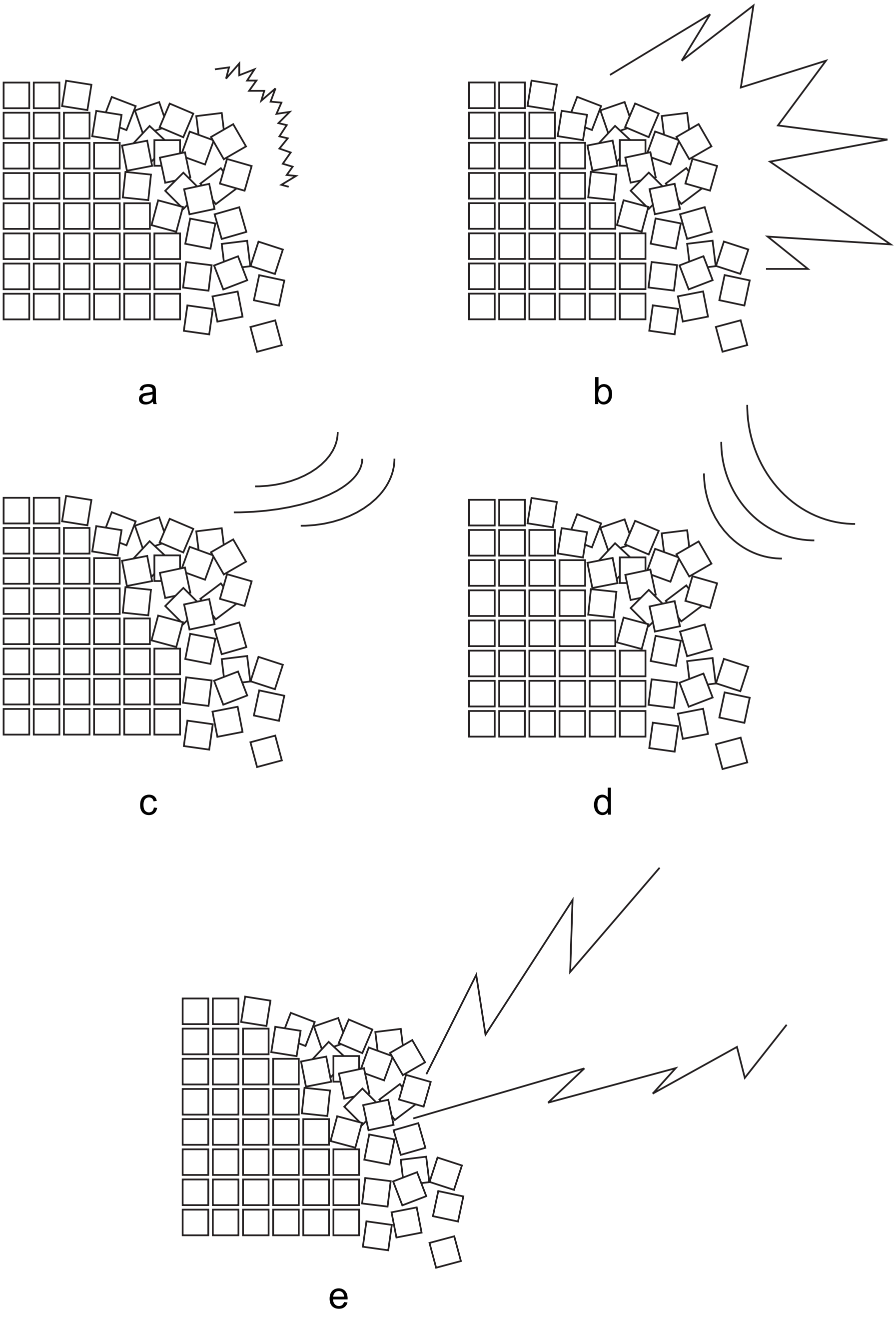


Figure 23. *Section 3 and 4*

Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning

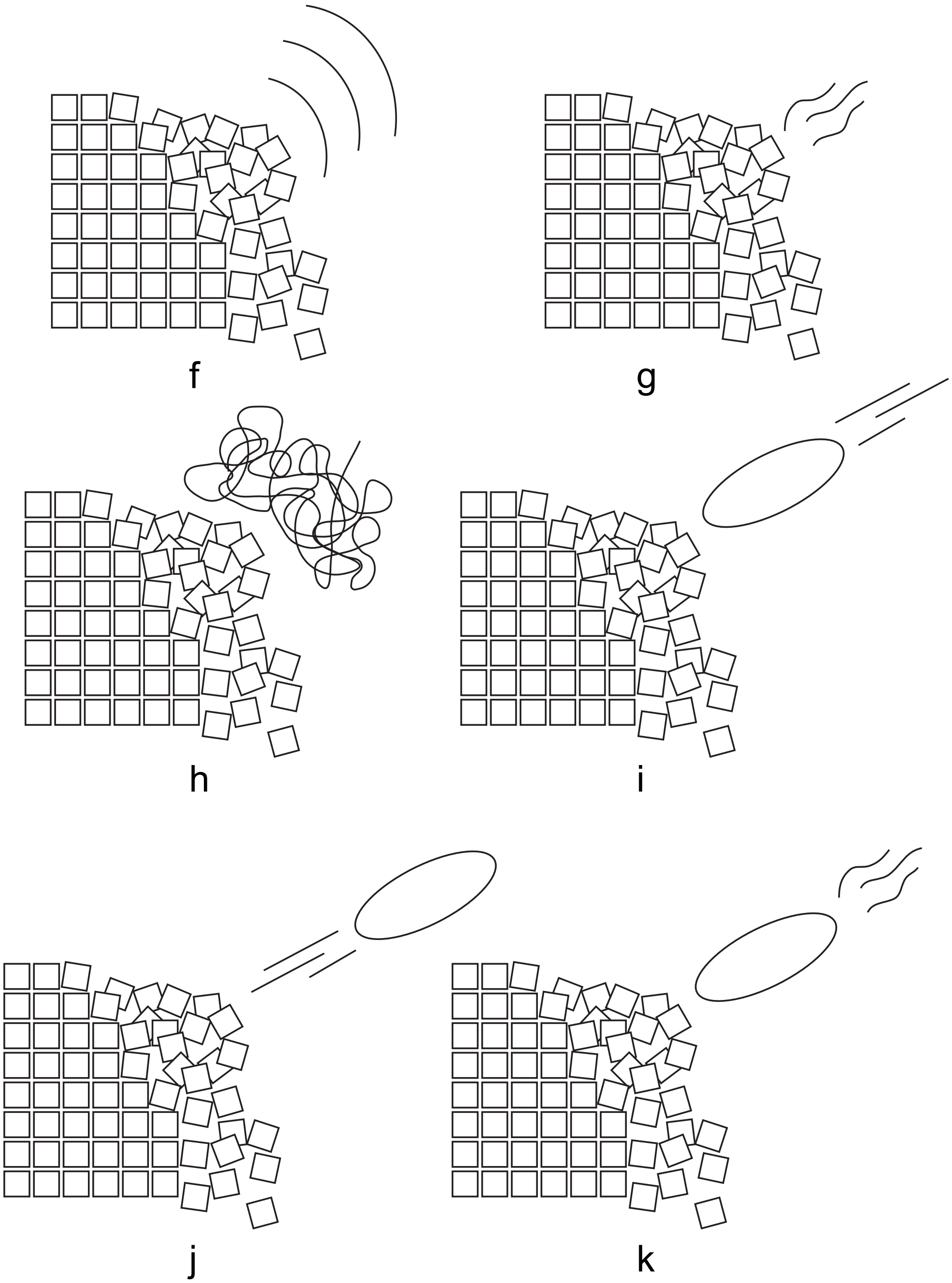


Figure 23. *Section 3 and 4*
Beyond the Gestalt principles of perceptual organization: The forms of grouping, shape and meaning