Suzuki, S. (2005). High-level pattern coding revealed by brief shape aftereffects. In C. Clifford and G. Rhodes (Eds.), Fitting the mind to the world: adaptation and aftereffects in high-level vision (Advances in Visual Cognition Series, Vol. 2), Oxford University Press.

5

High-Level Pattern Coding Revealed by Brief Shape Aftereffects

SATORU SUZUKI

5.1 Relating Neural Response Properties to Perception of Form Features

Visual form processing is often considered hierarchical in that patterns of illumination on the retina (detected by the rods and cones) are gradually transformed through subcortical and cortical processes into codings of increasingly global and complex patterns[1]. A challenge is to understand what neural codings are used at different processing stages and how they contribute to what people see from moment to moment. On the one hand, detailed neurophysiological research is necessary to understand the variety of responses, interactions and organizations existing in different visual areas of the brain. On the other hand, complementary behavioral research is necessary to infer principles (e.g. computational algorithms) that link the physiological properties and anatomical organizations of visual neurons to the coding of perceived forms.

[1] Strictly speaking, the traditional view that the visual form processing is hierarchical is true only with respect to feedforward neural activation. Increasing evidence suggests that consciously perceived patterns are closely associated with feedback (or reentrant) activation of V1 following activation of higher cortical visual areas (e.g. Zipser *et al.* 1996; Dehaene *et al.* 1998; Enns & Di Lollo 2000; Lamme 2001; Pascal-Leone & Walsh 2001).

At the initial stage of cortical visual processing, there is a relatively clear relationship between the variety of neural pattern selectivity and the perceptual features encoded. For example, cells in V1 (the first stage of cortical form processing) have small receptive fields (i.e. each neuron "looks at" a tiny portion of the retina) and are systematically tuned to different orientations and spatial frequencies (e.g. Hubel & Wiesel 1968; De Valois *et al.* 1982). Consistent with these systematic neural tunings, psychophysical studies using patterns designed to stimulate V1 cells (e.g. oriented bars, gratings, and Gabor patches) have revealed repulsive perceptual aftereffects for orientation and spatial frequency (e.g. Gibson & Radner 1937; Blakemore & Sutton 1969; Blakemore *et al.* 1970; Mitchell & Muir 1976; Magnussen & Kurtenbach 1980). Demonstrations of these repulsive aftereffects provided behavioral evidence that perception of orientation and spatial frequency are population coded on the basis of central tendencies of activity[2] of neural units tuned to orientation and spatial-frequency (see Fig. 5.1 for an illustration using tilt aftereffect as an example, and see Section 5.7 for details). The fact that these aftereffects are local (thus indicative of mediation by cells with small receptive fields) and that their tuning properties are similar to those of orientation and spatial-frequency tuned cells in V1, have further supported the idea that perception of local orientation and spatial frequency are population coded in V1 (e.g. Blakemore & Nachmias 1971; Braddick *et al.* 1978; Wenderoth & van der Zwan 1989).

Unfortunately, this clear relationship between organization of neural tuning and coding of perceived features quickly becomes murky as one examines higher cortical visual areas in the ventral visual pathway, V2 → V4 → IT (inferotemporal cortex), thought to mediate form perception (e.g. Ungerleider & Mishkin 1982; Mishkin *et al.* 1983). Though some tunings for orientation and/or spatial-frequency persist in V2, V4, and IT (e.g. Foster *et al.* 1985; Levitt *et al.* 1994; Desimone & Schein 1987; Vogels & Orban 1994), cells in higher visual areas respond preferentially to increasingly complex patterns. Even in V2 (an area adjacent to V1), cells begin to respond to local but complex geometric features such as subjective contours, bent or curved contours, contour intersections, and texture patterns (e.g. Peterhans & von der Heydt 1991; Sheth *et al.* 1996; Leventhal *et al.* 1998; Hegdé & Van Essen 2000). The V4 cells have larger receptive fields

[2] Here, it is assumed that perception of form features (presented in isolation or at the focus of attention) are primarily coded by neural population activity on the basis of activation strengths (spike rates) derived from classical receptive fields. It should be noted that recent physiological research has found that (1) neural responses as early as in V1 are influenced by grouping and figure-ground organizations existing well beyond the bounds of classical receptive fields (e.g. Zipser *et al.* 1996; Lamme *et al.* 1999; Nothdurft *et al.* 1999), and (2) some image properties may be coded in V1 as temporal patterns of neural activity (e.g. Ferster & Spruston 1995; Ikegaya *et al.* 2004) providing a means other than activation strengths (i.e. spike rates) to influence downstream neural activity (e.g. Lumer 1998; Fries *et al.* 2001; Reich *et al.* 2001; Suzuki & Grabowecky 2002*b*; Suzuki 2003*b*).

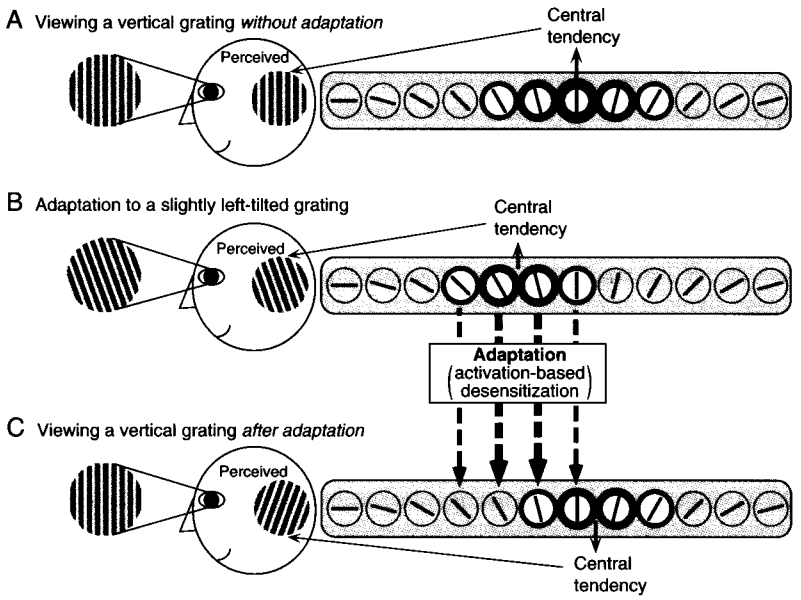

FIG. 5.1 An illustration of how the tilt aftereffect provides behavioral evidence that perceived orientation is coded on the basis of a central tendency of population activity of orientation-tuned cells. The circles with oriented lines represent orientation-tuned cells in the primary visual cortex (V1) with the indicated orientation preferences. A thicker circle indicates stronger neural activation. (A) When a vertical grating is viewed, the vertically tuned cells respond maximally, but cells that are tuned to the neighboring orientations also respond in a graded manner because of their tuning widths. Because the pattern of activation is symmetric about the vertical-tuned cell, the central tendency of the population activity falls at the vertical orientation, and the perceived orientation is vertical (see the grating depicted inside the schematic head of the viewer). (B) When a slightly left-tilted grating is viewed, the central tendency falls at the corresponding orientation (in this case between the preferred orientations of adjacent cells), and a slightly left-tilted orientation is perceived. Importantly, prolonged viewing of this grating results in adaptation (activation-based desensitization which is characteristic of most cortical cells) of the activated cells. (C) When a vertical grating is viewed following this adaptation, the central tendency of population activity shifts to a slightly right-tilted orientation due to the prior desensitization of the cells tuned to left-tilted orientations (dashed arrows). Thus, if perception of orientation is determined by the central tendency of population activity of orientation-tuned cells, the vertical grating should appear slightly tilted to the right; this indeed occurs and the phenomenon is known as the tilt aftereffect.

and respond to combinations and relative positions of curves and angles (e.g. Pasupathy & Connor 1999, 2001), texture patterns (e.g. Gallant *et al.* 1996) and luminance gradients (e.g. Hanazawa & Komatsu 2001). In inferotemporal cortex (IT), cells have very large receptive fields and they respond to a variety of simple and complex geometric patterns with substantial degrees of invariance

for position and size (e.g. Desimone, *et al.* 1984; Fujita *et al.* 1992; Logothetis & Sheinberg 1996; Tanaka 1996; Tsunoda *et al.* 2001). Because cells in these higher visual areas respond to a variety of different geometric features with little apparent systematicity (such as the systematic orientation and spatial-frequency tunings in V1), it has been difficult to understand how neural responses in these areas encode perceived features.

5.2 Some Basic Types of Neural Population Coding

At the most general level of conceptualizing neural population coding, perceived objects are encoded by overall patterns of distributed neural activity. For example, different classes of common objects (e.g. chairs, houses, bottles, shoes, scissors, and faces) evoke different patterns of neural activity distributed across the temporal visual areas (e.g. Haxby *et al.* 2000, 2001; Tsunoda *et al.* 2001). Objects can then be distinguished on the basis of differences in the general topography of distributed neural activity, providing virtually unlimited capacity for coding different objects. In this *general-distributed coding*, objects inducing less overlapped neural activities are more readily discriminated (e.g. Abdi *et al.* 2003). General-distributed coding is difficult to crack because one must unravel the principles that relate perceived properties of object features to specific topographies of distributed neural activity.

A subset of distributed coding might include *process-specific coding* where anatomically localized cell assemblies might encode image features that are useful for specific types of processing. Expert object identification (e.g. identification of faces and other objects, such as birds and cars, in which one is an expert), for example, is likely to require encoding of specific feature dimensions that are maximally sensitive to identification-relevant image variations (e.g. Young & Yamane 1992; Kobatake *et al.* 1998; Sigala & Logothetis 2002) but minimally sensitive to irrelevant image variations including spurious variations due to changes in illumination, size, orientation, and viewpoint. In contrast, perception of spatial layout, for example, imposes a different (and conflicting) set of computational demands, requiring viewpoint-dependent encoding of relative positions, sizes and orientations of surfaces and objects to support visually guided action. Human brain imaging (fMRI) results suggest that the processing of objects of expertise preferentially activates the fusiform area (e.g. Kanwisher *et al.* 1997; Gauthier *et al.* 1999, 2000; Gauthier 2000; Rhodes *et al.* 2004), whereas the processing of spatial layouts preferentially activates the parahippocampal area (e.g. Epstein *et al.* 1999, 2003) of the temporal lobe. Whereas any arbitrary categories of objects can be distinguished on the basis of general topographies of distributed neural activity, process-specific codings, utilizing specialized cell assemblies, might have developed to optimize discriminability for specific classes of behaviorally-relevant stimuli (see Section 5.6), and to anatomically segregate computationally conflicting processes.

The primary topic of this chapter, *central-tendency coding*, is a special case of process-specific coding. In central-tendency coding, a geometric feature is coded by the central tendency of activity of a population of neural units that are systematically tuned (with overlapped tuning functions) to different values of that feature. For simplicity, discussions in this chapter will be limited to codings of single geometric features (e.g. convexity), though central-tendency codings commonly occur in multi-dimensional feature spaces as most visual neurons are broadly tuned to multiple image features (e.g. V1 cells are simultaneously tuned to orientation, spatial frequency, position, and binocular disparity). The extraction of central tendency might be based on the population activity centroid, vector summation, and/or maximum-likelihood analyses (e.g. Lee *et al.* 1988; Vogels 1990; Young & Yamane 1992; Deneve *et al.* 1999).

Any coding based on patterns of population activity provides resistance against individual (uncorrelated) and global (correlated) fluctuations in neural activity on the basis of averaging and extraction of relative activity. Central-tendency codings in particular can facilitate high-resolution discriminations of the coded features because partially overlapped tuning functions allow interpolations across discretely sampled feature values. For example, in Fig. 5.1(B), the orientation of the grating stimulus is not identical to the preferred orientation of any of the twelve hypothetical orientation-tuned units. However, because each unit has a tuning around its preferred orientation (responding to neighboring orientations in a graded manner), the central tendency can extract the stimulus orientation that falls between the two units, allowing interpolations across discretely sampled orientations. It is thus likely that geometric features that are central-tendency coded in high-level processing are ones for which it is behaviorally important to discriminate their subtle variations.

5.3 Neurophysiological Evidence for Possible High-level Central-tendency Codings of Form Features

In addition to the systematic tunings for local orientation and spatial frequency in V1 and V2 (and to some degree in V4), sub-populations of V2 and V4 cells exhibit tunings for local curvature (Pasupathy & Connor 1999; Hegdé & Van Essen 2000). In V4, however, this "curvature tuning" is complicated by the fact that cells show additional selectivities for sequences of adjacent curves as well as positions of the preferred curves relative to the center of the stimulus pattern (e.g. Pasupathy & Connor 2001). In IT, cells tuned to "similar" geometric patterns tend to be anatomically clustered within "feature columns" (e.g. Fujita *et al.* 1992; Tanaka 1996; G. Wang *et al.* 1996; Y. Wang *et al.* 2000; Tsunoda *et al.* 2001). However, because somewhat arbitrary collections of shapes have been used as stimuli in most studies investigating response properties of IT cells and a single IT cell can exhibit apparently heterogeneous tunings for multiple patterns (e.g. Y. Wang *et al.* 2000), it is not clear which of the large variety of patterns that have been shown to preferentially stimulate individual IT cells indicate central-tendency

coded features. There is also evidence that at least some complex patterns (particularly objects with multiple parts) are coded in IT on the basis of part-based distributed patterns of neural activity, in which each sub-cluster of activity appears to represent a part and/or a combination of parts (e.g. Tsunoda *et al.* 2001; Brincat & Connor 2004).

A small set of potentially central-tendency coded features have been reported in IT, however. Face-related features are probably the most well-known example. Neural tunings for variations in global face attributes have been reported, such as roundness of facial outline and amount of hair (Young & Yamane 1992), and facial expression (Hasselmo *et al.* 1989). An optical imaging study also showed that anatomically adjacent regions of IT responded to faces rotated in depth by successive angles (Wang *et al.* 1996). Thus, identity, expression, and 3D orientation of faces appear to be central-tendency coded in IT.

Recently, potential systematic central-tendency codings in IT have also been reported for some non-face features. Kayaert *et al.* (2003, 2004) examined responses of IT cells to various non-accidental geometric features (i.e. features that are preserved across minor viewpoint changes) and to their metric variations, using generalized cones as the stimuli (a set of simple volumes believed by some to be the basic elements of 3D object perception; e.g. Marr 1982; Biederman 1987). They found that some IT cells showed selectivity for specific non-accidental feature "dimensions" such as aspect ratio (overall width and height), curvature, taper (convergence of contours) and convexity, while exhibiting broad tunings for metric variations within the dimensions to which they were sensitive.

5.4 Brief Shape Aftereffects (SAEs) as a "Psycho-anatomical" Tool for Probing High-level Central-tendency Codings of Global Form Features

While neurophysiological studies identify potential central-tendency coded features in terms of systematic neural tunings, complementary behavioral studies are necessary to show that perception of those features is indeed determined by central-tendency coding. Activation-based sensitivity reduction, known as adaptation, is a ubiquitous property of cortical neurons (e.g. for V1 cells, Vautin & Berkeley 1977; Albrecht *et al.* 1984; Hammond *et al.* 1989; Saul & Cynader 1989; Sclar *et al.* 1989; Bonds 1991; Carandini *et al.* 1998; Muller *et al.* 1999; for IT cells, Oram & Perrett 1992; Miller *et al.* 1993*b*; Lueschow *et al.* 1994; Vogels & Orban 1994; Vogels *et al.* 1995). Consequently, a psychophysical signature of central-tendency coding is a "repulsive aftereffect." Roughly speaking, prior exposure (i.e. adaptation) to a feature value (e.g. a left-tilted grating) that is deviated from the test feature value (e.g. a vertical grating) makes the test feature appear distorted in the opposite direction (e.g. it appears right tilted), due to reduction in the relative activity of cells tuned to the adapted value (see Fig. 5.1).

Demonstrations of repulsive aftereffects for local orientation and spatial frequency have confirmed central-tendency codings of local orientation and spatial

frequency in V1 (see Fig. 5.1 and Section 5.1). Similar demonstrations of repulsive aftereffects for global geometric features would indicate the existence of central-tendency codings for specific global features. Because neurophysiological results are currently ambiguous as to what global features are central-tendency coded in high-level visual areas, psychophysical demonstrations of global shape aftereffects should be especially helpful in interpreting and potentially guiding neurophysiological research on global shape coding. However, in comparison with the extensive prior research on local orientation and spatial frequency aftereffects, relatively few studies investigated global shape aftereffects (e.g. Regan & Hamstra 1992) prior to Suzuki and colleagues' demonstrations of brief shape aftereffects and recent demonstrations of face-related aftereffects (see Section 5.6). This was primarily due to the difficulty in dissociating global shape aftereffects, indicative of high-level coding of global forms, from local feature aftereffects (the orientation and spatial-frequency aftereffects discussed above, as well as contour-repulsion based figural aftereffects; e.g. Kohler & Wallach 1944; Sagara & Ohyama 1957), indicative of low-level coding of local image features. Isolating high-level aftereffects is difficult because complex shapes are made up of locally oriented contours which necessarily activate low-level processes.

In contrast with conventional methods of measuring aftereffects using prolonged stimuli (e.g. seconds to minutes of adaptation), Suzuki and colleagues noted that rapid sequences of stimuli, that is, (1) brief adaptation and (2) brief test, combined with (3) brief adapt-to-test ISI (inter stimulus interval), could produce robust repulsive aftereffects for global geometric features (see Fig. 5.2).

First, brief adaptors (less than ~150 ms) were found to be effective in producing high-level aftereffects with minimal influences from low-level aftereffects (details follow). This is consistent with the temporal properties of high-level and low-level cells. Inferotemporal (IT) cells in alert monkeys typically adapt to briefly presented stimuli; following a brief exposure to an adaptor stimulus (350–500 ms tested), IT cells reduce their sensitivity to a test stimulus (presented following 300–2000 ms of ISI), especially when the test stimulus is identical or similar to the adaptor stimulus (e.g. Miller *et al.* 1993*b*; Lueschow *et al.* 1994; Vogels & Orban 1994; Vogels *et al.* 1995). In contrast, most V1 cells in alert monkeys do not adapt to briefly presented (up to 500 ms) gratings of preferred orientations (Gur *et al.* 2004; interestingly, V1 cells in anesthetized monkeys adapt rapidly, Muller *et al.* 1999). Furthermore, both behavioral and physiological results indicate that V1 cells adapt slowly over many seconds (e.g. Albrecht *et al.* 1984). Circumventing low-level adaptation by using brief adaptors was particularly important for Suzuki and colleagues to demonstrate aftereffects on relatively simple geometric features consisting of oriented contours. Brief adaptation, however, is not an absolute requirement for generating a high-level aftereffect as evidenced by the fact that face-related aftereffects can be induced by adaptation durations ranging from a second to several minutes (see Rhodes *et al.* this volume).

Second, briefly presenting a test shape (with backward masking) is generally advantageous because it reduces adaptation effects initiated by the test shape,

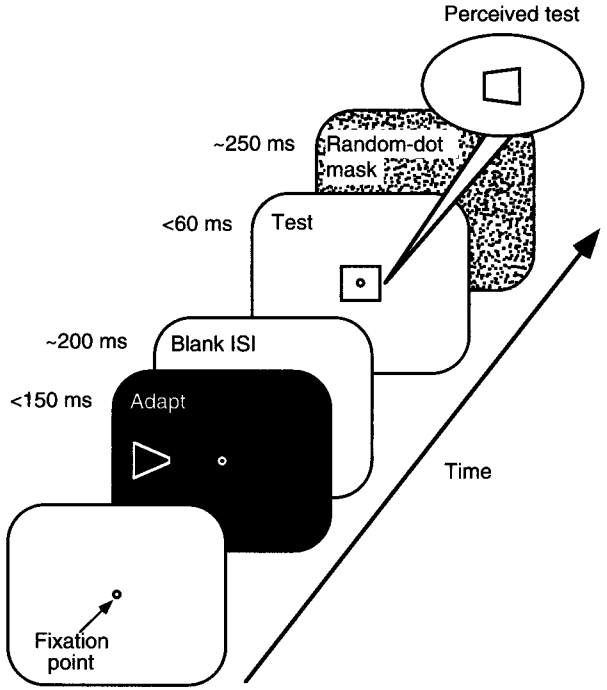

FIG. 5.2 A typical trial sequence used to demonstrate a global shape aftereffect (SAE); a taper aftereffect is shown as an example. The adaptor and test shapes are presented briefly (with an optimum adapt-to-test ISI of ~200 ms) in order to reveal aftereffects due to global-form adaptation unconfounded by low-level local aftereffects (Suzuki 2001). The adaptor display was often presented in reverse contrast (as shown in the figure) to prevent (or substantially reduce) perception of adapt-to-test apparent motion, though apparent motion does not contribute to SAEs (Suzuki & Cavanagh 1998).

and thus increases the sensitivity for detecting aftereffects induced by the adaptor (e.g. Wolfe 1984; Suzuki & Cavanagh 1997). Aftereffects tend to be strengthened when the test stimulus is weakened by reducing its duration, contrast, or presenting it away from the fovea (Suzuki & Cavanagh 1998). Third, in some cases the use of brief adapt-to-test ISI of about 200 ms was crucial to produce aftereffects that were spatially tolerant, that is, to produce aftereffects that diminished only gradually as the position and/or size of the adaptor shape increasingly deviated from the test shape (e.g. Rivest *et al.* 1998; Suzuki & Cavanagh 1998; Suzuki 2001). Because the sizes of neural receptive fields increase from averaging only 0.3–1° in V1, ~3° in V2, and 1.5–4° in V4 (in parafovea at retinal eccentricities of 2–5° from the fovea), to encompassing large portions of the visual field (a mean of 30° but up to 100°) in IT (e.g. Gross *et al.* 1972; Desimone & Gross 1979; Dow *et al.* 1981; Foster *et al.* 1985;

Desimone & Schein 1987; Gattas *et al.* 1988; Ito *et al.* 1995), aftereffects that show substantial spatial tolerance are likely to be mediated by central-tendency codings in high-level cortical visual areas. Suzuki and colleagues thus called these spatially tolerant aftereffects induced by rapid stimulus sequences, *shape aftereffects* (*SAEs*). The specific types of SAEs found so far suggest that the initial central-tendency coding of local orientation in V1 is elaborated in higher visual areas as central-tendency codings of global skew, taper, curvature, aspect ratio, and convexity (e.g. Suzuki & Cavanagh 1998; Suzuki 1999, 2001, 2003*a*).

For example, a skewed adaptor (e.g. a right-skewed parallelogram) induces an opposite skew on a subsequently flashed symmetric (e.g. rectangular) test shape (Fig. 5.3(A), top). Skew aftereffects are translation tolerant (i.e. occurring when the adaptor and test shapes are presented at completely non-overlapping locations), and they occur even when the adaptor and test shapes have no matching contours (Fig. 5.3(A), bottom), suggesting that skew is central-tendency coded as a global attribute beyond early coding of local orientation. Similarly, a tapered adaptor (e.g. a right pointing triangle) induces an opposite taper (Fig. 5.3(B); also see Fig. 5.2), a curved adaptor induces an opposite curvature (Fig. 5.3(C)), and an elongated adaptor (e.g. a tall ellipse) induces an opposite aspect ratio (Fig. 5.3(D)), on subsequently flashed symmetric test patterns. These aftereffects are translation and scale tolerant (up to 12° of adapt-test separation and 0.3–1.8 in adapt-test linear size ratio; Rivest *et al.* 1998; Suzuki & Cavanagh 1998; Suzuki 2001), suggesting that taper, curvature, and aspect ratio are also central-tendency coded as global attributes. Finally, a convex/concave adaptor (e.g. a diamond/hourglass shape) induces a concave/convex distortion on a subsequently flashed symmetric test pattern (Fig. 5.3(E), top/middle). These convexity aftereffects are scale tolerant (Fig. 5.3(E), bottom), suggesting that convexity is also central-tendency coded as a global attribute. Interestingly, unlike skew, taper, and aspect ratio aftereffects, convexity aftereffects are not translation tolerant, potentially reflecting the fact that the convex–concave distinction for a contour depends on the position of the contour relative to the region of interest. These aftereffects were typically measured using a method of adjustment (having the observer reproduce the appearance of the distorted test shape) or a method of cancellation (nulling the aftereffect using a staircase procedure).

To confirm that a rapid sequence of visual stimulation (Fig. 5.2) uniquely adapted high-level processing without affecting low-level processing, convexity aftereffects induced by convex and concave shapes (Fig. 5.3(E)) were compared with conventional tilt aftereffects induced by oriented gratings (which adapt local-orientation-tuned cells in V1) under closely matched conditions. With prolonged (2600 ms) adaptation, the four patches of oriented gratings locally produced narrowly-tuned tilt aftereffects (Fig. 5.4(A)), which were maximum when the adaptor orientation was ±15° from the test orientation, consistent with the narrow orientation tunings of V1 cells. With brief (134 ms or less) adaptation, however, these gratings produced

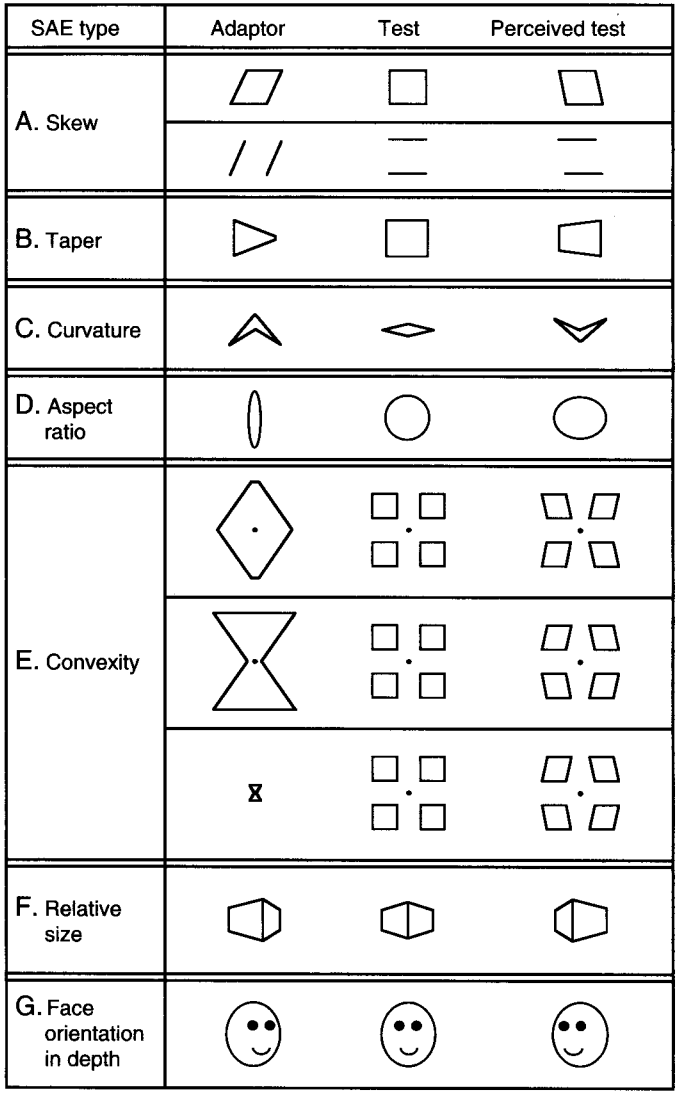

FIG. 5.3 Schematic examples of SAEs demonstrated so far. SAEs are generally translation and scale tolerant. However, in some cases, the magnitude of SAEs can be modulated by the relative position of the adaptor and test shapes (Suzuki & Cavanagh 1998). An array of squares rather than a solid shape was used as the test pattern for measuring convexity aftereffects (E), partly to demonstrate aftereffects acting globally on a grouped array, and partly to facilitate comparisons with local tilt aftereffects (Suzuki 2001).

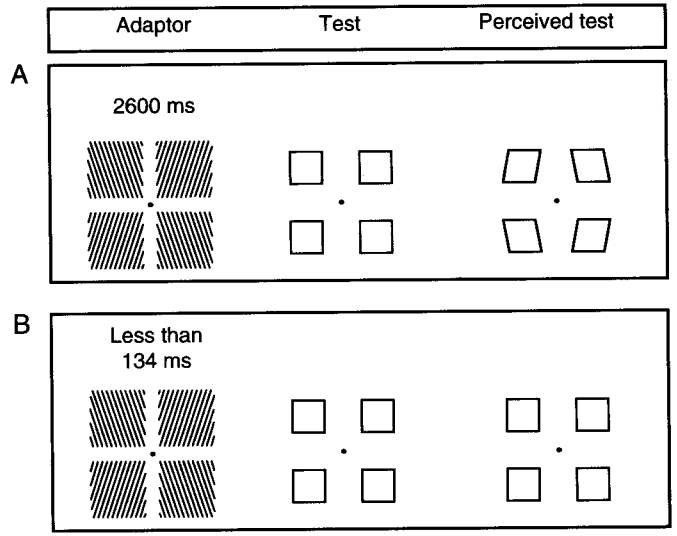

FIG. 5.4 An example of a grating adaptor used in Suzuki (2001) to demonstrate that brief adaptation dissociates global convexity aftereffects from local tilt aftereffects. (A) With prolonged adaptation (2600 ms), the four grating patches locally produce tilt aftereffects, resulting in the overall convex appearance of the four test squares, similar to aftereffects induced by a concave hourglass (see Fig. 3(E), middle). (B) With brief adaptation (less than 134 ms), however, the grating patches produce no aftereffect.

NO aftereffects (Fig. 5.4(B)). In contrast, the convex and concave shapes (Fig. 5.3(E)) produced strong aftereffects with adaptation as brief as 27 ms (undiminished in magnitude relative to 2600 ms adaptation). Control experiments have shown that this dissociation cannot be attributed to differences in spatial frequency contents, or to the possibility that both the convex and concave shapes and the oriented gratings adapted a common orientation-coding mechanism except that briefly presented grating stimuli somehow sub-optimally activated this mechanism (Suzuki 2001).

Parametric investigations have further demonstrated that brief aftereffects induced by the convex and concave shapes were (1) broadly tuned to convexity (maximum when orientations of the side contours were ±30 to ±60° from vertical), (2) indifferent as to the mode of contour definition (e.g. bright lines, high-pass-filtered lines, contours generated by the spatial inhomogeneity of visual sensitivity), (3) already maximum at low contrast energy (saturation within a 3-fold increase in contrast energy beyond detection), (4) strongly modulated by selective attention (see Section 5.5), and (5) relatively size tolerant. As suggested by Suzuki (2001, 2003*a*), these characteristics of brief convexity aftereffects are paralleled (sometimes even quantitatively) by the response characteristics of high-level pattern-selective neurons, especially in IT. Compared to their low-level counterparts, responses of high-level visual neurons tend to have broader orientation tuning (e.g. Desimone & Schein 1987;

Vogels & Orban 1994; Geisler & Albrecht 1997; McAdams & Maunsell 1999), be relatively indifferent to mode of contour definition (e.g. Sato *et al.* 1980; Rolls & Baylis 1986; Komatsu *et al.* 1992; Ito *et al.* 1994), saturate at lower contrast energy (e.g. Gross *et al.* 1972; Sato *et al.* 1980; Rolls & Baylis 1986; Sclar *et al.* 1990; Cheng *et al.* 1994), be strongly affected by selective attention (when the competing stimuli are presented within the receptive field of a cell; e.g. Luck *et al.* 1997; Chelazzi *et al.* 1998), and be scale tolerant (e.g. Ito *et al.* 1995; Tanaka 1996; Hikosaka 1999). Converging evidence thus suggests that SAEs are indeed mediated by central-tendency codings of global form features, perhaps in IT.

5.5 Attentional Modulations of SAEs

Global pattern processing necessarily pools information from lower-level processes to achieve both complex pattern selectivity and invariance for position and scale (see Riesenhuber & Poggio 2000, for a review). Because of the extensive convergence of neural signals, even small neural modulations that occur in lower-level processing due to attention, contextual factors, and short-term and long-term neural plasticity, can be substantially amplified in high-level processing. For example, though conventional tilt aftereffects (mediated by orientation-tuned cells in V1) are only weakly modulated by attention (Spivey & Spirn 2000), SAEs are strongly modulated by attention (see below). Furthermore, because cells that are tuned to global forms (with position and size tolerance) necessarily have large receptive fields, high-level coding has poor spatial resolution; for example, an IT cell will respond poorly to its preferred shape if other shapes are also present in the vicinity unless the preferred shape is attended (e.g. Miller *et al.* 1993*a*; Chelazzi *et al.* 1998). Thus, as people live in environments crowded with multiple objects, global form processing must be tightly coupled with attentional selection processes.

Because SAEs selectively probe adaptation of global-form coding mechanisms, they provide a unique means to study how attention modulates incoming signals to influence activations at the level of global-form processing. In neurophysiological research, the degree of attentional selection is often assessed in terms of linear attention weights, that is, how much the input signals from the attended contours are amplified relative to the input signals from the ignored contours in activating a given cell; this way of conceptualizing attention effects at the single-cell level is termed the biased competition model (e.g. Moran & Desimone 1985; Desimone & Duncan 1995; Luck *et al.* 1997; Reynolds *et al.* 1999). Comparable linear attention weights can be behaviorally derived from attentional modulations of SAEs (if one assumes that the magnitude of an SAE is monotonically related to activation of the relevant global-form coding neurons) using a simple formula,

$$w_{\text{attended}} - w_{\text{ignored}} = \frac{[\text{SAE from XY with X attended}] - [\text{SAE from XY with Y attended}]}{[\text{SAE from X alone}] - [\text{SAE from Y alone}]} \times 100\%,$$

where X and Y represent two adaptor shapes, XY represents an adaptor consisting of overlapping X and Y shapes, and w_{attended} and w_{ignored} represent linear attention weights that multiply the input signal when a shape is attended and ignored, respectively; note that although these attention weights are different for the two shapes (e.g. due to differences in their salience), the difference between the attended and ignored weights ($w_{\text{attended}} - w_{\text{ignored}}$) is the same for each shape under the assumptions of the biased competition model (see Suzuki 2003*a* for details).

For example, when overlapped hourglass and diamond shapes were used as the adaptor (Suzuki 2001, 2003*a*), attending to the (convex) diamond produced a concave aftereffect (Fig. 5.5(A)), whereas attending to the (concave) hourglass produced a convex aftereffect (Fig. 5.5(B)); the concave and convex directions of aftereffects were given opposite signs in the equation. The differential attention weights were obtained by expressing this attention-based difference in SAE from the overlapped adaptor as the percentage of the maximum possible difference in SAE from the diamond and hourglass adaptors presented alone. If attention completely suppressed the ignored shape, attending to the diamond or the hourglass in the overlapping adaptor should produce aftereffects identical to those obtained from the diamond or the hourglass adaptor presented alone. This would result in the numerator and denominator of the equation being equal, making the ratio unity (or 100 per cent). Thus, differential attention weights of 100 per cent would indicate complete attentional selection. Psychophysically inferred differential attention weights can be compared with those obtained in neurophysiological studies. Such a comparison is difficult with conventional behavioral measures of selective attention like modulations of response time and accuracy in probe detection tasks.

When different combinations of overlapped shapes were used, the differential attention weights obtained from SAEs were similar. Notably, attentional selection was similar for a combination of opposing shapes (e.g. an hourglass and a diamond that produce opposite convexity aftereffects) and a combination of non-opposing shapes (e.g. an hourglass that produces a convex aftereffect and horizontal lines that produce a vertical-elongation [aspect-ratio] aftereffect). This result ruled out the possibility that attentional selection in SAEs is primarily due to lateral-inhibition-type modulations of population activity within the specific central-tendency coding of the attended feature (e.g. suppression of convex-tuned units when a concave shape is attended). Note, however, that this type of within-coding interaction might still play a role in influencing the course of perceptual multi-stability (see Section 5.8). Interestingly, the differential attention weights derived from the SAE results were ~60 per cent under various conditions (Suzuki 2001, 2003*a*), comparable to the differential weights achieved at the level of V4 neurons (Reynolds *et al.* 1999). This might suggest that attentional modulations of SAEs are primarily due to modulations of contour processing up to V4.

Because SAEs probe high-level global processing and are sensitive to attentional modulations, SAEs also provide a useful paradigm for studying how attentional selection interacts with various image factors such as salience, scale,

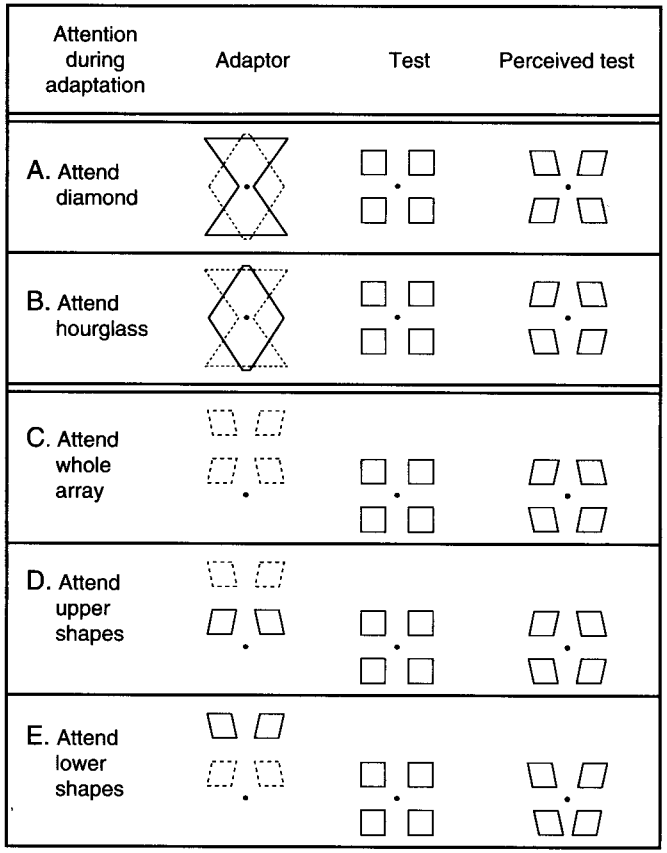

FIG. 5.5 Schematic examples of attentional modulations of SAEs. The attended portion of the adaptor is indicated by dashed contours (for illustration purposes only). When overlapped shapes were used as the adaptor (A and B), the two shapes were given different colors to facilitate attentional selection.

and grouping. For example, it would be beneficial if attention could differentially weight signals regardless of their luminance contrasts; this way, signals from dimmer, but potentially important images, can be selected for higher-level processing as well as signals from salient images. Physiologically, Reynolds and Desimone (2003) have demonstrated that attentional mechanisms can substantially boost the signals from attended low-contrast stimuli against ignored high-contrast stimuli, in controlling responses of V4 neurons. By measuring attentional modulations of SAEs while varying the relative contrasts of the overlapped adaptor shapes, Suzuki (2001) provided psychophysical evidence that the degree of attentional weighting was relatively independent of the salience of the stimuli. In addition, when the overall scale of the overlapped adaptors was varied from 2.3° to 15.8°, the differential attention weights

remained largely constant at ~60 per cent, suggesting that the efficiency of attentional selection is relatively scale invariant (Suzuki 2001).

In another SAE study, effects of attending to the whole and parts were assessed by using an adaptor consisting of an array of shapes (Suzuki 1999). For example, a concave array of four parallelograms as the adaptor was flashed in the upper visual field (Fig. 5.5(C–E)). Because the test array was flashed around the fixation point, the bottom row of the adaptor array and the top row of the test array were presented at the same location. Thus, local orientation aftereffects should make the top row of the test array appear tilted outward. To the contrary, when the whole concave adaptor array was attended (attended contours are indicated by dashed lines in Fig. 5.5), the top row of the test array appeared tilted inward and the test array as a whole appeared convex, consistent with a global convexity aftereffect (Fig. 5.5(C)). The result was similar when the top row of the adaptor array was attended (Fig. 5.5(D)), suggesting that when peripheral parts are attended, it is difficult to exclude parts that are closer to the fixation. Interestingly, attending to the bottom row of the adaptor array made the whole test array appear tapered downward (Fig. 5.5(E)), suggesting that attention to parts closer to the fixation can successfully exclude more peripheral parts, and that the selected parts of the adaptor array were processed as having an overall upward taper and the whole test array was processed as a unit. Thus, central-tendency codings of convexity and taper seem to be activated on the basis of the attended portion of an array.

It will be important in future research to gain an integrative understand of attention effects on global-form processing (e.g. studied with SAEs) and those on lower-level processing of luminance, contrast, and local features (e.g. edge orientation). For example, attention results in increased contrast gain and response synchrony in lower-level processes, which might contribute to the increased competitive edge for the attended signals in higher-level processes (Reynolds *et al.* 2000; Fries *et al.* 2001; Cameron *et al.* 2002; Reynolds & Desimone 2003). Future research must specify how signal weighting and synchrony are controlled by feedforward, feedback, and horizontal neural interactions at different stages of processing. Attention effects might also be different for cells with different response properties. For example, while examining attentional modulations of negative afterimages, Suzuki and Grabowecky (2003*b*) found that attentional modulation of brightness adaptation primarily acted on cells that are polarity-independent (responding to patterns whether they are dark or light) rather than cells that are polarity-dependent (responding selectively to dark or light).

5.6 What Global Form Features are Systematically Central-tendency Coded Beyond the Coding of Local Orientation in V1?

The types of SAEs demonstrated so far suggest that the initial central-tendency coding of local orientation in V1 is elaborated into central-tendency codings of

global configurations of orientated contours, including skew, taper, curvature, aspect ratio, and convexity (Fig. 5.3(A–E)), perhaps in IT. Others have reported position and scale tolerant aftereffects for various geometric distortions of faces, suggesting high-level central-tendency codings of face-related features. For example, Leopold *et al.* (2001) demonstrated that face-identity aftereffects were undiminished across adapt-to-test retinal separations (implemented as fixation shifts) of up to about half the width of the face (or 6°). Zhao and Chubb (2001) demonstrated that face-distortion aftereffects (originally reported by Webster & MacLin 1999) were still substantial (though reduced in magnitude by half) for adapt-to-test size changes by a factor of two (also see chapters in this volume by Leopold & Bondar, and Rhodes *et al.*). An important question is whether these psychophysically inferred central-tendency coded features represent a fundamental set of high-level image codings, or are merely the tip of the iceberg, and other central-tendency coded features will be uncovered so long as appropriate stimuli are used. Though a definitive answer awaits future research, there are several reasons to believe that these features may be principal.

First, despite the fact that systematic geometric tunings have rarely been reported in IT, there is apparent consistency between the few feature variations for which systematic neural tunings have been found in IT on the one hand, and the set of feature variations for which repulsive aftereffects have been found on the other. For example, neural tunings in IT as well as repulsive aftereffects have been reported for geometric variations in face-related features. In addition, the class of IT cells systematically tuned to variations in aspect ratio (overall height and width), curvature, taper (convergence of contours), and convexity could mediate the SAEs found for aspect ratio, curvature, taper, and convexity (see Sections 5.3 and 5.7).

Second, because high-resolution discriminations of face attributes are crucial for human survival, it is not surprising that face-related features are central-tendency coded; as discussed earlier, central-tendency coding based on overlapped tuning functions allows interpolations between discretely sampled feature values (see Section 5.2 and Fig. 5.1). More generally, if central-tendency coding is an adaptive strategy of the visual system to bestow fast global analyses (rather than part-based analyses) with enhanced perceptual resolution to processing of behaviorally important patterns, it would be beneficial for the visual system to have mechanisms to develop central-tendency codings for behaviorally relevant geometric variations through experience. Brain imaging research using fMRI suggests that coding of objects in which one is an expert (e.g. birds, cars, and artificial creatures such as "greebles") and coding of faces (in which almost everybody is an expert) share the same focal brain region in the fusiform area (e.g. Gauthier *et al.* 1999, 2000; Gauthier 2000). Consistent with this shared-resource idea, there is some evidence that acquired expertise in a non-face category (e.g. cars) can interfere with face processing (e.g. Gauthier *et al.* 2003). Recent fMRI results also suggest that different classes of objects (e.g. faces and other objects of expertise) are processed by relatively non-overlapping sub-populations of cells within the fusiform area (e.g. Rhodes *et al.* 2004).

Single-cell recording studies with monkeys have shown that, with practice, IT cells are capable of developing selectivity for novel complex patterns, such as for paper-clip-like objects made up of wires bent multiple times in different directions (Logothetis *et al.* 1995), schematic fish-like drawings with variations in fin and mouth shapes, schematic faces with variations in eye, nose, and mouth positions (Sigala & Logothetis 2002), a set of complex geometric shapes (Kobatake *et al.* 1998), and for a set of human faces (Young & Yamane 1992). Overall, converging evidence suggests that human fusiform area and monkey IT have the plasticity to develop neural tunings for complex geometric features (including face-related features) that are important for discriminating individuals within a trained class of objects (especially for objects that are structurally similar).

It is not clear, however, whether acquired neural selectivities result in central-tendency codings of the trained features. Indeed, well-practiced discriminations among geometrically distinct objects, not lying along specific geometric continua (e.g. a face versus a chair versus a trumpet), are likely to be based on differences in the general topography of distributed neural activity. Future research should thus examine SAEs in conjunction with pattern learning. If central-tendency codings develop with experience, experts should show repulsive aftereffects for the feature variations in which they are experts. For example, bird experts might show repulsive aftereffects for geometric variations that are crucial for their discrimination of birds. Strong support for this hypothesis would be obtained if a repulsive aftereffect was demonstrated for a trained feature variation for which no aftereffect was observed prior to the training. Such a result would confirm that central-tendency codings generally underlie perception of well-experienced, thus, behaviorally significant, geometric variations.

If central-tendency codings indeed develop with experience, it might be no accident that SAEs have been found for overall skew, taper, curvature, aspect ratio, and convexity. This is because judging 3D orientations of objects is a ubiquitous aspect of daily life, and these features represent a basic set of geometric variations that people routinely experience when flat or moderately curved surfaces are manipulated in depth (most 3D objects can be approximated as being composed of flat or moderately curved surfaces). For example, relative to the fronto-parallel view, a rotation about the vertical axis induces overall tall aspect ratio and laterally pointing taper, both of which increase with greater rotations away from the fronto-parallel view (e.g. from left to right in Fig. 5.6(A)). When an angled intersection of a pair of flat surfaces is rotated about the axis perpendicular to the intersection, an overall curvature is introduced as the surfaces are rotated away from the fronto-parallel view (Fig. 5.6(B)). When the bending angle is changed (while approximately in the fronto-parallel view), the degree of convexity varies. The degree of skew systematically varies as a slanted surface is moved about in space (Fig. 5.6(C)). These 3D-rotation-concomitant geometric distortions occur to flat or moderately curved surfaces of any shape, though the distortions are most evident for symmetric surfaces such as the rectangular surfaces shown in Fig. 5.6.

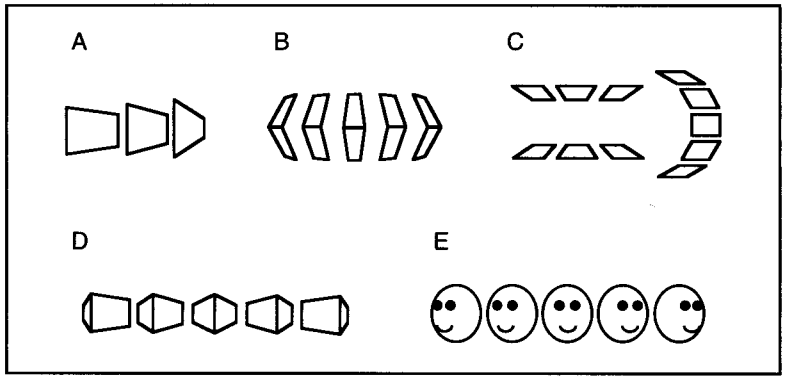

FIG. 5.6 Some geometric features that are systematically distorted when surfaces are rotated in depth. See text for details.

It is thus possible that central-tendency codings of global skew, taper, curvature, aspect ratio, and convexity might be an integral part of the mechanisms that mediate perception of 3D orientations of objects, and small degrees of viewpoint normalization (generating rotated representations to align to a reference viewpoint); for example, approximate 3D rotated views can be generated by applying these 2D geometric distortions via systematically shifting the neural population activities that encode these features. A mechanism like this might mediate 2D-based viewpoint interpolations which are hypothesized to facilitate recognition of objects across different viewpoints (e.g. Ullman 1998). If this is the case, then, repulsive aftereffects should be observed for other image features that also vary systematically as surfaces are rotated in depth. Relative size is an example. As an angled intersection of a pair of flat surfaces is rotated about the axis parallel to the intersection, the visible portion of one surface becomes larger and that of the other surface becomes smaller (Fig. 5.6(D)). Our preliminary results have demonstrated a (translation tolerant) repulsive aftereffect for this feature variation; for example, a flashed polygon showing more of the left surface makes a subsequently flashed symmetric polygon appear to show more of the right surface (see Fig. 5.3(F)).

Our preliminary results have also demonstrated a (translation tolerant) repulsive aftereffect for variations in 3D head orientation, using highly schematic faces (Fig. 5.6(E)); for example, a flashed face looking to the right makes a subsequently flashed forward-looking face appear to be looking to the left (see Fig. 5.3(G)). It is interesting to note that some IT cells show tunings for variations in 3D head orientation (see Section 5.3). Because knowing where a person is looking is of great behavioral significance (and perceived direction of gaze depends in part on perceived head orientation), it may be important to have a dedicated central-tendency coding for head orientation. In fact, Fang and He (2004) have recently reported seemingly object-specific repulsive aftereffects for head orientation as well as for 3D orientations of cars and stick figures, suggesting that the visual

system has specialized central-tendency codings of 3D orientation for at least some classes of objects. However, it would not be efficient to have dedicated central-tendency codings for 3D orientation of arbitrary objects. Because objects are composed of surfaces (often large, approximately flat and symmetric surfaces especially for artifacts), the visual system might have instead opted for central-tendency codings of basic geometric distortions that occur when surfaces are rotated and translated in 3D environments.

5.7 Opponent Coding as a Special Case of Central-tendency Coding

It is noteworthy that all of the potentially central-tendency coded global geometric variations discussed so far have two extreme ends and often include a "category boundary" in the middle where perception changes discontinuously (categorically) and where high-resolution discrimination is behaviorally desirable. For example, it is important to be able to quickly determine whether a person has a positive or negative expression (rather than how positive or how negative), whether a person is female or male (rather than how feminine or how masculine), or whether a person is looking to your left or to your right (rather than how far left or how far right). Even for the potential central-tendency codings of skew, taper, curvature, aspect ratio, and convexity, geometrically continuous variations in these features go through a categorical discontinuity at the point of symmetry. Skew, taper, and curvature switch from one direction to the opposite direction when they are varied across the respective null points of symmetry. As aspect ratio is changed continuously through the null point, it changes from elongation along one orientation (e.g. flat) to elongation along the orthogonal orientation (e.g. tall). Convexity switches from convex to concave or vice versa across the null point. Each null point, coinciding with the category boundary, can be considered "the norm" because it lies in the middle of a geometric continuum and represents an "average" of the two extremes.

These considerations suggest that underlying population codings of global geometric features may be based primarily on the relative activities of two opponent populations of cells, one broadly tuned to one side of the norm (e.g. tuned to a flat aspect ratio) and the other broadly tuned to the other side of the norm (e.g. tuned to a tall aspect ratio); the perceived feature value (e.g. the perceived aspect ratio) is determined by the relative activation of the two opponent pools of cells. This is a special case of central-tendency coding, where central tendency is computed on the basis of only two pools of cells broadly tuned to the extreme feature values (Fig. 5.7), rather than on the basis of multiple pools of cells tuned to intermediate feature values (Fig. 5.1). Regan and Hamstra (1992) originally proposed this *opponent-coding* model with respect to coding of aspect ratio. As illustrated in Fig. 5.7(A), the two opponent tuning functions (one broadly tuned to flat aspect ratio and the other broadly tuned to tall aspect ratio) intersect at the norm (neutral aspect ratio) where the slopes are steepest,

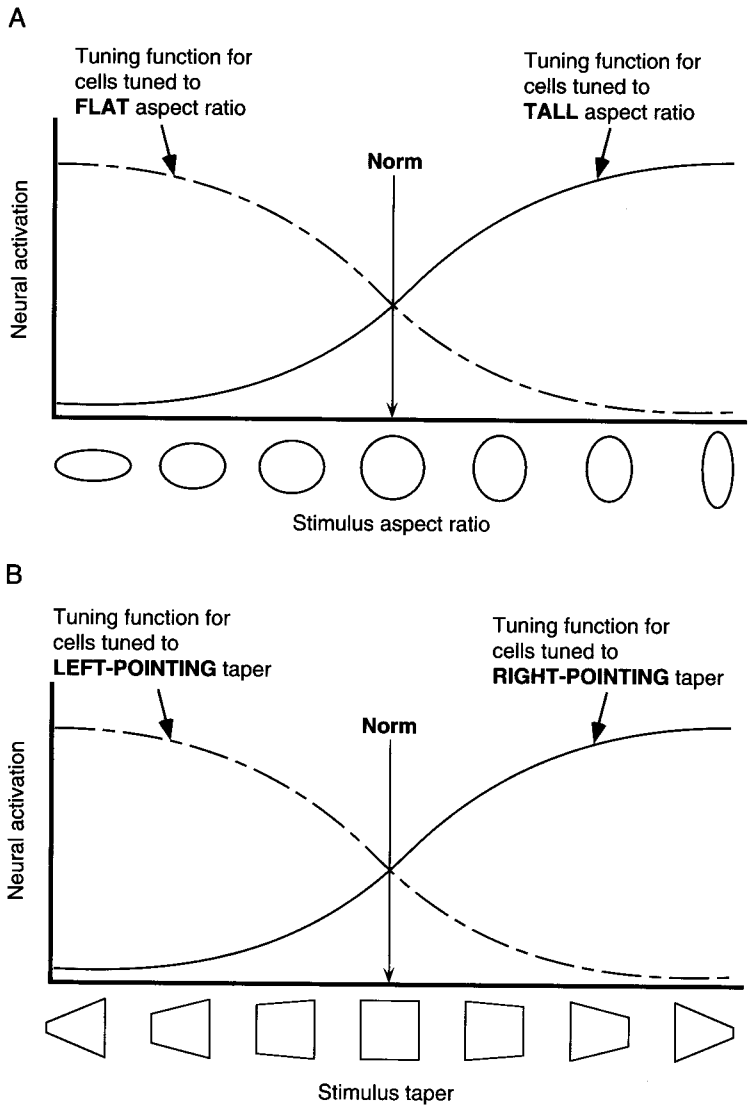

FIG. 5.7 Opponent coding based on two opponent pools of cells broadly tuned to the extreme values of a feature dimension. Two examples are illustrated: (A) opponent coding of aspect ratio (after Regan & Hamstra 1992), and (B) opponent coding of taper. The vertical lines indicate the norms (i.e. neutral aspect ratio, and neutral taper).

such that the relative activation of the opponent-tuned populations changes maximally when the feature value (aspect ratio) is varied near the norm (neutral aspect ratio). This is consistent with the fact that the thresholds for aspect-ratio discriminations are lowest at the neutral aspect ratio (Regan & Hamstra 1992). Regan and Hamstra's opponent-coding model can be applied to coding

of other global geometric features; an example is shown in Fig. 5.7(B) for the opponent coding of taper. An opponent coding would be an efficient version of a central-tendency coding especially when the primary perceptual demand is in detecting bi-directional categorical deviations from the norm, with high sensitivity primarily required at the category boundary.

An opponent coding (a central-tendency coding based on two opponent pools of cells) can be distinguished from a *multi-channel coding* (a central-tendency coding based on multiple pools of cells tuned to intermediate feature values) on the basis of the tuning properties of aftereffects. Multi-channel coding predicts repulsive aftereffects with respect to the *relative* feature values of the adaptor and test stimuli. For example, tilt aftereffects are thought to result from the multi-channel coding of local orientation by orientation-tuned cells in V1. As illustrated in Fig. 5.1, adaptation-based desensitization of orientation-tuned cells makes the test orientation appear tilted away from the adaptor orientation. Because there is nothing special about the vertical orientation in the multi-channel coding scheme illustrated in Fig. 5.1, the same repulsive aftereffect is predicted (and occurs) with respect to any test orientation. For example, if the test orientation was 30°, it would appear less than 30° if the adaptor orientation was greater than 30°, and it would appear greater than 30° if the adaptor orientation was less than 30°. Thus, a signature characteristic of multi-channel coding is that when the adaptor feature value (e.g. adaptor orientation) is varied while the test feature value (e.g. test orientation) is fixed, the relevant repulsive aftereffect (e.g. tilt aftereffect) should exhibit an anti-symmetric tuning curve centered at the test feature value (e.g. test orientation).

For example, several tuning curves of tilt aftereffects for different test orientations are illustrated in Fig. 5.8(A) (schematized from Mitchell & Muir 1976 and O'Toole & Wenderoth 1977). Suppose the test orientation is + 45° (the top tuning curve in Fig. 5.8(A)). If the adaptor orientation is identical to the test orientation (+45°), no aftereffect will occur and the test orientation will appear veridical at +45°; this is the node of the tuning curve. If the adaptor orientation is made slightly less than the test orientation (< +45°), the aftereffect will increase the perceived tilt of the test grating, causing the tuning curve to rise above +45°. In contrast, if the adaptor orientation is made slightly greater than the test orientation (> +45°), the aftereffect will decrease the perceived tilt of the test grating, causing the tuning curve to fall below +45°. Because the local orientation tunings of V1 cells are narrow (median orientation bandwidth = 30–40°; Geisler & Albrecht, 1997), these repulsive aftereffects will diminish as the adaptor orientation deviates further from the test orientation, causing the tuning curve to re-converge to +45°; in fact, the tuning curve typically peaks when the adaptor orientation is about ±15° from the test orientation and re-converges when the adaptor orientation is about ±40–60° from the test orientation (e.g. Mitchell & Muir 1976; O'Toole & Wenderoth 1977; Magnussen & Kurtenbach 1980; O'Shea *et al.* 1993).

Under some conditions, this repulsive tilt aftereffect flips to an attractive effect when the adaptor orientation deviates as far as about 50–90° from the test

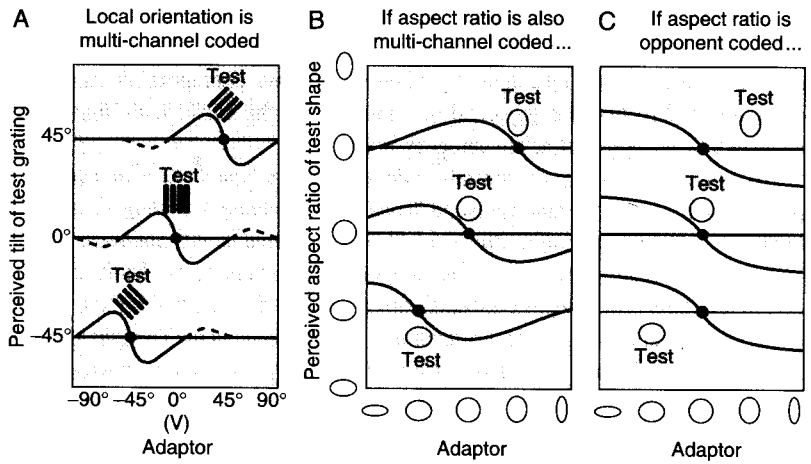

FIG. 5.8 Tuning curves of aftereffects expected from multi-channel coding (A and B) and from opponent coding (C) when values of the adaptor feature (orientation or aspect ratio) are varied for fixed values of the test feature. (A) Multi-channel coding of local orientation generates tuning curves of tilt aftereffect that are anti-symmetric about the test orientation, with the nodes (indicated by solid circles) occurring at the test orientations. The dashed portions indicate indirect tilt aftereffects. (B) If aspect ratio is also multi-channel coded, aspect-ratio aftereffects should produce tuning curves that are also anti-symmetric about the test aspect ratios (likely with broader tuning based on previous results). (C) Alternatively, if aspect ratio is opponent coded (see Fig. 5.7), aspect-ratio aftereffects should be anti-symmetric about the neutral aspect ratio regardless of the test aspect ratio.

orientation (the dashed portions of the tuning curves shown in Fig. 5.8(A)). This attractive effect is known as the *indirect* tilt aftereffect. It is typically smaller in magnitude (e.g. Gibson & Radner 1937; Campbell & Maffei 1971; O'Shea *et al.* 1993) and less reliable (e.g. Mitchell & Muir 1976) than the repulsive component of tilt aftereffects. Interestingly, the attractive component was abolished (1) when a square frame was placed around the test grating (e.g. Wenderoth & van der Zwan 1989), and (2) when an adaptor consisted of multiple gratings with no average tilt (Held & Shattuck 1971; Suzuki 2001; in these studies, the spatially averaged tilt across the multiple local gratings was the same as the test orientation [vertical]; see Fig. 5.4 as an example). Such sensitivity to global context (external frame and long-range spatial averaging) suggests that the attractive component of tilt aftereffects is mediated by a high-level mechanism involving cells with relatively large receptive fields. In contrast, the repulsive component of tilt aftereffects is largely unaffected by these manipulations of global context, consistent with its mediation by low-level orientation-tuned cells with small receptive fields (see Wenderoth & van der Zwan 1989 for a review of spatial and temporal dissociations between the repulsive and attractive components of tilt aftereffects). Clifford (in this volume) proposes a dual-process

model of orientation adaptation, in which response-based desensitization contributes to the repulsive component of tilt aftereffects (as described in this chapter; see Fig. 5.1), whereas the adaptive re-scaling of orientation-tuning bandwidths contributes to both the repulsive and attractive components. Regardless of the exact origin of the attractive component of tilt aftereffects, multi-channel coding is clearly associated with its characteristic anti-symmetric tuning curves centered at the test orientation (Fig. 5.8(A)).

If the geometric features revealed by the SAEs are also multi-channel coded by multiple neural units tuned to intermediate feature values, they should also exhibit similar anti-symmetric tuning curves centered at the test feature value. This prediction is illustrated in Fig. 5.8(B) for aspect ratio. If the adaptor aspect ratio is identical to the test aspect ratio, no aftereffects should occur (the node). If the adaptor aspect ratio is made flatter than the test aspect ratio, the aftereffect should make the test aspect ratio appear taller than it is, causing the tuning curve to rise above the veridical test aspect ratio. In contrast, if the adaptor aspect ratio is made taller than the test aspect ratio, the aftereffect should make the test aspect ratio appear flatter than it is, causing the tuning curve to fall below the veridical test aspect ratio. Note that the tunings illustrated in Fig. 5.8(B) are broader than in Fig. 5.8(A) because SAEs tend to have broad feature tunings (Suzuki & Rivest 1998; Suzuki 2001).

Qualitatively different tuning curves are expected if aspect ratio is opponent-coded based on the relative activity of the two opponent pools of cells, one broadly tuned to flat aspect ratios and the other broadly tuned to tall aspect ratios (Regan & Hamstra 1992; Suzuki & Rivest 1998; see Fig. 5.7(A)). A neutral adaptor would then adapt the flat and tall pools of cells equally, resulting in no aspect-ratio aftereffect. Thus, unlike in multi-channel coding where the node should follow the test aspect ratio (Fig. 5.8(B)), in opponent coding, the node of the tuning curve should be fixed at the neutral aspect ratio regardless of the test aspect ratio (Fig. 5.8(C)). A flat adaptor would preferentially adapt the flat-tuned pool of cells, making the test aspect ratio appear taller than it is, causing the tuning curve to rise above the veridical test aspect ratio. In contrast, a tall adaptor would preferentially adapt the tall-tuned pool of cells, making the test aspect ratio appear flatter than it is, causing the tuning curve to fall below the veridical test aspect ratio. Thus, if aspect ratio is opponent-coded, all tuning curves should be anti-symmetric about the neutral aspect ratio regardless of the test aspect ratio (Fig. 5.8(C)).

When a series of tuning curves was obtained for aspect ratio aftereffects by orthogonally varying aspect ratios of adaptor and test patterns in one study (Suzuki & Rivest 1998), the shapes of the obtained tuning curves, with their nodes occurring consistently at or near the neutral aspect ratio (Fig. 5.9), were largely consistent with those predicted by opponent coding (Fig. 5.8(C)), and clearly inconsistent with those predicted by multi-channel coding (Fig. 5.8(B)). Recent single-cell results by Kayaert *et al.* (2004) also support the idea that geometric features such as taper, convexity, and curvature may be opponent-coded in IT in a manner illustrated in Fig. 5.7. First, the IT cells they studied exhibited

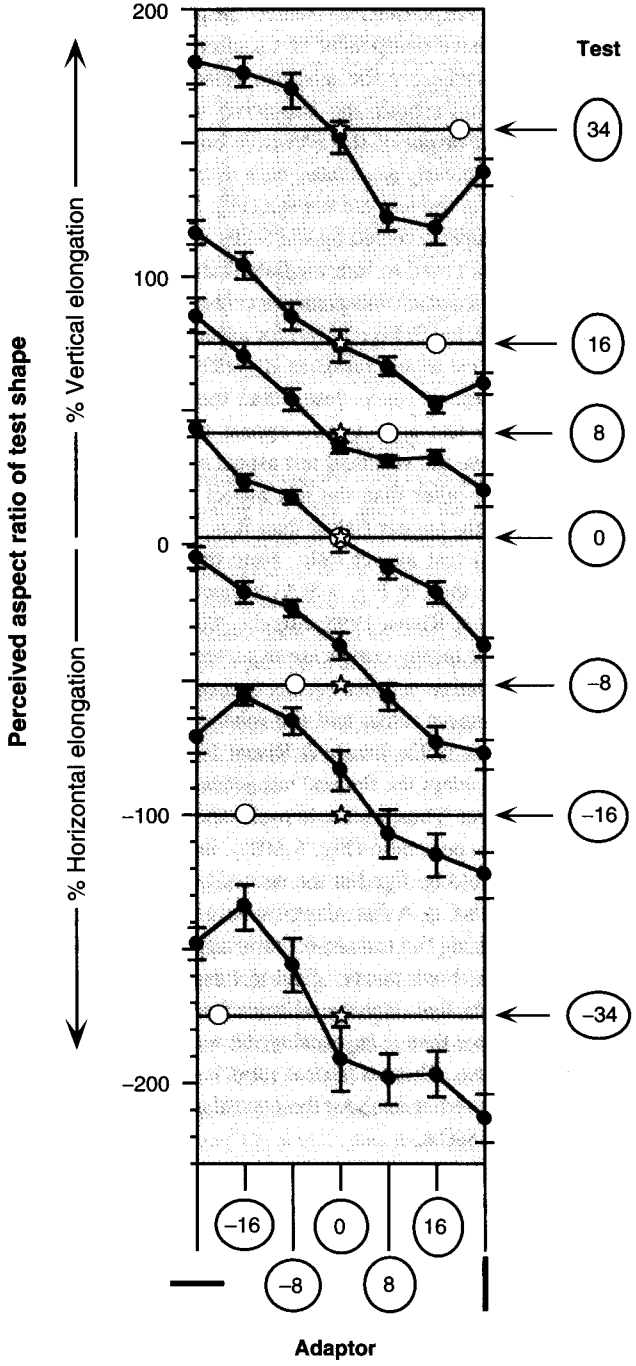

FIG. 5.9 Tuning curves for aspect-ratio aftereffects when the aspect ratios of the adaptor and test shapes were varied orthogonally (data from one observer shown). Aspect ratios were measured as % elongation, ((long axis – short axis)/short axis) × 100%, with positive values indicating vertical elongation and negative values indicating horizontal elongation. The adaptor and test shapes (3.5° in average diameter and presented at 6° eccentricity) were separated by 5°, and were presented in a rapid temporal sequence (see Fig. 5.2); the perceived aspect ratio was measured using a method of adjustment as in Suzuki and Cavanagh (1998). In the figure, the adaptor shapes are illustrated at the bottom and the test shapes on the right (with their % elongations indicated inside). The horizontal line drawn next to each test shape represents its perceived aspect ratio without adaptation (i.e. the baseline); note that the perceived aspect ratio of a briefly flashed (and backward-masked) ellipse tends to be exaggerated (consistent with Suzuki & Cavanagh, 1998). All of the tuning curves obtained for the six test aspect ratios (±8%, ±16%, and ±34% elongations) were more or less anti-symmetric about the neutral (circular) adaptor with baseline crossings (nodes) occurring at or near the neutral adaptor, generally consistent with the opponent-coding hypothesis (see open stars and Fig. 5.8(C) for the predicted baseline crossings), and inconsistent with the multi-channel-coding hypothesis (see open circles and Fig. 5.9(B) for the predicted baseline crossings).

relatively independent tunings for these three geometric features, suggesting that taper, convexity, and curvature are encoded as distinct global dimensions. Second, many of the recorded IT cells responded more strongly to more extreme taper, convexity, and curvature, consistent with the opponent tuning curves illustrated in Fig. 5.7.

There is also evidence that face identity may be partly opponent-coded. Research on identification of caricatured and anti-caricatured faces suggests that faces are encoded in terms of their specific deviations from the norm in the multi-dimensional space defined by face attributes (e.g. Lee *et al.* 2000). Demonstrations of face-specific aftereffects have provided evidence that various face-related dimensions such as simple deformations (e.g. Webster & MacLin 1999), socially relevant dimensions such as face gender, race, expression (e.g. Webster *et al.* 2004), and attractiveness (Rhodes *et al.* 2003) as well as dimensions along which individual faces deviate from the norm (e.g. Leopold *et al.* 2001; Rhodes *et al.* this volume), are represented by a type of central-tendency coding. Significantly, whereas adaptation to a face deviated from the norm caused a distortion on a test face in the opposite direction, adaptation to the norm face caused little distortion (e.g. Webster & MacLin 1999; Leopold *et al.* 2001, also see Leopold and Bondar in this volume). This is similar to aspect ratio aftereffects shown in Fig. 5.8(C) and 5.9 (exhibiting little aftereffect from a neutral-aspect-ratio adaptor). As discussed above, this property is consistent with opponent-coding tuning functions such as those shown in Fig. 5.7. In fact, Rhodes *et al.* (this volume) have applied the Regan and Hamstra's opponent coding model to norm-based coding of face-related features.

In distinguishing opponent coding and multi-channel coding, it is noteworthy that, in the case of a circular dimension such as line orientation (i.e. 180° rotation takes you back to 0° rotation), "opponent-type" aftereffects can be obtained from multi-channel coding. For example, Clifford (this volume) found that following adaptation to a specific orientation, the dots in a random-dot pattern appeared to be grouped into the orthogonal orientation. This orthogonal aftereffect is consistent with multi-channel coding of perceived orientation if one assumes that (1) the central tendency is computed as a vector summation along the entire circular dimension of orientation and (2) the orientation-to-vector-direction mapping is such that orthogonal orientations are represented by vectors with opposite directions (see Clifford, this volume, for details). The latter implies an opponent relationship between orthogonal orientations. However, the coding of perceived orientation is still based on central tendency computed over multiple orientation channels with no orientation being uniquely neutral. This is qualitatively different from the opponent coding discussed here (also in Rhodes *et al.* this volume, and Regan & Hamstra 1992; see Fig. 5.7), which is based on two opponent pools of cells and characterized by a special neutral value (e.g. neutral aspect ratio) adaptation which produces no aftereffects.

5.8 Opponent Coding and Perceptual Multi-stability

Indirect support for opponent codings of global skew, taper, curvature, and convexity has been found using a rather different experimental paradigm involving the phenomena of perceptual multi-stability. While voluntary attention allows for deliberate selection of behaviorally relevant patterns, the process of pattern selection also sustains meta-stable representations in that visual awareness spontaneously shifts from one object to another or from one image interpretation to another, in the absence of voluntary attention shifts or eye movements (see Attneave 1971; Taylor & Aldridge 1974; and Leopold & Logothetis 1999 for reviews). To study the dynamic process of spontaneous pattern selection involving specific sets of shapes, Suzuki & Grabowecky (2002*a*) devised stimuli that induced multi-stable binocular rivalry in which the perceived image spontaneously (and clearly) changed among four shapes.

In binocular rivalry, while the stimulus remains constant, perception changes dramatically due to perturbations in neural responses arising from the conflicting signals coming from the two eyes (see Levelt 1965, Blake 1989, and Blake & Logothetis 2002 for reviews). These perturbations could arise from non-linear inhibitory neural interactions coupled with adaptation and stochastic neural noise (e.g. Sugie 1982; Lehky 1988; Blake 1989; Laing & Chow 2002; Kim *et al.* 2005), synchronization and de-synchronization of neural activity (e.g. Lumer 1998; Srinivasan *et al.* 1999; Fries *et al.* 2001; Suzuki & Grabowecky 2002*b*), and/or from some other yet-to-be specified factors. In conventional binocular rivalry, perception alternates between only two images (one presented

to the left eye and the other presented to the right eye). Significantly, in the specific stimulus configurations used by Suzuki and Grabowecky (2002*a*), perception alternated among four distinct shapes, allowing examinations of transition probabilities (i.e. how visual awareness moves among multiple possible percepts) as well as temporal dynamics (i.e. the duration of each percept) of spontaneous pattern selection.

Note that activities of opponent-tuned pools of cells involved in a specific type of opponent coding (e.g. convex-tuned cells and concave-tuned cells) are likely to be highly interactive. Anatomically, inhibitory connections within IT tend to be short range, limiting strong interactions to primarily within a columnar region (where cells share related pattern selectivity) or across neighboring columns (e.g. Fujita & Fujita 1996; Y. Wang *et al.* 2000). Functionally, mutually inhibitory interactions between opponent-tuned pools of cells would enhance sensitivity for categorical discriminations (e.g. convex versus concave) by facilitating winner-take-all type flipping to one or the other category. If opponent-tuned pools of cells that mediate each opponent coding are indeed anatomically proximal and highly interactive, spontaneous perceptual transitions might be more prevalent between opponent shapes than between non-opponent shapes. Suzuki and Grabowecky (2002*a*) confirmed this prediction. A perceived image was more likely to change to an opponent shape (as well as to a shape with a similar degree of symmetry). In particular, when they used multi-stable stimuli in which perception changed among shapes with opposite skew, taper, curvature, and convexity, perceptual transitions were elevated between each pair of opponent shapes. For example, perception was more likely to change from a convex shape to a concave shape or from a left-curved shape to a right-curved shape, and less likely to change from a convex shape to a curved shape. Such phenomena of "perceptual trapping" within a pair of opponent shapes provided indirect support for opponent codings of skew, taper, curvature, and convexity, corroborating the SAE results (also see Leopold 2003; Rubin 2003; Suzuki & Grabowecky 2003*a*).

5.9 Putting the Codings Together

Ultimately, coding of perceived patterns should be understood on the basis of the overall neural activity throughout the ventral visual pathway (V1 → V2 → V4 → IT). Research on neural pattern tunings and visual aftereffects so far suggests that local orientation and spatial frequency are multi-channel coded in V1 (and up to V4), and certain global configurations of oriented contours such as skew, taper, curvature, aspect ratio, convexity, and relative size, as well as face-specific features, are opponent-coded in IT. Perhaps, any complex geometric features for which a behaviorally significant category boundary lies within a continuum of geometric variations might be opponent-coded in IT. Other object properties that are not conducive to coding in terms of systematic geometric variations around a norm, including highly trained patterns such as English

letters and Chinese characters, are probably coded as broadly distributed patterns of activity; the extent of distributed coding might be within IT (e.g. Tsunoda *et al.* 2001), perhaps involving the left inferior temporo-occipital region (Farah 1990), and/or across temporal lobe visual areas (e.g. Haxby *et al.* 2000, 2001), possibly involving lower visual areas such as V4 where there has as yet been no clear evidence of multi-channel or opponent codings of specific geometric features (see Sections 5.1 and 5.3).

As multiple visual areas are activated simultaneously upon viewing almost any visual pattern, a question arises as to how activities in different visual areas contribute to pattern perception. For example, in viewing a vase, local orientations of the luminance-defined contour edges would be multi-channel coded in V1 and V2, local curvatures of those contours and their configurations would be analyzed in V2 and V4, the overall convexity would be opponent-coded in IT, and other image properties such as brightness, color, texture, and 3D part structure (e.g. consisting of a long neck and a bulged base) would be coded (either separately or collectively) as patterns of distributed activity within specific visual areas and/or across multiple relevant areas. It is still largely a mystery as to how all these distributed parallel neural activities result in the rich visual awareness that people experience even in the simple act of appreciating a vase. Nevertheless, if the visual system is designed to adaptively optimize pattern perception for a given task, it is feasible that perception might be dominated by the neural activity involved in the type of coding(s) that maximizes the task performance. In other words, visual awareness might be flexible in "focusing on" whatever coding provides the highest signal-to-noise ratio with respect to the given task (e.g. Ahissar & Hochstein 2002; Paradiso 2002).

It is well-known that almost any perceptual performance (e.g. selecting a target pattern and identifying it) improves with practice, generally with a rapid initial improvement followed by a more gradual long-term improvement (e.g. Newell & Rosenbloom 1981; Heathcote *et al.* 2000). Depending on the task demand, a variety of mechanisms appear to contribute to practice effects, including (1) fine-tuning of the relevant feature detectors and their interactions at different levels of processing, (2) facilitation of attentional selection of task-relevant features while promoting grouping and suppression of irrelevant distractors, and (3) formation of increasingly task-optimized categorical representations (see Suzuki & Goolsby 2003 for a brief review). Practice effects might also optimize specific population codings by adjusting input weightings and decision criteria (e.g. Morgan 1992; Poggio *et al.* 1992). These various optimization processes engage once the observer becomes reasonably comfortable with the task. At the very beginning of performing a task, however, an important process might be to *hone in on the task-optimum coding* (this idea is articulated in the reverse-hierarchy model proposed by Ahissar and Hochstein [e.g. Ahissar & Hochstein 2002; Hochstein & Ahissar 2002], and is also related to the phenomena of abrupt perceptual learning [e.g. Ahissar & Hochstein 1997; Rubin *et al.* 1997]).

For example, suppose the observer discriminates shapes with overall convexity or concavity, and the shapes consist of variously oriented local

contour segments. On the one hand, if one were to perform a convex/concave discrimination task while the position and size of the shapes were varied, one's perception should "access" the opponent coding of convexity in IT, which is sensitive to convexity but tolerant for translation and scaling. On the other hand, if one were to perform a fine-grained orientation discrimination task on a particular contour segment at a particular location, one's perception should then access the high-resolution multi-channel coding of local orientation, which is unique to V1 (see Paradiso 2002). If a given task requires a specific geometric coding and the task-optimum coding is selected on the basis of its enhanced activation (e.g. Huk & Heeger 2000), then the initial process of selecting coding might be observed in a brain imaging study in terms of the relatively extensive and rapidly-changing patterns of neural activity at the very beginning of performing the task, quickly converging to localized and stable activity after the rapid initial adjustment period (see Petersen *et al.* 1998 for related results). The process of selecting a task-optimum coding is likely to be especially important when a task is novel. Perhaps poor perceptual learners are those who fail to select the optimum coding for the task; failing at this initial stage would not allow the process of gradual optimization and calibration of the coding to proceed.

5.10 Summary

An ultimate question of visual pattern perception is how neural activity spreads throughout the ventral visual pathway (with concurrent responses to different components and aspects of retinal stimulation) to generate a seemingly coherent and unitary visual awareness of objects and scenes. Extensive research on local aftereffects has linked the systematic organization of neural tunings for orientation and spatial frequency in early cortical visual areas to the central-tendency codings of the perception of those features. Organizations of neural tunings, however, become ambiguous when higher cortical visual areas are examined as cells in those areas integrate a variety of combinations of lower-level signals to achieve increasingly complex and heterogeneous pattern tunings. Indeed, common objects appear to be encoded primarily on the basis of complex and distributed patterns of neural activity. Individual clusters of neural activity might represent parts, combinations of parts, or global aspects of an object. It is largely unclear how these distributed clusters of activity, individually and collectively, map on to different geometric characteristics of an object (e.g. part shape, part structure, global geometry).

Recent demonstrations of global-shape aftereffects suggest that at least some well-learned global geometric attributes are central-tendency coded in high-level visual areas, perhaps because central-tendency codings afford fast and fine-tuned discriminations along the coded dimensions. For example, shape aftereffects (SAEs) using brief stimulus sequences have revealed central-tendency codings for a set of global geometric features, including skew, taper, curvature, aspect ratio, convexity, relative size, and face orientation in depth (Fig. 5.3).

The fact that these geometric features correspond to those that are systematically distorted when surfaces are rotated in depth (Fig. 5.6) suggests that central-tendency codings of these features might be an integral part of the mechanisms that mediate pictorial-cue-based extraction of 3D surface orientation and viewpoint normalization. Other researchers have demonstrated aftereffects for face-specific features, revealing central-tendency codings of face-related attributes. Interestingly, while low-level central-tendency codings of local orientation and spatial frequency are mediated by cells tuned to multiple intermediate feature values (e.g. Fig. 5.1)–multi-channel coding (perhaps, optimized for fine-grained analyses of local details), the tuning properties of global aftereffects suggest that each high-level central-tendency coding is mediated primarily by two opponent pools of cells broadly tuned to opposite feature directions across a norm (e.g. Fig. 5.7)–opponent coding (perhaps, optimized for norm-based categorical discriminations of complex patterns).

It is likely that different types of codings, such as general-purpose codings by distributed patterns of neural activity, specialized codings including central-tendency codings (e.g. multi-channel coding and opponent coding), as well as temporal codings based on the precise dynamics of neural firing (e.g. Ferster & Spruston 1995; Mechler *et al.* 1998; Sugase *et al.* 1999; Ikegaya *et al.* 2004), modify and/or develop adaptively with experience. A challenge is to specify what types of interactions among stimuli, perceptual demands, and preexisting neural circuits lead to developments of different codings, and how those codings are anatomically and physiologically organized within and across different areas of the brain. An equally challenging task is to understand how different codings, spread across the ventral visual pathway, selectively and/or interactively contribute to the rich perceptual experience of visual objects and scenes.

Acknowledgements

I would like to thank Marcia Grabowecky for her invaluable comments on early versions of the manuscript. This work was supported by a National Institutes of Health grant EY014110.

References

Abdi, H., Jiang, F., O'Toole, A.J., & Haxby, J. (2003). Pattern-based analyses for brain imaging data: applications to face processing. *Abstracts of the Psychonomic Society*, *8*, 26.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–6.

Ahissar, M., & Hochstein, S. (2002). The role of attention in learning simple visual tasks. In M. Fahle, & T. Poggio (ed.), *Perceptual Learning* (pp. 253–72). Cambridge, MA: MIT Press.

Albrecht, D.G., Farrar, S.B., & Hamilton, D.B. (1984). Spatial contrast adaptation characteristics of neurons recorded in the cat's visual cortex. *Journal of Physiology*, *347*, 713–39.

Attneave, F. (1971). Multistability in perception. *Scientific American*, *225*, 63–71.

Biedermen, I. (1987). Recognition by components: a theory of human image understanding. *Psychological Review*, *94*, 115–47.

Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*, 145–167.

Blake, R., & Logothetis, N.K. (2002). Visual competition. *Nature Neuroscience*, *3*, 1–11.

Blakemore, C, & Nachmias, J. (1971). The orientational specificity of two visual after-effects. *Journal of Physiology*, *213*, 157–74.

Blakemore, C., Nachmias, J., & Sutton, P. (1970). The perceived spatial frequency shift: Evidence for frequency selective neurons in the human brain. *Journal of Physiology*, *210*, 727–50.

Blakemore, C., & Sutton, P. (1969). Size adaptation: A new aftereffect. *Science*, *166*, 245–7.

Bonds, A.B. (1991). Temporal dynamics of contrast gain in single cells of the cat striate cortex. *Visual Neuroscience*, *6*, 239–55.

Braddick, O., Campbell, F.W., & Atkinson, J. (1978). Channels in vision: Basic aspects. In R. Held, H.W. Leibowitz, & L.-H. Teuber (ed.), *Handbook of Sensory Physiology: Vol. 8. Perception* (pp. 3–38). Berlin: Springer-Verlag.

Brincat, S.L., & Connor, C.E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, *8*(7), 880–6.

Cameron, E.L., Tai, J.C., & Carrasco, M. (2002). Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, *42*, 949–67.

Campbell, F. W., & Maffei, L. (1971). The tilt after-effect: a fresh look. *Vision Research*, *11*, 833–40.

Carandini, M., Movshon, J.A., & Ferster, D. (1998). Pattern adaptation and cross-orientation interactions in the primary visual cortex. *Neuropharmacology*, *37*, 501–11.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*(6), 2918–40.

Cheng, K., Hasegawa, T., Saleem, K.S., & Tanaka, K. (1994). Comparison of neural sensitivity for stimulus speed, length, and contrast in the prestriate visual cortical areas V4 and MT of the macaque monkey. *Journal of Neurophysiology*, *71*(6), 2269–80.

De Valois, R.L., Albrecht, D.G., & Thorell, L.G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–59.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dahaene-Lambertz, G. *et al.* (1998). Imaging unconscious semantic priming. *Nature*, *395*, 597–600.

Deneve, S., Latham, P.E., & Pouget, A. (1999). Reading population codes: a neural implementation of ideal observers. *Nature Neuroscience*, *2*(8), 740–45.

Desimone, R., Albright, T.D., Gross, C.G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, *8*, 2051–62.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.

Desimone, R., & Gross, C.G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, *178*, 363–80.

Desimone, R., & Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, *57*(3), 835–68.

Dow, B.M., Snyder, A.Z., Vautin, R.G., & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research*, *44*, 213–28.

Enns, J.T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*(9), 345–52.

Epstein, R., Graham, K.S., & Downing, P.E. (2003). View-point specific scene representations in human parahippocampal cortex. *Neuron*, *37*, 865–76.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, *23*, 115–25.

Fang, F., & He, S. (2004). Viewer-centered object representations in human visual system revealed by viewpoint aftereffect. *Abstracts of the Vision Sciences Society*, *4*, 71.

Farah, M.J. (1990). *Visual Agnosia: Disorders of Object Vision and What They Tell Us About Normal Vision*. Cambridge, MA: MIT Press.

Ferster, D. & Spruston, N. (1995). Cracking the neuronal code. *Science*, *270*, 756–7.

Foster, K.H., Gaska, J.P., Nagler, M., & Polen, D.A. (1985). Spatial and temporal frequency selectivity of neurons in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, *365*, 331–68.

Fries, P., Reynolds, J.H., Rorie, A.E., & Desimone, R. (2001). Modulation of oscillatory neural synchronization by selective visual attention. *Science*, *291*, 1560–63.

Fujita, I., & Fujita, T. (1996). Intrinsic connections in the macaque inferior temporal cortex. *Journal of Computational Neurology*, *368*, 467–86.

Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, *360*, 343–6.

Gallant, J.L., Connor, C.E., Rakshit, S., Lewis, J.W., & Van Essen, D.C. (1996). Neural responses to polar, hyperbolic, and cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, 76(4), 2718–39.

Gattass, R., Sousa, A.P.B., & Gross, C.G. (1988). Visuotopic organization and extent of V3 and V4 of the macaque. *Journal of Neuroscience, 8*(6), 1831–45.

Gauthier, I. (2000). What constrains the organization of the ventral temporal cortex? *Trends in Cognitive Sciences*, *4*(1), 1–2.

Gauthier, I., Curran, T., Curby, K.M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*(4), 428–32.

Gauthier, I., Skudlarski, P., Gore, J.C., & Anderson, A.W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–7.

Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., & Gore, J.C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*(6), 568–73.

Geisler, W.S., & Albrecht, D.G. (1997). Visual cortex neurons in monkey and cats: Detection, discrimination, and identification. *Visual Neuroscience*, *14*, 897–919.

Gibson, J.J., & Radner, M. (1937). Adaptation, aftereffect and contrast in the perception of tilted lines. *Journal of Experimental Psychology*, *20*, 453–67.

Gross, C.G., Rocha-Miranda, C.E., & Bender, D.B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, *35*, 96–111.

Gur, M., Kagan, I., & Snodderly, M.D. (2004). Lack of short-term adaptation in V1 cells of the alert monkey. *Abstracts of the Vision Sciences Society*, *4*, 255.

Hammond, P., Pomfrett, C.J.D., & Ahmed, B. (1989). Neural motion after-effects in the cat's striate cortex: orientation selectivity. *Vision Research*, *29*(12), 1671–83.

Hanazawa, A., & Komatsu, H. (2001). Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. *Journal of Neuroscience*, *21*(12), 4490–7.

Hasselmo, M.E., Rolls, E.T., & Baylis, G.C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioral Brain Research*, *32*, 203–18.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral visual cortex. *Science*, *293*, 2425–30.

Haxby, J.V., Ishai, A., Chao, L.L., Ungerleider, L.G., & Martin, A. (2000). Object-form topology in the ventral temporal lobe. *Trends in Cognitive Sciences*, *4*(1), 3–4.

Heathcote, A., Brown, S., & Mewhort, D.J.K. (2000). The power law repealed: the case for an exponential law of practice. *Psychonomic Bulletin & Review*, *7*, 185–207.

Hegdé, J., & Van Essen, D.C. (2000). Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience*, *20*, RC61.

Held, R., & Shattuck, S.R. (1971). Color- and edge-sensitive channels in the human visual system: tuning for orientation. *Science*, *174*, 314–16.

Hikosaka, K. (1999). Tolerances of responses to visual patterns in neurons of the posterior inferotemporal cortex in the macaque against changing stimulus size and orientation, and deleting patterns. *Behavioral Brain Research*, *100*, 67–76.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*, 791–804.

Hubel, D.H., & Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215–43.

Huk, A.C., & Heeger, D.J. (2000). Task-related modulation of visual cortex. *Journal of Neurophysiology*, *83*, 3525–36.

Ikegaya, Y., Aaron, G., Cossart, R., Aronov, D., Lampl, I., Ferster, D. *et al.* (2004). Synfire chains and cortical songs: temporal modules of cortical activity. *Science*, *304*, 559–64.

Ito, M., Fujita, I., Tamura, H., & Tanaka, K. (1994). Processing of contrast polarity of visual images in inferotemporal cortex of the macaque monkey. *Cerebral Cortex*, *5*, 499–508.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, *73*(1), 218–26.

Kanwisher, N.G., McDermott, J., & Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4301–11.

Kayaert, G., Op de Beeck, H., Biederman, I., & Vogels, R. (2004). Shape dimension-dependent coding of macaque IT neurons. *Abstracts of the Vision Sciences Society*, *4*, 70.

Kayaert, G., Biederman, I., & Vogels, R. (2003). Shape tuning in macaque inferior temporal cortex. *Journal of Neuroscience, 23*(7), 3016–27.

Kim, Y.-J., Grabowecky, M., & Suzuki, S. (2005). Stochastic resonance in binocular rivalry. Submitted.

Kobatake, E., Wang, G., & Tanaka, K. (1998). Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology*, *80*, 324–30.

Kohler, W., & Wallach, H. (1944). Figural aftereffects: an investigation of visual processes. *Proceedings of the American Philosophical Society*, *88*, 269–357.

Komatsu, H., Ideura, Y., Kaji, S., & Yamane, S. (1992). Color selectivity of neurons in the inferior temporal cortex of the awake macaque monkey. *Journal of Neuroscience*, *12*(2), 408–24.

Lamme, V.A.F. (2001). Blindsight: the role of feedforward and feedback corticocortical connections. *Acta Psychologica, 107*, 209–228.

Lamme, V.A.F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex, 9*, 406–413.

Lee, C., Rohrer, W.H., & Sparks, D.L. (1988). Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature, 332*, 357–60.

Lee, K., Byatt, G., & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychological Science, 11*(5), 379–85.

Lehky, S.R. (1988). An astable multivibrator model of binocular rivalry. *Perception, 17*, 215–28.

Leopold, D.A. (2003). Visual perception: shaping what we see. *Current Biology, 13*, R10–R12.

Leopold, D.A., & Logothetis, N.K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences, 3*, 254–64.

Leopold, D.A., O'Toole, A.J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after-effects. *Nature Neuroscience, 4*(1), 89–94.

Levelt, W.J.M. (1965). *On Binocular Rivalry*. Soesterberg, The Netherlands: Institute for Perception RVO-TNO.

Leventhal, A.G., Wang, Y., Schmolesky, M.T., & Zhou, Y. (1998). Neural correlates of boundary perception. *Visual Neuroscience, 15*, 1107–18.

Levitt, J.B., Kiper, D.C., & Movshon, J.A. (1994). Receptive fields and functional architecture of macaque V2. *Journal of Neurophysiology, 71*(6), 2517–42.

Laing, C.R., & Chow, C.C. (2002). A spiking neuron model for binocular rivalry. *Journal of Computational Neuroscience, 12*, 39–53.

Logothetis, N.K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology, 5*, 552–63.

Logothetis, N.K., & Sheinberg, D.L. (1996). Visual object recognition. *Annual Review of Neuroscience, 19*, 577–621.

Luck, S.J., Chelazzi, L., Hillyard, S., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology, 77*, 24–42.

Lueschow, A., Miller, E.K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex, 5*, 523–31.

Lumer, E.D. (1998). A neural model of binocular integration and rivalry based on the coordination of action-potential timing in primary visual cortex. *Cerebral Cortex, 8*, 553–61.

Magnussen, S., & Kurtenbach, W. (1980). Linear summation of tilt illusion and tilt aftereffect. *Vision Research, 20*, 39–42.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

McAdams, C.J., & Maunsell, J.H.R. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscicence, 19*(1), 431–41.

Mechler, F., Victor, J.D., Purpura, K.P., & Shapley, R. (1998). Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. *Journal of Neuroscience, 18*(16), 6583–98.

Miller, E.K., Gochin, P.M., & Gross, C.G. (1993a). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque monkey by addition of a second stimulus. *Brain Research, 616*, 25–9.

Miller, E.K., Li, L, & Desimone, R. (1993b). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience, 13*(4), 1460–78.

Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: two central pathways. *Trends in Neurosciences, 6*, 414–17.

Mitchel, D.E. & Muir, D.W. (1976). Does the tilt after-effect occur in the oblique meridian? *Vision Research, 16*, 609–13.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in extrastriate cortex. *Science, 229*, 782–4.

Morgan, M.J. (1992). Hyperacuity of those in the know. *Current Biology, 2*(9), 481–2.

Muller, J.R., Metha, A.B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science, 285*, 1405–8.

Newell, A., & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the power law of practice. In J.R. Anderson (ed.), *Cognitive Skills and Their Acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Nothdurft, H-C., Gallant, J.L., & Van Essen, D.C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience, 16*, 15–34.

O'Shea, R.P., Wilson, R.G., & Duckett, A. (1993). The effects of contrast reversal on the direct, indirect, and interocularly-transferred tilt aftereffect. *New Zealand Journal of Psychology, 22*(2), 94–100.

O'Toole, B., & Wenderoth, P. (1977). The tilt illusion: Repulsion and attraction effects in the oblique meridian. *Vision Research, 17*(3), 367–74.

Oram, M.W., & Perrett, D.I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology, 68*(1), 70–84.

Paradiso, M.A. (2002). Perceptual and neuronal correspondence in primary visual cortex. *Current Opinion in Neurobiology, 12*, 155–161.

Pascal-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*, 510–21.

Pasupathy, A., & Connor, C.E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology, 82*, 2490–502.

Pasupathy, A., & Connor, C.E. (2001). Shape representation in area V4: position-specific tuning for boundary conformation. *Journal of Neurophysiology, 86*, 2505–19.

Peterhans, E., & von der Heydt, R. (1991). Subjective contours—bridging the gap between psychophysics and physiology. *Trends in Neurosciences, 14*, 112–9.

Petersen, S.E., van Mier, H., Fiez, J. A., & Raichle, M.E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences USA, 95*, 853–60.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, 256*, 1018–21.

Regan, D., & Hamstra, S.J. (1992). Shape discrimination and the judgment of perfect symmetry: dissociation of shape from size. *Vision Research, 32*, 1845–64.

Reich, D.S., Mechler, F., & Victor, J.D. (2001). Temporal coding of contrast in primary visual cortex: when, what, and why. *Journal of Neurophysiology, 85*, 1039–50.

Reynolds, J.H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19*(5), 1736–53.

Reynolds, J.H., & Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron, 37*, 853–63.

Reynolds, J.H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*, 703–14.

Rhodes, G., Byatt, G., Michie, P.T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience*, *16*(2), 189–203.

Rhodes, G., Jeffery, L., Watson T.L., Clifford C.W.G., & Nakayama K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychological Science*, *14*(6), 558–66.

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, *3*, 1199–204.

Rivest, J., Intriligator, J., Suzuki, S., & Warner, J. (1998). A shape distortion effect that is size invariant. *Investigative Ophthalmology & Visual Science (Suppl.)*, *39*(4), S853.

Rolls, E.T. & Baylis, G.C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, *65*, 38–48.

Rubin, N. (2003). Binocular rivalry and perceptual multi-stability. *Trends in Neurosciences*, *26*(6), 289–91.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology*, *7*, 461–7.

Sagara, M., & Ohyama, T. (1957). Experimental studies of figural aftereffects in Japan. *Psychological Bulletin*, *54*, 327–38.

Sato, T., Kawamura, T., & Iwai, E. (1980). Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. *Experimental Brain Research*, *38*, 313–19.

Saul, A.B., & Cynader, M.S. (1989). Adaptation in single units in visual cortex: The tuning of aftereffects in the spatial domain. *Visual Neuroscience*, *2*, 593–607.

Sclar, G., Lennie, P., & DePriest, D.D. (1989). Contrast adaptation in striate cortex of macaque. *Vision Research*, *29*(7), 747–55.

Sclar, G., Maunsell, J.H.R., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, *30*(1), 1–10.

Sheth, B.R., Sharma, J., Rao, S.C., & Sur, M. (1996). Orientation maps of subjective contours in visual cortex. *Science*, *274*, 2110–5.

Sigala, N., & Logothetis, N.K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*, 318–20.

Spivey, M.J., & Spirn, M.J. (2000). Selective visual attention modulates the direct tilt aftereffect. *Perception & Psychophysics*, *62*(8), 1525–33.

Srinivasan, R., Russell, D.P., Edelman, G.M., & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *Journal of Neuroscience*, *19*(13), 5435–48.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*, 869–73.

Sugie, N. (1982). Neural models of brightness perception and retinal rivalry in binocular vision. *Biological Cybernetics*, *43*, 13–21.

Suzuki, S. (1999). Influences of contexts on a non-retinotopic skew-contrast effect. *Investigative Ophthalmology & Visual Science (Suppl.)*, *40*(4), S812.

Suzuki, S. (2001). Attention-dependent brief adaptation to contour orientation: a high-level aftereffect for convexity? *Vision Research*, *41*(28), 3883–902.

Suzuki, S. (2003a). Attentional selection of overlapped shapes: a study using brief shape aftereffects. *Vision Research*, *43*, 549–61.

Suzuki, S. (2003b). The high and low of visual awareness. *Neuron*, *39*, 883–84.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: an attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 443–63.

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1315–41.

Suzuki, S., & Goolsby, B.A. (2003). Sequential priming is not constrained by the shape of long-term learning curves. *Perception & Psychophysics*, *65*(4), 632–48.

Suzuki, S., & Grabowecky, M. (2002a). Evidence for perceptual "trapping" and adaptation in multistable binocular rivalry. *Neuron*, *36*, 143–57.

Suzuki, S., & Grabowecky, M. (2002b). Overlapping features can be parsed on the basis of rapid temporal cues that produce stable emergent percepts. *Vision Research*, *42*, 2669–92.

Suzuki, S., & Grabowecky, M. (2003a). Response: Binocular rivalry and perceptual multi-stability. *Trends in Neurosciences*, *26*(6), 287–9.

Suzuki, S., & Grabowecky, M. (2003b). Attention during adaptation weakens negative afterimages. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(4), 793–807.

Suzuki, S., & Rivest, J. (1998). Interactions among "aspect-ratio channels." *Investigative Ophthalmology & Visual Science (Suppl.)*, *39*(4), S855.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–39.

Taylor, M.M., & Aldridge, K.D. (1974). Stochastic processes in reversing figure perception. *Perception & Psychophysics*, *16*(1), 9–27.

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience*, *4*(8), 832–8.

Ullman, S. (1998). Three-dimensional object recognition based on the combination of views. *Cognition*, *67*, 21–44.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J.W. Mansfield (ed.), *Analysis of Visual Behavior* (pp. 549–86). Cambridge, MA: MIT Press.

Vautin, R.G., & Berkley, M.A. (1977). Responses of single cells in cat visual cortex to prolonged stimulus movement: Neural correlates of visual aftereffects. *Journal of Physiology*, *40*(5), 1051–65.

Vogels, R. (1990). Population coding of stimulus orientation by striate cortical cells. *Biological Cybernetics*, *64*, 25–31.

Vogels, R., & Orban, G.A. (1994). Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *Journal of Neurophysiology*, *71*(4), 1428–51.

Vogels, R., Sary, G., & Orban, G.A. (1995). How task-related are the responses of inferior temporal neurons? *Visual Neuroscience*, *12*, 207–14.

Wang, G., Tanaka, K., & Tanifuji, M. (1996). Optical imaging of functional organization in the monkey inferotemporal cortex. *Science*, *272*, 1665–8.

Wang, Y., Fujita, I., & Maruyama, Y. (2000). Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nature Neuroscience*, *3*(8), 807–13.

Webster, M., Kaping, D., Mizokami, Y., & Dumahel, P. (2004). Adaptation to natural face categories. *Nature*, *428*, 558–61.

Webster, M., & MacLin, O.H. (1999). Figural after-effects in the perception of faces. *Psychonomic Bulletin & Review*, *6*(4), 647–53.

Wenderoth, P., & van der Zwan, R. (1989). The effects of exposure duration and surrounding frames on direct and indirect tilt aftereffects and illusions. *Perception & Psychophysics*, *46*(4), 338–44.

Wolfe, J.M. (1984). Short test flashes produce large tilt aftereffects. *Vision Research*, *24*(12), 1959–64.

Young, M.P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, *256*, 1327–31.

Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research*, *41*, 2979–94.

Zipser, K., Lamme, V.A.F., & Shiller, P.H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, *16*(22), 7376–89.