

A Pluralist Perspective on Shape Constancy

E. J. Green

Abstract

The ability to perceive the shapes of things as enduring through changes in how they stimulate our sense organs is vital to our sense of stability in the world. But what sort of capacity *is* shape constancy, and how is it reflected in perceptual experience? This paper defends a *pluralist* account of shape constancy: There are multiple kinds of shape constancy centered on geometrical properties at various levels of abstraction, and properties at these various levels feature in the content of perceptual experience, governing patterns of apparent shape similarity. I propose that the varieties of shape constancy are subserved by the syntactic complexity of perceptual shape representations. By assigning discrete constituents to various abstract shape parameters, these representations attune us to the preservation of certain abstract shape properties through changes in more determinate shape properties. Finally, I draw broader lessons concerning the nature and function of perceptual constancy.

- 1 *Introduction*
- 2 *Object-Centered Representation and Robust Constancy*
- 3 *Skeletal Approaches to Shape Representation*
- 4 *The Varieties of Shape Constancy*
 - 4.1 *Determinate shape*
 - 4.2 *Compositional structure*
 - 4.3 *Skeletal shape*
 - 4.4 *Skeletal topology*
 - 4.5 *Two concerns*
 - 4.6 *Summing up*
- 5 *Conclusion*

1 Introduction

The ability to perceive aspects of the environment as remaining stable despite variation in sensory input is a remarkable feat. As you move about your kitchen, your table seems to retain its shape and size although its orientation, distance, and retinal projection are in constant flux. Accounting for the compresence of stability and variation in perceptual experience is a central aim in the philosophy of perception. The present paper examines this phenomenon in the case of shape. I develop an account of shape constancy that illuminates salient but poorly understood aspects of spatial experience.

I advocate a *pluralist* account on which shape constancy is a complex capacity composed of various “sub-constancies.”¹ There are multiple varieties of shape constancy centered on geometrical properties at varying levels of abstraction. This pluralism is grounded in the syntactic complexity of perceptual shape representations. By assigning discrete constituents to different abstract shape parameters, these representations attune us to deep resemblances in abstract shape despite variation in fine-grained, determinate shape. More abstract constancies are exercised through changes that “peel away” further constituents of these representations.

The most familiar sort of shape constancy, involving stability of perceived shape across shifts in viewpoint, is shown in figure 1. While this capacity is undeniably important, here I call attention to forms of invariance in shape perception exercised not merely through viewpoint shifts, but through changes in an object’s intrinsic structure. Consider figures 2-4. Each depicts objects that differ perceptibly in determinate shape, but there is also a salient resemblance between them—some deep structural characteristic that they seem to share, despite superficial differences. The challenge is to explain these kinds of shape similarity and the perceptual representations underlying them.



Figure 1. A coin continues to be perceived as circular despite a change in viewpoint.

¹ Pluralist conceptions have recently been developed for other constancies, notably color (Foster [2011]; Wright [2013]; Davies [2022]).

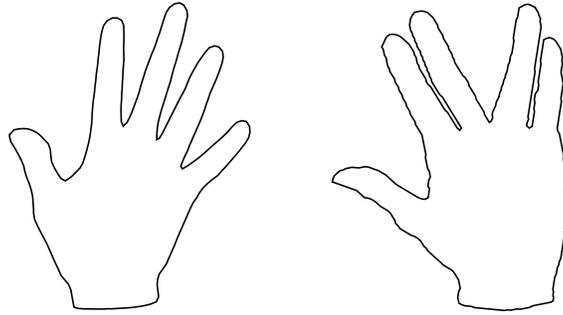


Figure 2. After Lowet et al. ([2018]). Thanks to Chaz Firestone for this figure.

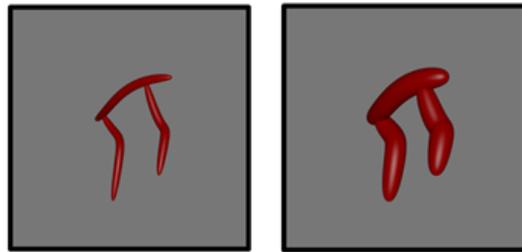


Figure 3. *Source:* Ayzenberg and Lourenco ([2019]). Reprinted under the Creative Commons Attribution 4.0 International License.

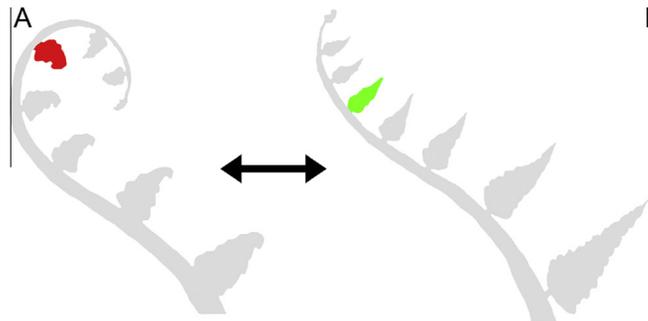


Figure 4. *Source:* Schmidt and Fleming ([2016]). Reprinted under the Creative Commons CC-BY-NC-ND license.

I propose that each of these cases involves a distinctive form of constancy. We perceive an abstract shape property as invariant through changes in determinate geometrical structure.

Accordingly, we must broaden our conception of shape constancy, and perceptual constancy more

generally. Constancy is commonly construed as an ability to perceive certain *intrinsic* properties of objects as invariant through changes in *relational*, *extrinsic*, or *situational* properties—for example, orientation or illumination.² This conception undergirds many standard characterizations of shape constancy. For example:

Shape constancy is defined usually as the relative constancy of the perceived shape of an object despite variations in its orientation. (Epstein and Park [1963], p. 265)

Shape constancy...is the ability to perceive an object as being the same shape despite changes in its orientation relative to the observer. (Slater and Morison [1985], p. 337)

If I am right, such characterizations are too narrow in an important respect. Perceptual constancy, and shape constancy in particular, is exercised not merely through changes in extrinsic properties such as orientation or illumination, but also through changes in an object's intrinsic properties—specifically, changes altering more volatile intrinsic properties while preserving more stable ones.

The plan is as follows. Section 2 distinguishes *minimal* from *robust* conceptions of shape constancy and argues that our perceptual systems achieve robust constancy. Section 3 introduces a prominent approach to object-centered representation—*skeletal representation*—and draws out its implications for the format of perceptual shape representations. Section 4 argues that skeletal representations underpin four varieties of shape constancy, which I label constancies for *determinate shape*, *compositional structure*, *skeletal shape*, and *skeletal topology*. I assemble evidence for each and argue that they figure in perceptual experience. Section 5 discusses implications for the nature and function of perceptual constancy.

2 Object-Centered Representation and Robust Constancy

I will understand perceptual constancy as a representational capacity—roughly, a capacity to perceptually represent properties of objects as remaining invariant despite relevant changes in

² See Schellenberg ([2008]), Allen ([2018]), and O'Dea ([2020]).

proximal stimulation received from them.³ In shape perception, this involves a capacity to perceptually represent shape properties as remaining invariant despite variation in projected 2D retinal shape.

We can distinguish *minimal* from *robust* conceptions of shape constancy. Let's say a perceiver or perceptual system exhibits *minimal* shape constancy when it perceptually represents some single shape property under relevantly different conditions of proximal stimulation—e.g., across significant variation in an object's retinal projection.⁴ If you continue to perceptually represent the surface of your table as rectangular despite variation in the shape or size of its retinal projection, then you exhibit minimal shape constancy.

Let's say a perceiver or perceptual system exhibits *robust* shape constancy when it perceptually represents some single shape property in the *same way*, or via *tokens of a single representation-type*, under relevantly different conditions of proximal stimulation. Robust shape constancy requires minimal shape constancy, but not vice versa. Consider your table again. As you move around the table, its 2D projection changes, and you continue to perceptually represent its rectangularity; thus, you exhibit minimal shape constancy. However, it is an open question whether you perceptually represent its rectangularity in the *same way* across perspectives.

Suppose that perception represents the table's rectangularity by encoding its spatial layout within a viewer-centered reference frame—for example, a frame anchored to the viewer's cyclopean eye, with cardinal axes fixed by the viewer's up, down, left, and right directions. Then the table's layout in this frame—and thus the way its rectangularity is represented—changes with any shift in perspective on the table. Thus, one might exhibit minimal shape constancy with respect to the table, but not robust constancy. Some prominent models of object recognition positing only viewer-

³ Representational conceptions of constancy are popular (Rock [1983]; Hilbert [2005]; Burge [2010]; Rescorla [2014]; Cohen [2015]; Green [2019]), but not universally accepted (Gibson [1979]; Olin [2016]; Buccella [2021]).

⁴ For more on the distinction between relevant and irrelevant proximal changes, see (Green [2023]).

centered shape representations are suited to the view that shape constancy is *merely* minimal—minimal but not robust (Ullman and Basri [1991]; Bülthoff and Edelman [1992]; Edelman and Duvdevani-Bar [1997]).

Conversely, suppose that perception represents the table's rectangularity within an *object-centered* reference frame determined by the table's axes of symmetry. If so, it is possible to token the same representation-type despite shifts in the table's orientation. The table's layout relative to its intrinsic axes remains unchanged through these shifts, while its layout in a viewer-centered frame does not (Marr and Nishihara [1978]; Erdogan and Jacobs [2017]). A representation-type that encodes the table's layout relative to its intrinsic axes may be tokened (veridically) across different perspectives on the table.⁵

In what follows, I'll be concerned with robust shape constancy. This approach departs from some other discussions of shape constancy, which analyze it as a capacity to represent a single shape property differently (say, under different modes of presentation) from different perspectives (Burge [2010], [2014]; see also Thompson [2010]; Schellenberg [2018]). While I don't deny that we *can* form viewer-centered representations that vary with perspective (Briscoe [2009]; Lande [2018]), I suggest that such representations occur *alongside* object-centered representations, which subservise robust shape constancy.

I cannot offer a sustained defense of object-centered representation here. However, I adduce two lines of support. First, certain patterns of shape *misperception* are governed by objects' intrinsic axes in ways that are easily explainable on object-centered models but not viewer-centered models. Consider *mirror-image confusion*, where perceivers confuse an object with its mirror image (for instance,

⁵ The object-/viewer-centered dichotomy is somewhat idealized. Broadly object-centered models can incorporate viewer-centered elements. For example, Biederman's ([1987]) geon scheme encodes part shapes in an object-centered manner, but characterizes categorical relations between parts (for example, *above*) relative to the viewer or direction of gravity.

“b” vs. “d”). Human children (Rudel and Teuber [1963]), adults (Nickerson and Adams [1979]), and even octopuses (Sutherland [1960]) exhibit mirror-image confusions in shape recognition.

Furthermore, discrimination of mirror-reversed objects can be selectively impaired: Vannuscorps *et al.* ([2022]) examined a patient who, when asked to draw a sample shape, often produced mirror-reversed versions, and confused letters with their mirror-reversed counterparts (see also Turnbull and McCarthy [1996]). Moreover, fMRI data suggest that posterior fusiform sulcus—a high-level ventral area—is sensitive to most shape differences but not to mirror-reversals (Dilks *et al.* [2011]). Mirror-image confusions also commonly occur in normally-sighted adults when either reidentifying or drawing objects at their original orientations (Gregory and McCloskey [2010]; Chaisilprungraung *et al.* [2019]).

Object-centered schemes encode a shape’s boundary elements via their distances and directions from corresponding points or segments of the shape’s intrinsic axis. Mirror-image confusions are analyzable as cases where distances of boundary elements from the axis are coded correctly but directions are systematically flipped. Conversely, if the representation of shape is wholly viewer-centered, it is unclear why systematic mirror-image reversals would occur (Gregory and McCloskey [2010]; Chaisilprungraung *et al.* [2019]),⁶ since any systematic errors should presumably be governed by the intrinsic axes of the perceiver, not the object.

Unilateral neglect, a condition caused by parietal lobe lesions, offers another line of evidence. Neglect patients fail to attend to one side of space—typically the left. However, the “left” is definable in viewer-centered, object-centered, or scene-centered frames of reference (Humphreys *et al.* [2013]). Driver and Halligan ([1991]) found evidence for object-centered neglect. Participants judged the sameness/difference of two shapes, which could differ with respect to elements either on the right or left of their respective axes of elongation. Neglect patients struggled to discriminate

⁶ More precisely, it is unclear why we would observe mirror-reversals *without* systematic errors of localization.

shapes when the differentiating elements landed on the left, and crucially this pattern persisted when elements on the left of the *axis* fell within the subject's *right* hemifield (see also Tipper and Behrmann [1996]).⁷

Thus, there is evidence that the visual system forms object-centered shape representations. Because object-centered representations can endure through relevant changes in an object's retinal projection, they provide a basis for robust shape constancy.

3 Skeletal Approaches to Shape Representation

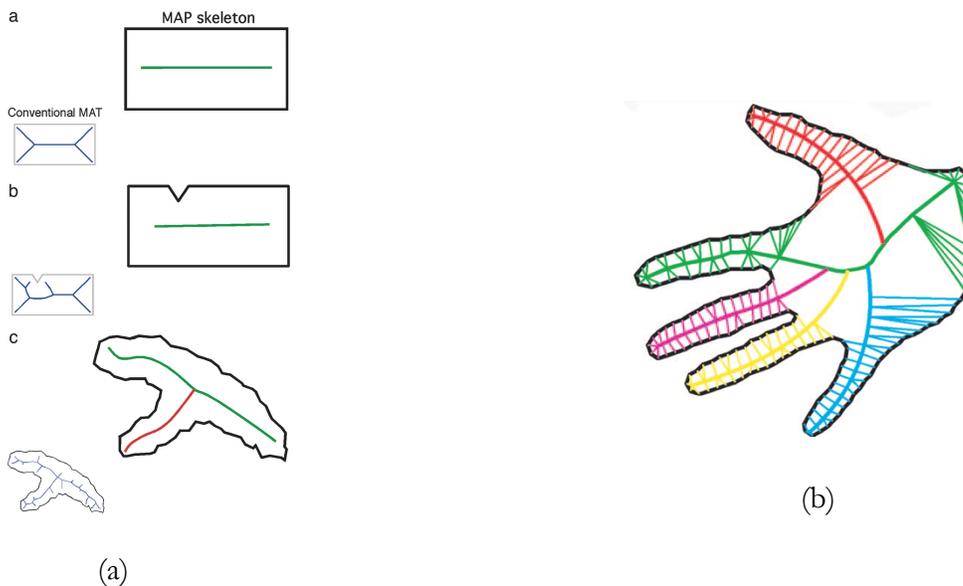
Object-centered representations encode the layout of an object's boundaries in relation to its intrinsic axes. However, objects have many intrinsic axes that might serve this function. Mirror-symmetric shapes have axes of bilateral symmetry. And all shapes have axes of elongation, defined either as the line that minimizes the sum of squared distances to all points in the shape (Cheng and Gallistel [2005]), or as the line connecting the two points on the shape that are farthest apart (Quinlan and Humphreys [1993]).

While symmetry and elongation axes can be employed in shape representation (Quinlan and Humphreys [1993]; Sekuler and Swimmer [2000]), I will primarily focus on a different axis whose significance is steadily gaining support. The *medial axis* comprises the set of points within a shape's interior that have two or more nearest neighbors on the shape's boundary (Blum [1973]; Kimia [2003]). It resembles a "skeleton" from which the shape is "grown." The medial axis of a complex shape consists of multiple *branches*, which often correspond to the shape's perceived part structure, wherein distinct parts are associated with distinct branches.

⁷ Although see Buxbaum et al. ([1996]) and Filimon ([2015]) for reservations about whether object-based neglect requires object-centered representation.

Unfortunately, the *raw* medial axis is highly sensitive to small contour perturbations.

Introducing a small notch to an object’s contour produces additional branches to its medial axis that don’t correspond to its perceived part structure (see figs. 5a, 5c). To overcome this difficulty, various authors have proposed “pruned” or “smoothed” variants of the medial axis (Shaked and Bruckstein [1998]; Ayzenberg *et al.* [2019]). Thus, Feldman and Singh ([2006]) impose a Bayesian prior over skeletons that assigns higher probability to smoother axes and penalizes extra branches. Their likelihood function expresses how well each candidate skeleton “explains” the shape’s boundary points.⁸ The model selects the skeleton which maximizes the product of prior and likelihood. Importantly, it aligns well with perceived part structure. Thus, in figure 5b separate axis branches correspond to separate fingers of the hand. I’ll use *skeleton* to denote this sort of pruned medial axis whose branches accord with perceived part structure.



⁸ Specifically, each axis point is construed as “sprouting” a rib perpendicular to the axis. The lengths of the ribs exhibit probabilistic noise, inducing a distribution over possible boundary points conditional on a given axis point. Boundary points are explained better when they receive higher probability in this distribution.

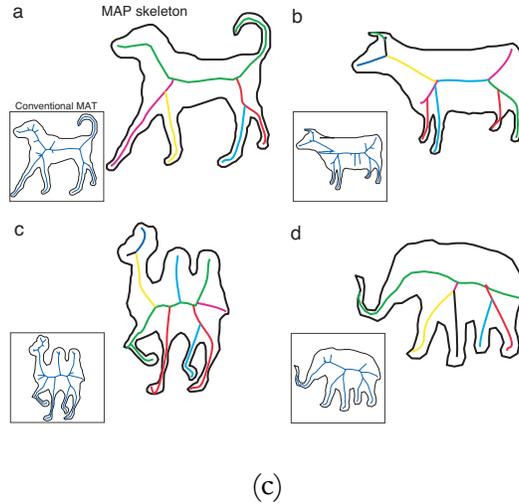


Figure 5. (a) Left column shows the “raw” medial axis, while right column shows the results of Feldman and Singh’s ([2006]) Bayesian smoothing. (b) Skeleton of the human hand, together with “ribs” linking points on the axes with the boundary points they explain. (c) Skeletons of some nonhuman animals (with contrasting raw medial axes). *Source:* Feldman and Singh ([2006]). Copyright (2006) National Academy of Sciences.

The *format* of a representational system consists in principles governing the relationship between the syntactic properties of representations belonging to the system and the contents that those representations represent. For example, certain theories of *analog* format classify a system as analog when there is a monotonic mapping from syntactic magnitudes of representations belonging to the system to the magnitudes that they represent (Beck [2019]). I now consider the format of skeletal representations, highlighting three features.

First, skeletal representations possess a *tree* format (fig. 6; Siddiqi *et al.* [1999]; Feldman and Singh [2006]; Feldman *et al.* [2013]; Green [2019]). A shape’s skeleton is organized into a root branch, followed by children, grandchildren, etc. The root branch belongs to the main “body” of the shape, while its children belong to parts protruding from the main body. Grandchildren offshoot from the children, and so on. For a human body, the root branch might belong to the torso, children to upper limbs, and later descendants to lower limbs. The syntactic structure of a skeletal

representation mirrors these relations among parts of the shape it represents. The representation is organized as a tree composed of discrete *nodes*, each of which represents some part of the shape. A complex shape is encoded compositionally, by concatenating these nodes, which represent individual parts, with *edges* representing relations between parts.

Nodes higher in the tree represent parts whose axes are ancestors of the axes of parts represented by nodes lower in the tree. Each node is complex, representing multiple geometrical properties of the part it characterizes: specifically, the structure of the part's axis and the structure of the boundaries (contours or surfaces) surrounding the axis. Furthermore, while the presence of an edge simply encodes a branching relation between parts (a relation that holds between parts P1 and P2 iff P2's axis branches from P1's axis), I conjecture that edges are *tagged* with information about spatial relations between parts—for example, the point along the parent's axis at which the child's axis branches, and angles between the axes.

Second, skeletal representations *extract* properties of axes. They possess separable constituents that represent the structure of axes and nothing else (and nothing more specific).⁹ By analogy, the sentence “the red cup is on the blue table” possesses separable constituents, “red” and “blue,” that represent the properties *red* and *blue*, and nothing else. These constituents are “separable” because they can remain unaltered while other constituents are discarded or modified (see: “the red ball abuts the blue box”). Extraction is not unique to sentences. A photograph of a red cup on a table contains a constituent (a spatial part) depicting the cup, which represents fine-grained information about the cup's color (some determinate shade of red), and can remain unaltered while other parts of the picture are modified. However, it does not extract the property *red*, since the same constituent also encodes more specific color properties (Kulvicki [2007], [2015]).

⁹ This notion of extraction derives from Dretske ([1981], ch. 7; see also Matthen [2005], ch. 3; Kulvicki [2007], [2015]).

A node of a skeletal representation-type contains a separable constituent that encodes the structure of a part’s axis (e.g., the curvature at each segment of its axis) and nothing else about the shape of that part.¹⁰ As such, skeletal representations explicitly encode commonalities between objects that share the same skeleton but differ in other ways. The properties of the outer contours or surfaces surrounding the part’s axis—which I’ll call the “boundary” structure of that part—are specified relative to its axis. For example, the representation might combine a representation of axis structure with a representation of the average width of the part, or perhaps with more detailed representations of the distance to the boundary from each point on the axis (compare Blum’s [1973] “sym-function”).

Third, beyond extracting the metric structure of axes and boundaries, I suggest that skeletal representations extract some of their *abstract* geometrical properties, representing these properties via separable constituents. For instance, in addition to its curvature at each segment, the axis may be attributed categorical properties like *straight* or *curved* (Biederman [1987], [2000]; Amir *et al.* [2012]). Likewise, the boundaries around the axis may be represented as simply *parallel*, *tapered*, or *symmetric* (Amir *et al.* [2012]; Hafri *et al.* [2023], [forthcoming]).

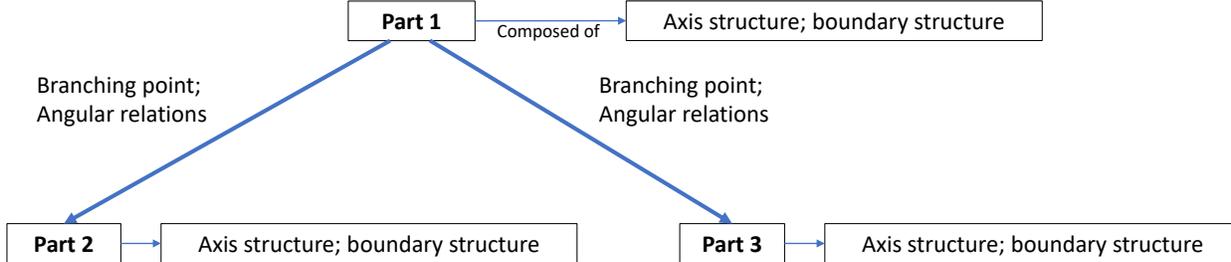


Figure 6. Structure of a skeletal representation-type. Notice that nodes within the tree are complex. For example, the node “Part 1” is *composed of* a representation of part 1’s axis structure combined with a representation of its boundary structure.

¹⁰ The constituents that represent part axes are not *primitive* constituents. They are complex constituents of a larger skeletal representation-type.

Notice that “skeletal” representations encode more than just the skeleton. They also encode boundary structures of parts in relation to their axes and additional spatial relations between parts or axes. The label “skeletal representation” simply connotes the central role that skeletal axes play in this system—enabling an object-centered specification of boundary structure.

There is copious evidence for skeletal representations, much of which will be discussed in section 4. Here are three preliminary lines of support:

First, medial axis structure influences low-level visual phenomena like contrast sensitivity (Kovács and Julesz [1994]; Kovács *et al.* [1998]) and texture segregation. For example, subjects are better at segregating texture-defined shapes from their backgrounds when the texture elements within the shape are oriented parallel to its medial axis (Harrison and Feldman [2009]).

Second, when people are asked to tap a shape wherever they like, taps tend to cluster around medial axes in ways that cannot be explained by alternative intrinsic axes (Firestone and Scholl [2014]). Moreover, Ayzenberg *et al.* ([2019]) found that the tapping patterns were best explained by a pruned model: taps clustered around axis branches marking significant structural aspects of the shape, but not branches reflecting contour noise, supporting the psychological reality of pruned medial axes.

Third, skeletal shape *complexity* influences various psychological processes. Objects with more complex skeletons (with higher curvature or more branches) are more quickly identified among objects with simpler skeletons than vice versa (Sun and Firestone [2021]). Skeletal complexity also influences the detection of closed shapes amidst noise (Wilder *et al.* [2016]), and even aesthetic preferences—wherein shapes of intermediate complexity are more appealing than either very simple or complex shapes (Sun and Firestone [2022]).

Note, finally, that skeletal representation presupposes other capacities equally critical to shape perception. Gestalt processes of figure-ground organization, grouping, and contour integration help define the “units” to which shape properties are attributed (Palmer and Rock [1994]; Feldman [2007]; Wagemans *et al.* [2012]; Lande [2023]). Moreover, while a virtue of the Feldman/Singh model is that it predicts the perceived part structure of objects, the perception of parthood is itself a rich capacity that follows known organizational principles—e.g., minima and short-cut rules (Hoffman and Richards [1984]; Singh *et al.* [1999]). Such principles might independently contribute to shape representation, perhaps supplementing computations of skeletal axes in specifying part structure.

4 The Varieties of Shape Constancy

This section argues that skeletal representations underlie four varieties of robust constancy: constancies for *determinate shape*, *compositional structure*, *skeletal shape*, and *skeletal topology*. More abstract constancies are exercised through changes that “peel away” further constituents from skeletal representations. Table 1 lists these constancies alongside the component features of the complex properties they target. While they may be used with different meanings elsewhere, I intend the labels “determinate shape,” “compositional structure,” “skeletal shape,” and “skeletal topology” to denote just those properties described in the right-hand column.

Variety of Constancy	Complex Property the Constancy Targets
Constancy for determinate shape	Axis structures, boundary structures, axis branching points, angular relations between axes
Constancy for compositional structure	Axis structures, boundary structures, axis branching points
Constancy for skeletal shape	Axis structures, axis branching points
Constancy for skeletal topology	Number of axes, descendancy relations among axes

Table 1. Varieties of shape constancy (left) and the complex properties they target (right), described in terms of the features that compose them.

Pluralism about shape constancy is the view that our perceptual systems exhibit robust constancy with respect to shape properties at varying levels of abstraction. This view is distinct from *pluralism about shape representation*, which holds that our perceptual systems employ multiple systems of shape representation—e.g., involving different formats or reference frames. For instance, Hummel ([2013]) maintains that perception employs both template-like systems of shape representation (Bülthoff and Edelman [1992]) and geon-based systems (Biederman [1987]), utilizing the former for unattended objects and the latter for attended objects (Stankiewicz *et al.* [1998]; Thoma *et al.* [2004]; although see Guggenmos *et al.* [2015]).

Pluralism about shape constancy is compatible with, but doesn't entail, pluralism about shape representation. A perceptual system might employ distinct systems for representing determinate shape properties and abstract, determinable properties. If both systems generate representation-types that remain invariant through proximal variation, then the perceptual system would exhibit multiple constancies *by way of* multiple representational systems. However, a *single* representational system might facilitate shape constancies at multiple levels of abstraction if it explicitly encodes both more and less abstract shape properties. I suggest that skeletal representations do this. They assign separable constituents to shape properties of varying

abstractness. Because these constituents remain invariant through different kinds of variation, skeletal representations underlie constancies for both more and less abstract shape properties.

I am also inclined toward pluralism about systems of shape representation. For instance, I believe the visual system produces both viewer-centered and object-centered shape representations. However, one needn't accept this view to accept pluralism about shape constancy.

4.1 Determinate shape

The first variety of shape constancy is constancy for *determinate shape*. An object's determinate shape consists in the intrinsic shapes of its parts (their axis and boundary structures) together with the angular relations between them (e.g., the angle between a person's arm and their torso). Determinate shape constancy is exercised when a perceiver or perceptual system tokens the *same overall skeletal representation-type* despite relevant variation in proximal stimulation. Paradigm cases include shifting viewpoint on an object or seeing the object undergo a rigid translation or rotation.

Determinate shape constancy is a *robust* constancy. It requires representing an object's determinate shape via the same representation-type under different proximal conditions. Importantly, the hypothesis that we exhibit determinate shape constancy explains similarities and differences among shape experiences that are not easily explainable otherwise. The contents of object-centered representations are inherited by perceptual experiences.

Section 2 assembled evidence for object-centered representations, but did not explore their fit with shape phenomenology. However, one might worry that the fit is poor. Shapes *appear different* at different orientations. Coins viewed head-on appear different from slanted coins. Because object-centered representations remain invariant through changes in orientation, they run afoul of this basic datum. Rather, shape representation-types must be tethered to specific perspectives on objects, so objects elicit different representation-types at different orientations.

However, perspectival variation in shape appearance does not impugn object-centered representations. First, even if shape phenomenology were *wholly* determined by viewer-centered representations, our perceptual systems might still form *unconscious* object-centered representations. Such representations might still feature in object recognition, categorization, or action guidance.

More importantly, inspection of the patterns of similarity among shape experiences actually offers compelling *support* for object-centered representations. Such patterns are naturally explained by the view that our visual systems produce *both* viewer-centered and object-centered representations that both contribute to shape phenomenology.

Consider the triangle and the square in figure 7, and consider the effect of 45°-rotation on our experiences of them. Plausibly, there is a larger change in the square's appearance than the triangle's (Peacocke [1992], pp. 76-77; Macpherson [2006]). While the triangles appear different, the squares appear *more* different. The hypothesis that experience reflects both viewer-centered and object-centered representations can explain this datum. Viewer-centered representations ground aspects of appearance that vary in both cases, while object-centered representations ground aspects that vary *only* in the square case. When an object has a single salient axis, that axis will be used to encode its shape regardless of orientation. Because the triangle has only one salient axis (its symmetry axis), it elicits the same object-centered representation-type at both orientations. However, when an object has multiple salient axes, perception may prioritize the axis nearest to vertical in viewer-centered coordinates. Because the square has multiple salient axes (e.g., symmetry axes bisecting both its corners and its sides), it elicits different object-centered representation-types at different orientations. Consistent with this, isosceles triangles visually prime their 45°-rotated counterparts, while squares do not (Humphreys and Quinlan [1988]).

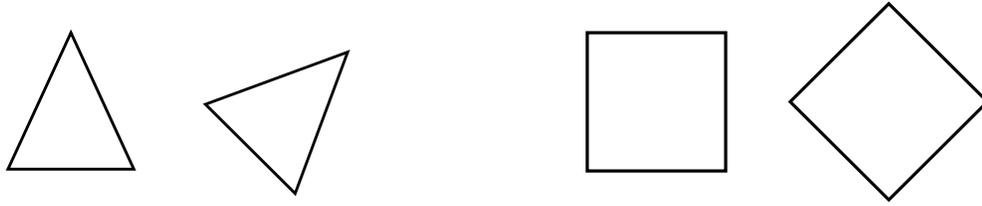


Figure 7. An isosceles triangle and a square/diamond.

I suggest, then, that we exhibit determinate shape constancy for the triangle but not the square, and this difference is phenomenologically manifest.

A study by Baker and Kellman ([2018]) provides further evidence for determinate shape constancy. Participants briefly saw a shape composed of black and white dots followed by a mask (fig. 8). Afterward, another shape appeared, which either matched the first or was slightly deformed, and subjects reported whether it was the same or different. On same-shape trials, the second shape was either at the same position and scale as the first, or was rotated or rescaled. Accuracy gradually improved with longer exposure to the first shape: participants were at chance for 30 ms exposures, but above chance at 50 ms or longer. Performance flattened for exposures of 110 ms or longer. Critically, this pattern held regardless of whether the shape underwent an orientation/size change. Baker and Kellman ([2018], p. 1300) conclude: “That results were similar across differing transformation types...suggests that a common abstract shape representation was used in the task.” In other words, by around 110 ms, the visual system tokened a representation-type that was invariant to changes in picture-plane orientation (see also Quinlan and Allen [2018]).

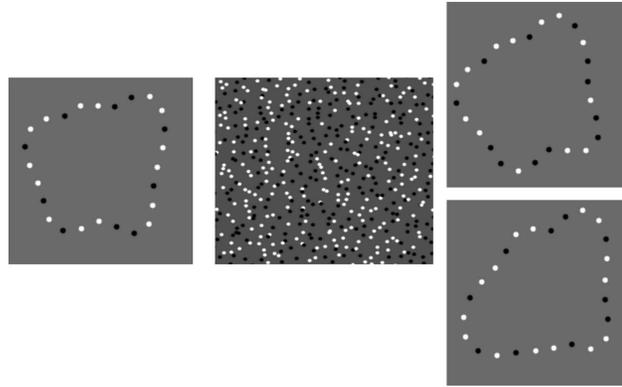


Figure 8. Trial structure in Baker and Kellman ([2018]). Top right shape is a rotated version of the left shape without any shape change. Bottom right shape involves both rotation and shape change. Reprinted with permission from (Baker and Kellman [2018]).

Thus, there is evidence that the visual system generates representations of determinate shape that remain stable through relevant changes in retinal projection. While the foregoing evidence concerns constancy through picture-plane rotations, determinate shape constancy may also be achieved through rotations in depth, provided the object has readily identifiable part structure and the same parts are visible throughout the rotation (Biederman and Gerhardstein [1993]; Biederman and Bar [1999]; Pizlo and Stevenson [1999]). Constancy through depth rotation is significantly aided by stereoscopic depth cues (Bennett and Vuong [2006]; Oliver *et al.* [2018]; although see Pizlo *et al.* [2008]).

Nonetheless, it is questionable how often determinate shape constancy is achieved. Human perception exhibits notable biases, limiting the situations in which perfect determinate shape constancy is possible. A given interval appears shorter when oriented in the sagittal plane (extending away in depth) than when oriented in the frontal plane (perpendicular to the line of sight) (Wagner [1985]; Loomis *et al.* [2002]; Wagner and Gambino [2016]). Likewise, volumetric shapes often appear compressed along the viewer's depth axis (Tittle *et al.* [1995]; Todd and Norman [2003]; Todd [2004]; Hill [2020]; Yu *et al.* [2021]). Rotating an object in depth alters its apparent determinate shape, since this changes the object's orientation with respect to the viewer's depth axis, and so the

direction along which it is perceptually compressed. Thus, determinate shape constancy is typically imperfect through depth rotations.

I regard determinate shape constancy as an ideal that is often only approximated. Fortunately, skeletal representations offer resources for understanding approximations to determinate shape constancy. Recall the third feature of skeletal representations highlighted earlier: They extract abstract shape properties, encoding them by separable, dedicated constituents. One relevant abstract property is *affine* shape (Todd [2004]; Bennett [2012]; Warren [2012]; Green [2017]; Todd and Petrov [2022]). Unlike *metric* shape properties, which are preserved only under similarity transformations (rigid transformations and uniform scaling), affine properties are preserved under all affine transformations—roughly, transformations preserving collinearity, parallelism, and ratios of lengths of parallel line segments. These include compressing, stretching, or shearing along a direction.¹¹ Examples of affine properties include *parallelogram*, *ellipse*, or *triangle*. Because compression along a direction—say, the depth axis—preserves affine shape, distortion of perceived metric shape is compatible with accurate perception of affine shape. Someone might perceptually represent a slanted square as a compressed rectangle (getting its metric shape wrong), but also simply as a parallelogram (getting its affine shape right).

There is evidence that our visual systems represent affine shape. Shapes are more discriminable when they differ in affine shape than when they differ only in metric properties, even when they differ equally in other respects (e.g., in pixel-level overlap) (Todd *et al.* [1998], [2014]; Amir *et al.* [2012]; Green [2017]). Moreover, perceptually-based judgments about an object's affine

¹¹ Two external objects can be affine-equivalent even if their retinal images are not. A head-on square and a slanted oblong rectangle are affine-equivalent. However, the square projects an image with two pairs of parallel sides while the rectangle projects a trapezoid with non-parallel sides, so their images are not affine-equivalent. I suggest that perception represents affine properties of external objects, not their projected images.

properties are both more accurate and more invariant across changes in viewpoint than judgments about metric properties (Phillips *et al.* [2003]).

Perceptual representation of affine shape is not *always* perfect. A rectangular object might appear trapezoidal when presented at an extreme slant. This would mark a misperception of affine shape (since rectangles and trapezoids are not affine-equivalent). The crucial point is that perception of affine shape is significantly *more* accurate and viewpoint-invariant than perception of metric shape.

Thus, when an object is rotated in depth, constituents of its representation encoding *metric* properties often do *not* remain invariant, but constituents encoding abstract properties might. More generally, while perfect constancy for determinate metric shape may be relatively rare, constancy for affine abstractions from determinate shape may be fairly common.

I conjecture that the remaining forms of constancy are also imperfect, particularly under depth rotation. However, the foregoing strategy can be generalized. We do not achieve *perfect* constancy for the relevant property under depth rotation, but may achieve near-perfect constancy for a natural (e.g., affine) abstraction from it. Systematic biases in spatial perception do not preclude shape constancy, but demand a nuanced understanding of it (Todd [2004]; Bennett [2012]; Hatfield [2016]).

4.2 Compositional structure

In a perceptive passage, J. L. Austin ([1962], p. 67) observes that when we seek to identify the “real” shape of a cat, multiple answers present themselves:

[W]hat is the real shape of a cat? Does its real shape change whenever it moves? If not, in what posture *is* its real shape on display? Furthermore, is its real shape such as to be fairly smooth-outlined, or must it be finely enough serrated to take account of each hair? It is pretty obvious that there is *no* answer to these questions—no rules according to which, no procedure by which, answers are to be determined.

Austin mentions two challenges in specifying the cat’s real shape. First, it adopts various postures, which affect how its limbs are oriented relative to its body. Either (i) we say that its real shape changes whenever its posture changes, or (ii) we choose some specific posture as displaying its real shape, but any choice seems arbitrary. Second, the cat’s outline can be characterized at many levels of specificity. The finest description incorporates every hair, while a coarser description would treat its outline as a smooth curve. Austin could have expanded the second challenge to pose a parallel issue to the first. The precise arrangement of hairs on the cat’s body is volatile, changing whenever the cat runs quickly or gets wet. Assuming a fine-grained description, then either (i) the cat’s real shape changes whenever it gets wet, or (ii) we choose some level of wetness in which its real shape is on display, but any choice seems arbitrary.

This subsection analyzes a form of constancy exercised through changes in the cat’s posture. The next analyzes a form exercised through changes in its fine-grained outline. At certain levels of analysis, the cat’s shape remains unaltered through these shifts, and we perceptually represent properties at these levels, enabling us to reliably reidentify the cat despite postural changes. Thus, my answer to Austin is that the cat has multiple “real” shapes. Some are volatile, others are resilient, and we perceptually represent both. By analogy, stretching a square into an oblong rectangle alters its determinate shape while preserving certain more abstract shape properties (e.g., *quadrilateral*). Just as it makes no sense to ask whether the “real” shape of the square changes or survives the stretching, as if there were a single answer, it makes no sense to ask whether the cat’s real shape changes or remains constant when it shifts posture.

An object’s *compositional structure* consists in its decomposition into parts, the determinate intrinsic shapes of its parts, and the “joints” at which its parts are connected—construed as the point along the parent part’s axis at which the child’s axis branches (see Green [2019], [2022]). Compositional structure abstracts away from the precise angular relations between parts. As a

person walks, her limbs rotate with respect to her torso, so the angular relations between these parts change. Nonetheless, the determinate intrinsic shapes of the parts remain approximately unchanged, as do their joint locations. The notion of compositional structure can be precisified using skeletal representations. Figure 8 crosses out elements of the representation encoding features that are inessential to compositional structure. The remaining elements encode properties—namely, axis and boundary structures of parts and the points where branches sprout from their parents—that together comprise compositional structure.

Compositional structure is distinct from determinate shape, although it is an abstraction from determinate shape (just as *quadrilateral* is an abstraction from *square*). An object’s determinate shape depends on both the determinate shapes of its parts and the angles between them, while compositional structure abstracts away from the latter feature. Any objects that share determinate shape also share compositional structure, but the converse does not hold. Moreover, compositional structure remains invariant under a wider range of transformations than determinate shape.

Compositional structure is not a property traditionally studied in geometry. Indeed, it is geometrically “messy” insofar as it is not preserved under the sorts of transformations standardly used to individuate shape properties, which coincide with different general geometries (e.g., Euclidean, affine, or projective geometry) (Todd and Petrov [2022]).¹² The perceptual significance of compositional structure derives not from pure mathematics, but from contingent facts about our environments: objects often change in ways that preserve compositional structure but not determinate shape, making it adaptive for our perceptual systems to extract compositional structure in order to reidentify objects across such changes.

¹² While compositional structure is not itself a metric property, it is partially composed of metric properties—namely, the metric shapes of an object’s parts. It consists of both metric properties and pure topological properties (viz., the points of contact between parent and child axes). Thus, compositional structure spans traditional levels of abstraction in geometry.

Skeletal representations facilitate robust constancy for compositional structure, or *structure constancy* for short. Structure constancy is exercised when a perceiver or perceptual system perceptually represents compositional structure via tokens of the same representation-type despite relevant variation in proximal conditions. The representation-type here is a separable complex constituent possessed by many skeletal representation-types—a constituent consisting of just those elements that remain uncrossed in figure 9. Because skeletal representations extract the axis and boundary structures of parts, representing them independently from angular relations between parts, structure constancy can be exercised through changes that “peel away” those constituents of the representation that encode angular relations, while determinate shape constancy cannot.

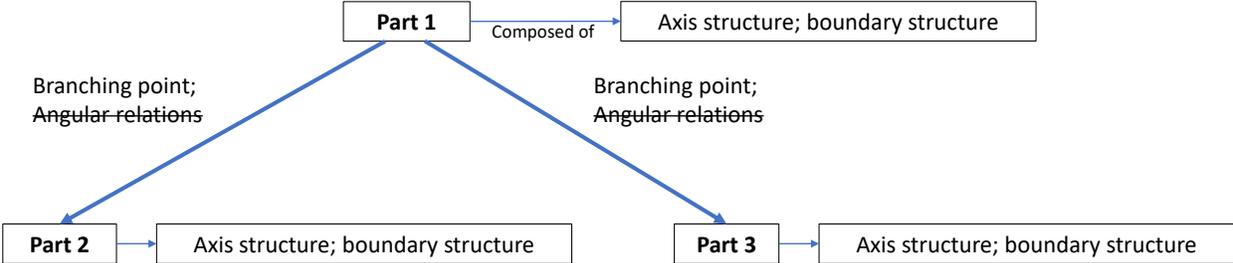


Figure 9. Uncrossed elements represent aspects of compositional structure. Crossed-out elements do not.

4.2.1 Compositional structure in perceptual experience

There is an introspective case that perceptual experience represents compositional structure. Objects that share compositional structure are, other things being equal, more similar in apparent shape than those that do not, even when both pairs are equally alike in other respects.

Consider a *base object*, A, and two *test objects*, B and C. B and C are equally different from A (or perhaps B is more different) along low-level measures such as the amount of pixel-level overlap between them. However, objects A and B share compositional structure while A and C do not. In such cases, introspection suggests that A and B appear more similar vis-à-vis shape than A and C.

For example, figures 10a and 10b share compositional structure, while figures 10c and 10d differ from them. In figure 10c, the smaller part’s axis offshoots from a lower point along the larger part’s axis (altering joint location), while in figure 10d, the shape of the smaller part has changed from rectangular to triangular.^{13, 14} Importantly, explicit similarity judgments exhibit the same pattern: shapes that share compositional structure are systematically rated as more similar than shapes that differ either in part shapes or joint locations (Barenholtz and Tarr [2008]).

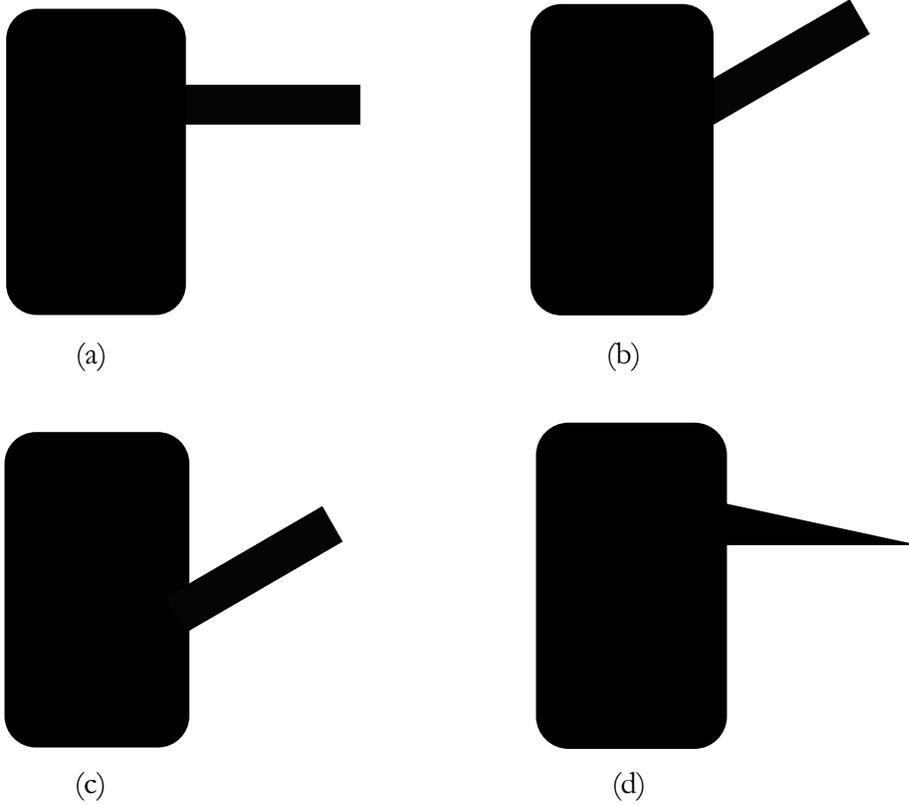


Figure 10. (a) and (b) share compositional structure, while (c) and (d) both differ from them.

¹³ Figures 10b and 10c are equally different from figure 10a in lower-level respects because the same amount of rotation has been applied to the “arm.” The only difference is that, in figure 10b, the arm has rotated about its joint with the “body,” while in figure 10c it has rotated about its endpoint.

¹⁴ These patterns would be less salient if figures 10a and 10d were made more alike by, say, making the “arm” in figure 10d a trapezoid. However, the figures would still differ in compositional structure because their arms would differ in determinate shape (namely, in the boundary structures surrounding their axes). If experience represents compositional structure, then the figures *should* appear more similar under these conditions. Since compositional structure consists partly in the determinate shapes of an object’s parts, objects should appear more similar as their parts’ determinate shapes become more alike.

The hypothesis that visual experience represents compositional structure offers an attractive explanation for this pattern of apparent shape similarity. While all four objects are experienced as differing in determinate shape, figures 10a and 10b are experienced as sharing a property that figures 10a and 10c (and 10a and 10d) are not. Thus, figures 10a and 10b are more alike in apparent shape.

Return to the cat. As the cat walks, its compositional structure often (though not always) remains stable. Its limbs rotate about its torso, but the determinate shapes of the limbs and the points at which its body parts are connected remain approximately unchanged, due to its anatomical skeleton. So, if perceptual experience represents compositional structure, then it represents a feature that endures through non-rigid articulations of the cat.

If visual experiences represent compositional structure, then they plausibly inherit some of their contents from skeletal representations. After all, skeletal representations *extract* the features constitutive of compositional structure, representing them separately from features that are not constitutive of its compositional structure. Not just any system of shape representation predicts the above patterns of shape similarity. For example, template-based models where shape is encoded as a vector of local feature locations within a viewer-centered reference frame (Ullman and Basri [1991]; Bülthoff and Edelman [1992]; Edelman and Duvdevani-Bar [1997]) do not predict them because shapes that share compositional structure are deemed more similar than those that do not even when local feature change is equated in both cases (Barenholtz and Tarr [2008]; Lowet *et al.* [2018]; see also Stankiewicz and Hummel [1996]).

Buccella ([2021]) objects to the distinction between determinate shape constancy and structure constancy. Buccella argues that if we countenance this distinction, then we are committed to the presence of two constancy *mechanisms*—one activated by rigid objects (for determinate shape), the other by non-rigid objects in motion (for compositional structure). Accordingly, we face the

unenviable task of explaining “how the retinal image gets disambiguated and...how the visual system ‘decides’ which constancy mechanism to activate in each case” (Buccella [2021], p. 15).

However, a reply is available. While we must differentiate determinate shape constancy from structure constancy, we needn’t hold that these constancies depend on distinct constancy *mechanisms*. An attractive feature of the account advocated here is that it illustrates how a *single* constancy mechanism with syntactically complex outputs may ground *multiple* constancy *capacities*. The mechanism produces skeletal representations based on proximal cues. Because of their syntactic structure, these representations enable constancies for shape properties at different levels of abstraction. A common representation-type facilitates *both* determinate shape constancy and structure constancy, along with the others to come. We needn’t posit multiple constancy mechanisms (though multiple mechanisms may exist), and thus needn’t posit a stage where the system decides whether to engage mechanisms for determinate shape or structure constancy.

4.2.2 Further support for structure constancy

The case for structure constancy is bolstered by two further data: (i) the visual system represents the shapes of parts independently of one another, and (ii) the visual system differentiates changes in the relations between parts that alter their joint locations from changes that merely alter angles between parts.

Regarding (i): A raft of data indicates that the visual system divides objects into parts and that represents their shapes separately from one another, allowing us to appreciate commonalities between objects composed of same-shaped parts in different arrangements. Cacciamani et al. ([2014]) found that visual priming transfers between objects with the same parts in different arrangements. When shown an ambiguous figure-ground display, participants were more likely to see a region as figure if they were primed with an object containing the same parts as the region but

arranged differently. Second, the ability to represent part shapes can be preserved while the ability to represent their relations is impaired. For example, certain patients with integrative agnosia reliably distinguish complex objects composed of differently shaped parts (e.g., a cube atop a rod versus a cone atop a sphere), but cannot distinguish objects that differ only in how their parts are arranged (Behrmann *et al.* [2006]).

These data are explainable on the hypothesis that the visual system produces shape representations containing discrete constituents encoding the shapes of midsized parts. This format makes explicit the commonalities between objects composed of the same parts in different arrangements, since their representations have discrete, separable constituents in common. Moreover, the mechanism that produces these representations may be partially damaged without going fully offline. It may retain the ability to form representations of part shapes (nodes) while losing the ability to form representations of their relations (edges).

Regarding (ii): Changes modifying the joint locations between parts are more salient than equal-magnitude changes that merely modify angles between parts. Barenholtz and Tarr ([2008]) presented a base shape followed by two test shapes, and subjects judged which appeared more similar to the base. The objects consisted of a main “body” and a “limb.” In one test shape, the limb was rotated about its joint (preserving joint location), while in the other it was rotated about its endpoint (altering joint location and thus compositional structure). Subjects reliably rated the former shape as more similar to the base. More recently, Lowet *et al.* ([2018]) found that unfamiliar shapes composed of line segments were more easily discriminable when joint locations between line segments were altered than when only the angles between them changed.

Skeletal representations can explain these results. Such representations prioritize joint locations between parts because they extract this information. Separable constituents encode the point along a parent axis at which its child axis sprouts, and this information is represented with

high fidelity. In contrast, information about angles between parts is represented independently, perhaps in a more coarse-grained way.

Thus, our visual systems recover a type of property that cats, staplers, and other deformable objects retain through changes in posture. This capacity grounds a form of constancy distinct from constancy for determinate shape. Our visual systems represent compositional structure via tokens of the same representation-type despite relevant proximal variation. Moreover, compositional structure is represented in experience, governing comparative similarities among shape experiences.

4.3 Skeletal shape

I turn to *skeletal shape constancy*. An object’s skeletal shape consists in the shapes of its parts’ axes together with their branching points. Objects can share the same skeletal shape despite differing in their determinate shapes (and compositional structures) due to differences in the boundaries (contours or surfaces) surrounding the axes. The features constitutive of skeletal shape are shown in figure 11. Skeletal shape is represented by a complex constituent of a skeletal representation-type consisting of (i) constituents of nodes that encode structural features (length, curvature, etc.) of part axes, and (ii) tags on the edges between nodes that encode the point(s) along each parent axis where its children offshoot. Elements representing features inessential to skeletal shape have been crossed out.

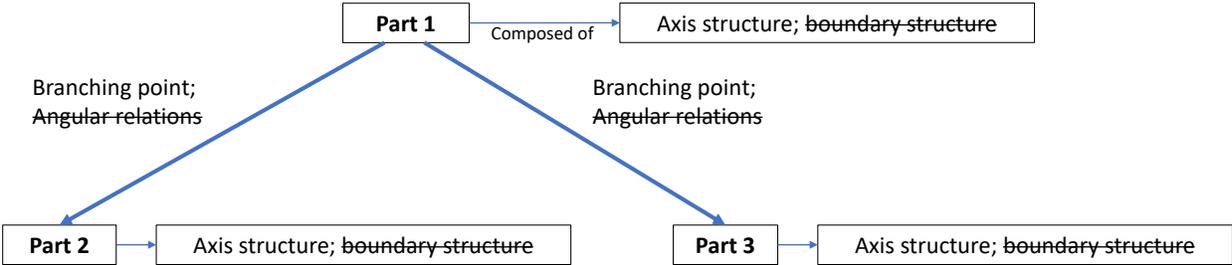


Figure 11. Uncrossed elements represent aspects of skeletal shape. Crossed-out elements do not.

Skeletal shape is distinct from compositional structure, though it is an abstraction from compositional structure. An object's compositional structure depends not just on the shapes of its component axes, but also on the structure of the boundaries (contours or surfaces) surrounding the axes. Skeletal shape abstracts away from the latter feature—though, like compositional structure, it also abstracts away from the angles between connected parts. Thus, any two objects that share compositional structure also share skeletal shape, but the converse is not true.

Skeletal shape constancy is exercised when a perceiver or perceptual system perceptually represents skeletal shape via tokens of the same representation-type despite variation in proximal conditions. The relevant representation-type is a constituent possessed by many skeletal representation-types. Because skeletal representations encode axis structure independently from boundary structure, skeletal shape constancy can be exercised through changes that “peel away” constituents of the representation that encode boundary structure, while constancies for compositional structure and determinate shape cannot.

4.3.1 Skeletal shape in perceptual experience

My argument for the experiential representation of skeletal shape parallels that given for compositional structure.

Consider figure 12. While all three shapes differ from one another in both determinate shape and compositional structure, there is a salient respect in which the top shape appears more similar to the bottom-left shape than the bottom-right. The hypothesis that perceptual experiences inherit some of their contents from skeletal representations explains these patterns of apparent shape similarity, while various alternative theories do not. For example, there is no easy way to explain these patterns by appeal to the viewer-centered coordinates of local boundary features. There is also no easy way to explain them by appeal to more qualitative aspects of boundaries (e.g., whether they

are “sharp” or “smooth,” etc.), since such aspects are, if anything, more similar between the objects that differ in skeletal shape (Ayzenberg and Lourenco [2019], pp. 6-7).

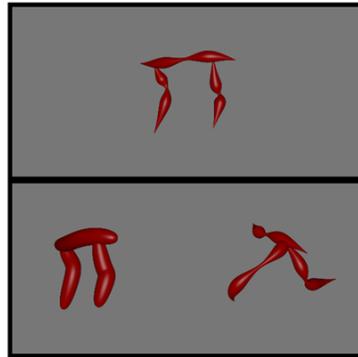


Figure 12. Similarity versus difference in skeletal shape. *Source:* Ayzenberg and Lourenco ([2019]). Reprinted under the Creative Commons Attribution 4.0 International License.

Ayzenberg and Lourenco ([2019]) gathered shape discrimination evidence that corroborates these introspective judgments. They presented a pair of objects like those shown in figure 12 at varying orientations in depth and asked participants to judge whether the objects were the same. Levels of discriminability were predicted by skeletal dissimilarity¹⁵: objects with more dissimilar skeletons tended to be easier to discriminate. Skeletal dissimilarity also explained unique variance in subjects’ discrimination performance that low-level image-based measures and contemporary convolutional networks like AlexNet could not explain (see also Destler *et al.* [2019]).¹⁶

Ayzenberg and Lourenco ([2022]) found these same patterns in human infants. 6- to 12-month-old infants were habituated to a novel shape, and then shown two new objects. Both differed in determinate shape from the habituation stimulus, but only one shared its skeletal shape. Infants dishabituated to the object with the new skeleton, but not to the one with the same skeleton. Thus,

¹⁵ Calculated as “the mean Euclidean distance between each point on one skeleton and the closest point on the second skeleton following maximal alignment” (Ayzenberg and Lourenco [2019], p. 2).

¹⁶ Baker *et al.* ([2020]) gather further evidence that convolutional neural networks prioritize texture over shape in classification (although see Hermann *et al.* [2020]).

there is evidence that skeletal shape is recovered in perception and broadcast to processes of shape discrimination and categorization.

Return to the cat. As it runs quickly or gets wet, its determinate shape changes. Its compositional structure changes too, since its parts adopt new determinate shapes. However, its skeletal shape remains virtually constant. If visual experience represents skeletal shape, then it represents an aspect of shape that endures through variation in the cat's intrinsic structure, and accounts for our sense of stability through these changes.

4.3.2 Further support for skeletal shape constancy

The perceptual representation of skeletal shape is supported by patterns of apparent shape similarity, discriminability, and categorization. However, one concern with this evidence is that the behavioral data might result from the recovery of skeletal shape in cognition, not perception. However, there is also evidence that canonical visual brain areas code for skeletal shape, bolstering the case that skeletal shape representations are perceptual.

First, Hung *et al.* ([2012]) found neurons in the inferotemporal cortex (IT) of macaques that were tuned to particular complex medial-axis structures. Their responses were more heavily driven by medial axis structure than by surface curvature. IT is generally regarded as the homologue of human lateral occipital cortex (LOC), which functions to represent shape for recognition (Grill-Spector *et al.* [2001]).

Similarly, Lescroart and Biederman ([2013]) found that BOLD activity patterns in areas V3 and LOC of humans were selective for medial axis structures across changes in orientation, independent of changes in the boundaries surrounding the axes.¹⁷ Ayzenberg *et al.* ([2022])

¹⁷ The claim is *not* that information about position/orientation is discarded at higher levels—it is not (DiCarlo and Maunsell [2003]). Rather, it is represented alongside object-centered information.

reinforced this conclusion by showing that skeletal shape similarity explained unique variance in the response patterns in areas V3 and LOC (see also Papale *et al.* [2020]; Ayzenberg and Behrmann [2022b]). Skeletal encoding again generalized across changes in surface form. Finally, they found that the response patterns in V3 and LOC predicted the discrimination data from Ayzenberg and Lourenco ([2019]), with similar responses correlated with lower discriminability.¹⁸

The fact that responses in visual brain areas are correlated with skeletal shape does not entail that skeletal shape is perceptually represented. Correlation is not sufficient for representation. The neuroscience should be interpreted as *complementing* the behavioral evidence, which suggests that skeletal shape guides visual discrimination, recognition, and categorization. The best overall conclusion, I contend, is that skeletal shape is perceptually represented and made available to these processes.

4.4 Skeletal topology

I turn finally to constancy with respect to *skeletal topology*. My case for this constancy is somewhat more speculative, but it is a natural extrapolation from the first three varieties of constancy.

Many objects are not just globally non-rigid; they are *piece-wise* non-rigid. Unlike the cat's limbs, which remain roughly the same shape as it moves, the parts of many objects undergo non-rigid deformations like bending or twisting—think of a snake, a fern, or a lizard's tail. This is true even of parts of the cat—its torso and tail deform non-rigidly when it walks or curls up to nap. Nonetheless, such changes seem *systematic*, in contrast to the deformation of, say, an amorphous blob.

¹⁸ Another question concerns how representations of skeletal shape in LOC are generated. Recent evidence suggests that dorsal-stream regions may encode the spatial arrangement of part axes and transmit this information to the ventral stream, where it is combined with ventrally encoded information about local features into a representation of overall shape (Ayzenberg and Behrmann [2022a], [2022b]). Thus, while LOC is a ventral region, computations of skeletal shape may not be confined to the ventral stream.

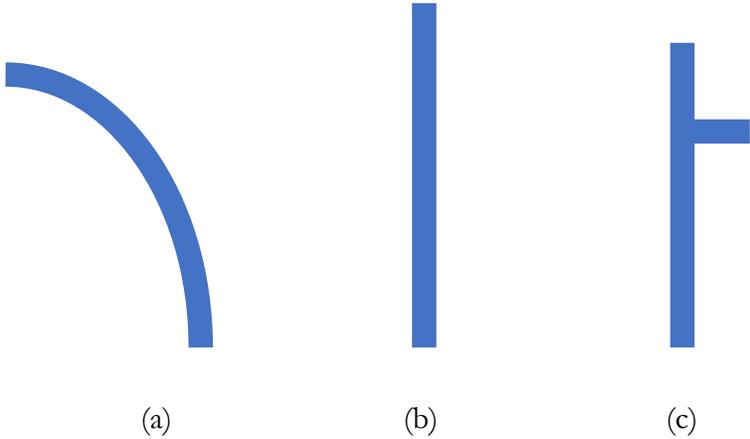
The foregoing changes alter the intrinsic shapes of an object’s parts—both their axes and their boundaries—and thus demand revisions to most constituents of a skeletal representation. Nevertheless, the changes leave the overall *topology* of ancestral relations among parts intact. More specifically, an object’s *skeletal topology* consists of (i) the number of axes in its skeleton, and (ii) the pattern of ancestral relations among these axes (compare Destler *et al.* [2023], p. 4). To make this notion more concrete, imagine the deformation of a lizard as its body bends and its tail curls. The lizard’s torso and tail adopt new shapes, and their skeletons deform non-rigidly (they become more curved). However, the lizard’s torso continues to appear as the “main” part of its body, and the number of parts seen as “descending” from its torso (tail, limbs, etc.) remains unchanged.

Skeletal topology is distinct from skeletal shape, though it is an abstraction from skeletal shape. Skeletal shape involves not just an object’s number of axes and their ancestral relations, but specific geometrical properties of each axis (for instance, length and curvature), and specific relations between axes (namely, the points along each parent axis at which its children offshoot). Thus, any two objects that share the same skeletal shape also share the same skeletal topology, but the converse is not the case.

Constancy for skeletal topology is exercised when we perceptually represent an object’s skeletal topology via tokens of the same representation-type despite variation in proximal conditions. Here, the relevant representation-type is the general tree structure, which is shared between skeletal representations that attribute different axis and boundary structures to each part. I conjecture that skeletal topology is unique among the abstract shape properties discussed here insofar as it is not encoded by a discrete, separable constituent that represents the property independently of other properties. Rather, skeletal topology is encoded by a *structural aspect* of the representation: namely, the number of nodes it contains, and the positions of edges between those nodes. I suggest that this structural aspect of the representation is accessible by consuming processes.

Another unique feature of skeletal topology is that it is a purely categorical property that does not exhibit continuous variation. While it is possible to measure comparative differences in skeletal topology (Sebastian and Kimia [2005]; Destler *et al.* [2023]), it is not true that whenever two objects differ in skeletal topology, there is always a third skeletal topology lying “between” them. If object A is a one-part object and object B is a two-part object consisting of a parent and a single child, no skeletal topology lies between them. Accordingly, skeletal topology differs from, say, skeletal shape, which depends on continuous properties of an object’s component axes (such as length or curvature) that can be manipulated parametrically (Ayzenberg and Lourenco [2019]). Skeletal representations thus have the virtue of explicitly encoding information about both continuous and categorical geometrical features.

Consider figures 13a-c. The central figure differs from both the left and right figures, but there is plausibly a respect in which the central figure appears more like the one on the left. I suggest that this difference is explained, at least in part, by the experiential representation of skeletal topology. The objects on the center and left both have one skeletal axis, while the object on the right has two axes. Figures 13d-13f illustrate that these patterns of apparent similarity cannot be easily explained in terms of pixel-level overlap between the shapes.



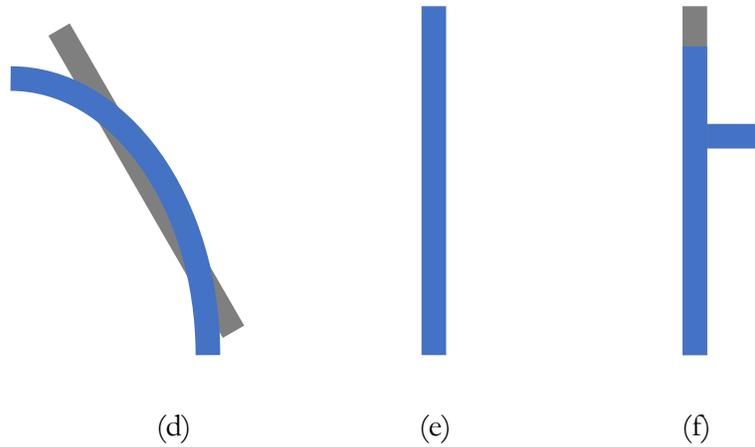


Figure 13. (a)-(c): The central shape differs from the right shape in skeletal topology, but shares skeletal topology with the left shape. (d)-(f): It is difficult to understand the respect of apparent similarity between figures (a) and (b) in terms of pixel-level overlap.

Within computer vision, Sebastian *et al.* ([2004]) developed an algorithm for shape recognition based on comparing the topology of graphs formed from medial-axis branches. All shapes with the same topological configuration of axis branches are treated as equivalent, and the “distance” between two shapes is measured by the number of topological changes to the skeleton needed to transform one shape’s graph into the other. They found that this approach outperformed algorithms based on boundary curvature in recognizing shapes through articulation of parts and partial occlusion (Sebastian and Kimia [2005]).

A recent study by Destler *et al.* ([2023]) suggests that skeletal topology also contributes to apparent shape similarity in humans. Participants were shown a grid of 36 novel shapes and asked to indicate which belonged to the same category as a reference shape. Categorization judgments were well-predicted by a model that compared the skeletal topologies of the objects. Objects were more often categorized together when fewer axis additions or deletions were needed to transition from one object’s skeletal topology to the other’s.

Further evidence derives from judgments of structural stability through non-rigid transformations—e.g., bending or twisting. Several studies have explored our ability to match

objects as being “the same” across such distortions (Spröte and Fleming [2016]; Schmidt and Fleming [2016]), and to infer the kind of distortion an object has undergone (Spröte *et al.* [2016]; Schmidt and Fleming [2018]; Fleming and Schmidt [2019]; Schmidt *et al.* [2019]).

Spröte and Fleming ([2016]) showed participants a test object at the center of a screen and four candidate objects at the four corners (fig. 14). Participants were asked to identify which (if any) candidate matched the test object. When there was a match, either the objects had the same determinate shape or one was a bent version of the other, and they were told to ignore bending when identifying matches. Importantly, participants identified matches in the bent condition at well-above-chance rates. It is worth noting that the objects in this study were unfamiliar and highly complex, and so weren’t matched perfectly even *without* a bend.

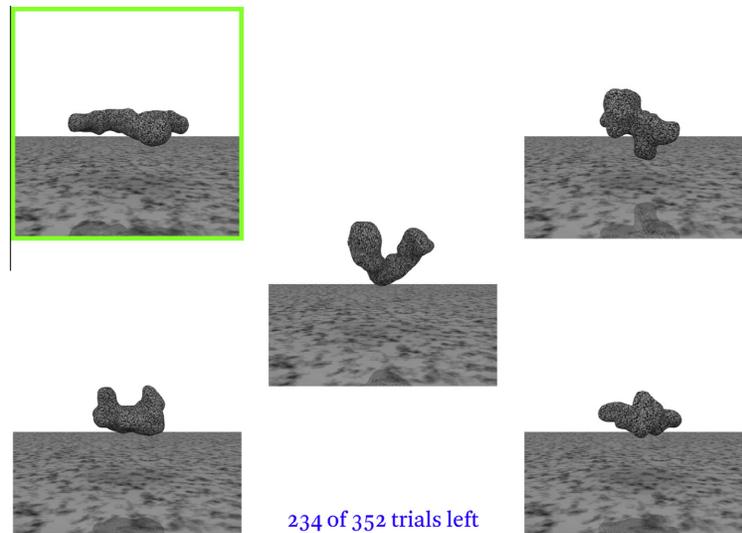


Figure 14. Example stimulus used by Spröte and Fleming ([2016]). The test object is at center, and the “unbent” version of it is at the upper left. Reprinted with permission from (Spröte and Fleming [2016]).

Perceptual representation of skeletal topology may facilitate the matching of shapes across non-rigid changes like bends—particularly when the shapes are unfamiliar. If shape representations make an object’s skeletal topology readily accessible to consuming processes, then it should be

possible to determine whether two objects have the same skeletal topology, assuming the objects are not too complex (viz., by assessing whether their skeletal representations possess the same tree structure). This would provide a basis for deciding whether two objects possess the same shape, modulo a simple non-rigid distortion.

One might object that there is a simpler explanation of our ability to match objects across bends: Perhaps such matches rely on correspondences in the ordering of local surface features like bumps, indentations, or corners. As an object is bent, the ordering of its convexities and concavities remains generally stable, and we can use this pattern to determine whether two objects are the same modulo a bend.

Participants do sometimes rely on local-feature correspondences when comparing objects across non-rigid distortions. Schmidt and Fleming ([2016]) presented shapes that differed by various non-rigid changes (shears, bends, etc.). A series of dots was shown on one shape, and participants had to move dots on the other shape to locations “corresponding” to the dots on the first shape. Their responses were highly consistent and well-predicted by a model which first matched salient landmarks between the shapes (for example, curvature extrema), then placed points at the same positions relative to corresponding landmarks. Perhaps Spröte and Fleming’s ([2016]) subjects also used local feature-matching to determine whether two shapes were bent versions of one another.

However, it is doubtful that local feature-matching fully explains shape matching across non-rigid distortions. For, two shapes with similar curvature extrema in the same order can look *very* different. Figure 15 shows an example modeled after (Sebastian and Kimia [2005], fig. 14). The two shapes have a similar sequence of curvature extrema. However, the figure on the right does not look like a bent version of the figure on the left. They look like completely different shapes. Plausibly, one reason they look so different is that they possess skeletal trees with different topology.



Figure 15. Two shapes that have similar orderings of curvature extrema but appear very different.

Matching shapes across nonrigid distortions may rely on both skeletal topology and local feature matching. Perhaps skeletal topology is used to determine whether two objects are *viable candidates* for correspondence—whether one *could* be a bent version of the other. Once this initial global matching is performed, we may turn to local feature matching to determine how to match smaller contour segments between the shapes.

4.5 Two concerns

Before closing, I consider two potential concerns with the foregoing approach.

First, while I've argued that abstract varieties of shape constancy are reflected in online perceptual experience, an alternative view would maintain that abstract shape properties are not represented in perception, but only later on—perhaps when encoding shapes into memory. However, two data speak in favor of the perceptual view. First, skeletal shape influences paradigmatically low-level perceptual processes, such as contrast sensitivity (Kovács and Julesz [1994]; Kovács *et al.* [1998]), texture segregation (Harrison and Feldman [2009]), and figure-ground segregation (Froyen *et al.* [2010]), suggesting that it is represented perceptually as well. Second, the perceptual view arguably fits better with the phenomenology of shape. When we look at the pairs of objects in figures 2-4 above, we don't merely *remember* them as bearing some structural similarity—the similarities seem immediately present in experience. Thus, while it is notoriously difficult to

resolve the boundary between perception and memory, I think we have grounds for placing skeletal representations, and the constancies they engender, on the perception side.

Second, one might worry that the account has been supported only by confirming its positive predictions. I identified several properties that skeletal representations extract, and adduced evidence that perception exhibits constancies for these properties. But this leaves *negative* predictions unexplored. Are there any possible forms of shape constancy that the skeletal approach predicts should be *absent* in human perception?

One problem with evaluating negative predictions in this context is that even if there are forms of constancy that *skeletal* representations cannot explain, other representations might explain them instead (recall that I am inclined toward pluralism about shape representation). Nonetheless, I now highlight a possible form of constancy that skeletal representations are unfit to explain, and which also turns out to be not very prominent in human perception.

While all viable systems of shape representation are compositional insofar as complex representations are composed from semantically significant constituents (Hummel [2000], [2013]), a signature of skeletal representations is that their constituents (nodes) characterize a shape's *spatial parts*. Not all schemes have this feature. One that lacks it is *Fourier description*. On this approach, a shape is first encoded as a vector of local features (e.g., a sequence of turning angles). The Fourier transform of this vector outputs a series of components characterizing the entire shape via its amplitude and phase at various frequencies (Zahn and Roskies [1972]; Zhang and Lu [2004]). These components comprise the shape's Fourier description (FD). Importantly, the constituents of an FD do *not* represent a shape's spatial parts. Rather, as figure 16 illustrates, each represents a global contributor to the frequency content of the entire shape (Elder [2018]).

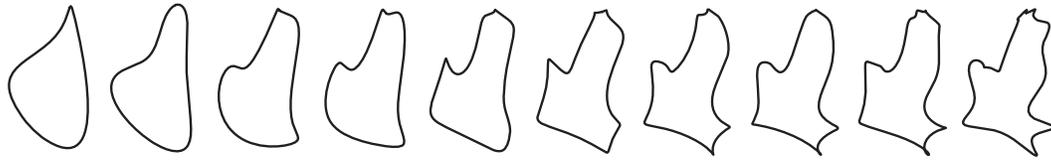


Figure 16. FD-based approximations of a complex shape using 1(left) to 10 (right) components. Reprinted with permission from (Elder [2018]).

We might imagine a system that displayed constancy for Fourier components. The system would be selectively sensitive to similarities between objects that, while distinct, share certain components of their Fourier series. Skeletal representations would be poorly suited for this type of constancy. When a shape has multiple parts, each node of its skeletal representation encodes a spatial part of the shape; there is no single node dedicated to the shape “as a whole.” Thus, there is nowhere in the tree to “put” representations of a shape’s FD components, since there is no spatial part that they characterize. Accordingly, the FD approach is sometimes offered as a “holistic” alternative to part-based schemes (Cortese and Dyre [1996]).

However, this blindness to part structure is a severe limitation of FD, since it precludes sensitivity to similarities between shapes that share parts but differ in their spatial arrangement (Arguin and Saumier [2004]). Moreover, the FD approach predicts that the visual system should represent individual Fourier components of shapes generated by combining multiple frequencies, but there is scant evidence for this (Brincat and Connor [2004]; Elder [2018]). For instance, neurons in inferotemporal cortex are typically insensitive to shared Fourier components between shapes that differ in other respects (Albright and Gross [1990]).

Thus, skeletal representations are unsuited to subserve constancy for Fourier components, and this form of constancy is also not very prominent in human vision.

4.6 Summing up

I have elucidated four varieties of shape constancy. Each targets a complex property represented by a constituent or aspect of a skeletal representation-type. Skeletal representations extract information about abstract shape properties, allowing constancies for them through proximal variation caused by changes to an object's determinate shape. Abstract constancies are exercised through changes that alter or peel away constituents of these representations while leaving other constituents or aspects of the representation unchanged.

The four types of constancy are exercised through different kinds of variation. While determinate shape constancy is paradigmatically exercised through changes in perspective or rigid transformations of an object, more abstract constancies can be exercised through nonrigid transformations as well, and also across differences between category exemplars. For instance, two cats that differ in determinate shape might have very similar skeletal shapes, and even identical skeletal topologies.

I do not intend this to be the last word on the varieties of shape constancy. There may be further constancies more abstract than those considered, or which fall between them. An advantage of skeletal representations is that they extract a wide range of shape features, potentially enabling constancies for various complex properties comprising features from different levels of abstraction—e.g., composites of metric and topological properties. Properties that appear “messy” from the standpoint of geometry can be adaptive for perception to represent, particularly when they endure through intrinsic changes that objects in our environment often undergo. Future work may unearth further constancies underpinned by skeletal representations, or perhaps identify constancies demanding entirely different systems of representation.

5 Conclusion

This paper has developed a pluralist perspective on shape constancy. Our perceptual systems extract shape properties at varying levels of abstraction and exercise distinct constancies for properties at these various levels. These constancies are explained by the syntactic structure of perceptual shape representations, which assign separable constituents to various abstract shape properties. I've argued that the varieties of shape constancy are not confined to subpersonal processing, but manifest in experience as well.

I close with three broader implications concerning the nature and function of perceptual constancies.

First, my account fits with an emerging consensus that constancy mechanisms do not function to *discount* variable properties like illumination (*pace* Helmholtz [1867/1962]), but to represent these properties *separately* from more stable properties, enabling information about these properties to be accessed independently (Hilbert [2005]; Matthen [2010]; Brown [2014]; O'Dea [2022]). Likewise, mechanisms of shape perception do not aim to *discount* information about a cat's posture, but rather to represent its shape in a manner that makes its present posture accessible separately from its posture-invariant structure.

Second, the account uncovers a connection between perceptual constancy and the format of perceptual representation. It is possible to exhibit constancies for perceptible dimensions at multiple levels of abstraction. However, doing so generally requires a representational format that *extracts* abstract properties rather than simply “nesting” them in the representation of more determinate properties (*pace* Kulvicki [2007]). As such, shape representations are unlike familiar imagistic representations like drawings or color photographs which lack discrete constituents for abstract, categorical properties, instead nesting information about these properties in the more determinate

properties they depict (Quilty-Dunn [2020]). Rather, shape representations assign separable representation-types to abstract properties.

Third, it is sometimes suggested that the primary aim of constancy mechanisms is to separate *intrinsic* properties of objects, like size or reflectance, from *relational* properties like distance or illumination. I believe this is only a crude approximation to the truth. The more fundamental aim is to separate *more variable* properties of objects from *less variable* properties preserved through changes in the former. Many relational properties, such as depth or orientation, are indeed highly variable, and our visual systems do form representations that endure through changes in them. However, on closer inspection, we find that certain intrinsic properties are highly variable too. To reidentify objects through these variations, we must differentiate volatile intrinsic properties from stable ones. The pluralist approach illuminates how this challenge is met in the case of shape.

Acknowledgements

Thanks to Alex Byrne, Tyler Brooke-Wilson, Will Davies, Chaz Firestone, Jake Quilty-Dunn, and two anonymous reviewers for valuable comments and/or discussion. Thanks also to audiences at Ruhr-Universität Bochum, the University of Salzburg, and the 2022 SPP/ESPP meeting at the University of Milan, where portions of this material were presented.

References

- Albright, T. D. and Gross, C. G. [1990]: ‘Do Inferior Temporal Neurons Encode Shape by Acting as Fourier Descriptor Filters?’, *Proceedings of the International Conference on Fuzzy Logic and Neural Networks*, **2**, pp. 375-78.
- Allen, K. [2018]: ‘Perceptual Constancy and Apparent Properties’, in F. Dorsch and F. Macpherson (eds.), *Phenomenal Presence*, Oxford: Oxford University Press, pp. 39-57.
- Amir, O., Biederman, I. and Hayworth, K. J. [2012]: ‘Sensitivity to Nonaccidental Properties Across Various Shape Dimensions’, *Vision Research*, **62**, pp. 35-43.
- Arguin, M. and Saumier, D. [2004]: ‘Independent Processing of Parts and of Their Spatial Organization in Complex Visual Objects’, *Psychological Science*, **15**, pp. 629-633.
- Austin, J. L. [1962]: *Sense and Sensibilia*, Oxford: Oxford University Press.
- Ayzenberg, V. and Behrmann, M. [2022a]: ‘Does the Brain’s Ventral Visual Pathway Compute Object Shape?’, *Trends in Cognitive Sciences*, **22**, pp. 1119-32.
- Ayzenberg, V. and Behrmann, M. [2022b]: ‘The Dorsal Visual Pathway Represents Object-Centered Spatial Relations for Object Recognition’, *Journal of Neuroscience*, **42**, pp. 4693-710.
- Ayzenberg, V., Chen, Y., Yousif, S. R. and Lourenco, S. F. [2019]: ‘Skeletal Representations of Shape in Human Vision: Evidence for a Pruned Medial Axis Model’, *Journal of Vision*, **19**, pp. 1-21.
- Ayzenberg, V., Kamps, F. S., Dilks, D. D. and Lourenco, S. F. [2022]: ‘Skeletal Representations of Shape in the Human Visual Cortex’, *Neuropsychologia*, **164**, 108092.

- Ayzenberg, V. and Lourenco, S. F. [2019]: ‘Skeletal Descriptions of Shape Provide Unique Perceptual Information for Object Recognition’, *Scientific Reports*, **9**, pp. 1-13.
- Ayzenberg, V. and Lourenco, S. [2022]: ‘Perception of an Object’s Global Shape is Best Described by a Model of Skeletal Structure in Human Infants’, *eLife*, **11**, e74943.
- Baker, N. and Kellman, P. J. [2018]: ‘Abstract Shape Representation in Human Visual Perception’, *Journal of Experimental Psychology: General*, **147**, pp. 1295-1308.
- Baker, N. and Kellman, P. J. [2021]: ‘Constant Curvature Modeling of Abstract Shape Representation’, *Plos One*, **16**, e0254719.
- Baker, N., Lu, H., Erlikhman, G. and Kellman, P. J. [2020]: ‘Local Features and Global Shape Information in Object Classification by Deep Convolutional Neural Networks’, *Vision Research*, **172**, pp. 46-61.
- Barenholtz, E. and Tarr, M.J. [2008]: ‘Visual Judgment of Similarity Across Shape Transformations: Evidence for a Compositional Model of Articulated Objects’, *Acta Psychologica*, **128**, pp. 331-38.
- Beck, J. [2019]: ‘Perception is Analog: The Argument From Weber’s Law’, *The Journal of Philosophy*, **116**, pp. 319-49.
- Behrmann, M., Peterson, M.A., Moscovitch, M. and Suzuki, S. [2006]: ‘Independent Representation of Parts and the Relations Between Them: Evidence From Integrative Agnosia’, *Journal of Experimental Psychology: Human Perception and Performance*, **32**, pp. 1169-84.
- Behrmann, M. and Williams, P. [2007]: ‘Impairments in Part–Whole Representations of Objects in Two Cases of Integrative Visual Agnosia’, *Cognitive Neuropsychology*, **24**, pp. 701-30.
- Bennett, D. J. and Vuong, Q. C. [2006]: ‘A Stereo Advantage in Generalizing Over Changes in Viewpoint on Object Recognition Tasks’, *Perception and Psychophysics*, **68**, pp. 1082-93.
- Bennett, D. J. [2012]: ‘Seeing Shape: Shape Appearances and Shape Constancy’, *The British Journal for the Philosophy of Science*, **63**, pp. 487-518.
- Biederman, I. [1987]: ‘Recognition-by-Components: A Theory of Human Image Understanding’, *Psychological Review*, **94**, pp. 115-47.
- Biederman, I. [2001]: ‘Recognizing Depth-Rotated Objects: A Review of Recent Research and Theory’, *Spatial Vision*, **13**, pp. 241-54.
- Biederman, I. and Bar, M. [1999]: ‘One-Shot Viewpoint Invariance in Matching Novel Objects’, *Vision Research*, **39**, pp. 2885-99.
- Biederman, I. and Gerhardstein, P. C. [1993]: ‘Recognizing Depth-Rotated Objects: Evidence and Conditions for Three-Dimensional Viewpoint Invariance’, *Journal of Experimental Psychology: Human Perception and Performance*, **19**, pp. 1162-82.
- Blum, H. [1973]: ‘Biological Shape and Visual Science (Part I)’, *Journal of Theoretical Biology*, **38**, pp. 205-87.
- Briscoe, R. E. [2009]: ‘Egocentric Spatial Representation in Action and Perception’, *Philosophy and Phenomenological Research*, **79**, pp. 423-60.
- Brown, D. [2014]: ‘Colour Layering and Colour Constancy’, *Philosophers’ Imprint*, **14**, pp. 1-31.
- Buccella, A. [2021]: ‘The Problem of Perceptual Invariance’, *Synthese*, **199**, pp. 13883-905.
- Bülthoff, H. H. and Edelman, S. [1992]: ‘Psychophysical Support for a Two-Dimensional View Interpolation Theory of Object Recognition’, *Proceedings of the National Academy of Sciences*, **89**, pp. 60-64.
- Burge, T. [2010]: *Origins of Objectivity*, Oxford: Oxford University Press.
- Burge, T. [2014]: ‘Reply to Rescorla and Peacocke: Perceptual Content in Light of Perceptual Constancies and Biological Constraints’, *Philosophy and Phenomenological Research*, **88**, pp. 485-501.
- Buxbaum, L. J., Coslett, H. B., Montgomery, M. W. and Farah, M. J. [1996]: ‘Mental Rotation May

- Underlie Apparent Object-Based Neglect', *Neuropsychologia*, **14**, pp. 113-26.
- Cacciamani, L., Ayars, A.A. and Peterson, M.A. [2014]: 'Spatially Rearranged Object Parts Can Facilitate Perception of Intact Whole Objects', *Frontiers in Psychology*, **5**, pp. 1-11.
- Chaisilprungraung, T., German, J. and McCloskey, M. [2019]: 'How are Object Shape Axes Defined? Evidence From Mirror-Image Confusions', *Journal of Experimental Psychology: Human Perception and Performance*, **45**, pp. 111-24.
- Cheng, K. and Gallistel, C. R. [2005]: 'Shape Parameters Explain Data from Spatial Transformations: Comment on Pearce et al. (2004) and Tommasi and Polli (2004)', *Journal of Experimental Psychology: Animal Behavior Processes*, **31**, pp. 254-259.
- Cohen, J. [2015]: 'Perceptual Constancy', in M. Matthen (ed.), *The Oxford Handbook of Philosophy of Perception*, Oxford: Oxford University Press, pp. 621-39.
- Cortese, J. M. and Dyre, B. P. [1996]: 'Perceptual Similarity of Shapes Generated from Fourier Descriptors', *Journal of Experimental Psychology: Human Perception and Performance*, **22**, pp. 133-43.
- Davies, W. [2022]: 'The Paradox of Colour Constancy: Plotting the Lower Borders of Perception', *Noûs*, **56**, pp. 787-813.
- Destler, N., Singh, M. and Feldman, J. [2019]: 'Shape Discrimination Along Morph-Spaces', *Vision Research*, **158**, pp. 189-99.
- Destler, N., Singh, M. and Feldman, J. [2023]: 'Skeleton-Based Shape Similarity', *Psychological Review*, DOI: 10.1037/rev0000412.
- Di Carlo, J. J. and Maunsell, J. H. R. [2003]: 'Anterior Inferotemporal Neurons of Monkeys Engaged in Object Recognition Can Be Highly Sensitive to Object Retinal Position', *Journal of Neurophysiology*, **89**, pp. 3264-78.
- Dilks, D. D., Julian, J. B., Kubilius, J., Spelke, E. S. and Kanwisher, N. [2011]: 'Mirror-Image Sensitivity and Invariance in Object and Scene Processing Pathways', *Journal of Neuroscience*, **31**, pp. 11305-12.
- Dretske, F. [1981]: *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.
- Driver, J. and Halligan, P. W. [1991]: 'Can Visual Neglect Operate in Object-Centred Co-ordinates? An Affirmative Single-Case Study', *Cognitive Neuropsychology*, **8**, pp. 475-96.
- Edelman, S. and Duvdevani-Bar, S. [1997]: 'A Model of Visual Recognition and Categorization', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **352**, pp. 1191-202.
- Elder, J. H. [2018]: 'Shape from Contour: Computation and Representation', *Annual Review of Vision Science*, **4**, pp. 423-50.
- Epstein, W., Hatfield, G. and Muise, G. [1977]: 'Perceived Shape at a Slant as a Function of Processing Time and Processing Load', *Journal of Experimental Psychology: Human Perception and Performance*, **3**, pp. 473-83.
- Epstein, W. and Park, J. N. [1963]: 'Shape Constancy: Functional Relationships and Theoretical Formulations', *Psychological Bulletin*, **60**, pp. 265-88.
- Erdogan, G. and Jacobs, R. A. [2017]: 'Visual Shape Perception as Bayesian Inference of 3D Object-Centered Shape Representations', *Psychological Review*, **124**, pp. 740-61.
- Feldman, J. [2007]: 'Formation of Visual "Objects" in the Early Computation of Spatial Relations', *Perception and Psychophysics*, **69**, pp. 816-27.
- Feldman, J. and Singh, M. [2006]: 'Bayesian Estimation of the Shape Skeleton', *Proceedings of the National Academy of Sciences*, **103**, pp. 18014-19.
- Feldman, J., Singh, M., Briscoe, E., Froyen, V., Kim, S. and Wilder, J. [2013]: 'An Integrated Bayesian Approach to Shape Representation and Perceptual Organization', in S.J. Dickinson and Z. Pizlo (eds.), *Shape Perception in Human and Computer Vision*, New York: Springer, pp. 55-70.

- Filimon, F. [2015]: ‘Are All Spatial Reference Frames Egocentric? Reinterpreting Evidence for Allocentric, Object-Centered, or World-Centered Reference Frames’, *Frontiers in Human Neuroscience*, **9**, 648.
- Firestone, C. and Scholl, B. J. [2014]: “‘Please Tap the Shape, Anywhere You Like’: Shape Skeletons in Human Vision Revealed by an Exceedingly Simple Measure’, *Psychological Science*, **25**, pp. 377-86.
- Fleming, R. W. and Schmidt, F. [2019]: ‘Getting “Fumpered”: Classifying Objects by What Has Been Done to Them’, *Journal of Vision*, **19**, pp. 1-12.
- Foster, D. H. [2011]: ‘Color Constancy’, *Vision Research*, **51**, pp. 674-700.
- Froyen, V., Feldman, J. and Singh, M. [2010]: ‘A Bayesian Framework for Figure-Ground Interpretation’, in J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, and A. Culotta (eds.), *Advances in Neural Information Processing Systems*, Vancouver: Curran Associates, vol. 23, pp. 631-39.
- Gibson, J. J. [1979]: *The Ecological Approach to Visual Perception*, Boston, MA: Houghton Mifflin.
- Green, E. J. [2017]: ‘A Layered View of Shape Perception’, *The British Journal for the Philosophy of Science*, **68**, pp. 355-87.
- Green, E. J. [2019]: ‘On the Perception of Structure’, *Noûs*, **53**, pp. 564-92.
- Green, E. J. [2022]: ‘The Puzzle of Cross-Modal Shape Experience’, *Noûs*, **56**, pp. 867-96.
- Green, E. J. [2023]: ‘Perceptual Constancy and Perceptual Representation’, *Analytic Philosophy*, DOI: 10.1111/phib.12293.
- Gregory, E. and McCloskey, M. [2010]: ‘Mirror-Image Confusions: Implications for Representation and Processing of Object Orientation’, *Cognition*, **116**, pp. 110-29.
- Grill-Spector, K., Kourtzi, Z. and Kanwisher, N. [2001]: ‘The Lateral Occipital Complex and its Role in Object Recognition’, *Vision Research*, **41**, pp. 1409-22.
- Guggenmos, M., Thoma, V., Cichy, R. M., Haynes, J. D., Sterzer, P. and Richardson-Klavehn, A. [2015]: ‘Non-Holistic Coding of Objects in Lateral Occipital Complex with and Without Attention’, *Neuroimage*, **107**, pp. 356-63.
- Hafri, A., Gleitman, L. R., Landau, B. and Trueswell, J. C. [2023]: ‘Where Word and World Meet: Language and Vision Share an Abstract Representation of Symmetry’, *Journal of Experimental Psychology: General*, **152**, pp. 509-27.
- Hafri, A., Green, E. J. and Firestone, C. [forthcoming]: ‘Compositionality in Visual Perception’, *Behavioral and Brain Sciences*.
- Harrison, S. J. and Feldman, J. [2009]: ‘The Influence of Shape and Skeletal Axis Structure on Texture Perception’, *Journal of Vision*, **9**, pp. 1-21.
- Hatfield, G. [2016]: ‘Perceiving as Having Subjectively Conditioned Appearances’, *Philosophical Topics*, **44**, pp. 149-78.
- von Helmholtz, H. [1867/1962]: *Helmholtz’s Treatise on Physiological Optics*, New York: Dover.
- Hermann, K., Chen, T. and Kornblith, S. [2020]: ‘The Origins and Prevalence of Texture Bias in Convolutional Neural Networks’, *Advances in Neural Information Processing Systems*, **33**, pp. 19000-15.
- Hilbert, D. [2005]: ‘Color Constancy and the Complexity of Color’, *Philosophical Topics*, **33**, pp. 141-58.
- Hill, C. S. [2020]: ‘Appearance and Reality’, *Philosophical Issues*, **30**, pp. 175-91.
- Hoffman, D. D. and Richards, W. A. [1984]: ‘Parts of Recognition’, *Cognition*, **18**, pp. 65-96.
- Howard, I. P., Fujii, Y., Allison, R. S. and Kirillos, R. [2014]: ‘Shape Constancy Measured by a Canonical-Shape Method’, *Vision Research*, **94**, pp. 33-40.
- Hummel, J. E. [2000]: ‘Where View-Based Theories Break Down: The Role of Structure in Shape Perception and Object Recognition’, in E. Dietrich and A. Markman (eds.), *Cognitive Dynamics*:

- Conceptual Change in Humans and Machines*, Hillsdale, NJ: Lawrence Erlbaum, pp. 157-85.
- Hummel, J. E. [2013]: ‘Object Recognition’, In D. Reisburg (ed.), *Oxford Handbook of Cognitive Psychology*, Oxford: Oxford University Press, pp. 32-46.
- Humphreys, G. W. and Quinlan, P. T. [1988]: ‘Priming Effects Between Two-Dimensional Shapes’, *Journal of Experimental Psychology: Human Perception and Performance*, **14**, pp. 203-20.
- Hung, C. C., Carlson, E. T. and Connor, C. E. [2012]: ‘Medial Axis Shape Coding in Macaque Inferotemporal Cortex’, *Neuron*, **74**, pp. 1099-113.
- Kimia, B.B. [2003]: ‘On the Role of Medial Geometry in Human Vision’, *Journal of Physiology-Paris*, **97**, pp. 155-90.
- Kovács, I., Fehér, A. and Julesz, B. [1998]: ‘Medial-Point Description of Shape: A Representation for Action Coding and its Psychophysical Correlates’, *Vision Research*, **38**, pp. 2323-33.
- Kovács, I. and Julesz, B. [1994]: ‘Perceptual Sensitivity Maps Within Globally Defined Visual Shapes’, *Nature*, **370**, pp. 644-46.
- Kulvicki, J. [2007]: ‘Perceptual Content is Vertically Articulate’, *American Philosophical Quarterly*, **44**, pp. 357-69.
- Kulvicki, J. [2015]: ‘Analog Representation and the Parts Principle’, *Review of Philosophy and Psychology*, **6**, pp. 165-80.
- Lande, K. J. [2018]: ‘The Perspectival Character of Perception’, *The Journal of Philosophy*, **115**, pp. 187-214.
- Lande, K. J. [2023]: ‘Contours of Vision: Towards a Compositional Semantics of Perception’, *The British Journal for the Philosophy of Science*, <https://doi.org/10.1086/725094>.
- Lescroart, M. D. and Biederman, I. [2013]: ‘Cortical Representation of Medial Axis Structure’, *Cerebral Cortex*, **23**, pp. 629-37.
- Loomis, J. M., Philbeck, J. W. and Zahorik, P. [2002]: ‘Dissociation Between Location and Shape in Visual Space’, *Journal of Experimental Psychology: Human Perception and Performance*, **28**, pp. 1202-12.
- Lowet, A. S., Firestone, C. and Scholl, B. J. [2018]: ‘Seeing Structure: Shape Skeletons Modulate Perceived Similarity’, *Attention, Perception, and Psychophysics*, **80**, pp. 1278-89.
- Macpherson, F. [2006]: ‘Ambiguous Figures and the Content of Experience’, *Noûs*, **40**, pp. 82-117.
- Marr, D. and Nishihara, H. K. [1978]: ‘Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes’, *Proceedings of the Royal Society of London, B: Biological Sciences*, **200**, pp. 269-94.
- Matthen, M. [2005]: *Seeing, Doing, and Knowing*, Oxford: Oxford University Press.
- Matthen, M. [2010]: ‘How Things Look (and What Things Look that Way)’, in B. Nanay (ed.), *Perceiving the World*, Oxford: Oxford University Press, pp. 226-53.
- Nickerson, R. S. and Adams, M. J. [1979]: ‘Long-Term Memory for a Common Object’, *Cognitive Psychology*, **11**, pp. 287-307.
- O’Dea, J. [2022]: ‘Perceptual Constancy and the Dimensions of Perceptual Experience’, *Phenomenology and the Cognitive Sciences*, **21**, pp. 421-34.
- Olin, L. [2016]: ‘Burge on Perception and Sensation’, *Synthese*, **193**, pp. 1479-508.
- Oliver, Z. J., Cristino, F., Roberts, M. V., Pegna, A. J. and Leek, E. C. [2018]: ‘Stereo Viewing Modulates Three-Dimensional Shape Processing During Object Recognition: A High-Density ERP Study’, *Journal of Experimental Psychology: Human Perception and Performance*, **44**, pp. 518-34.
- Palmer, S. and Rock, I. [1994]: ‘Rethinking Perceptual Organization: The Role of Uniform Connectedness’, *Psychonomic Bulletin and Review*, **1**, pp. 29-55.
- Papale, P., Leo, A., Handjaras, G., Cecchetti, L., Pietrini, P. and Ricciardi, E. [2020]: ‘Shape Coding in Occipito-Temporal Cortex Relies on Object Silhouette, Curvature, and Medial

- Axis', *Journal of Neurophysiology*, **124**, pp. 1560-70.
- Peacocke, C. [1992]: *A Study of Concepts*, Cambridge, MA: MIT Press.
- Phillips, F., Todd, J. T., Koenderink, J. J. and Kappers, A. M. [2003]: 'Perceptual Representation of Visible Surfaces', *Perception and Psychophysics*, **65**, pp. 747-62.
- Pizlo, Z., Li, Y. and Steinman, R. M. [2008]: 'Binocular Disparity Only Comes into Play When Everything Else Fails: A Finding with Broader Implications than One Might Suppose', *Spatial Vision*, **21**, pp. 495-508.
- Pizlo, Z. and Stevenson, A. K. [1999]: 'Shape Constancy from Novel Views', *Perception and Psychophysics*, **61**, pp. 1299-1307.
- Quilty-Dunn, J. [2020]: 'Perceptual Pluralism', *Noûs*, **54**, pp. 807-38.
- Quinlan, P. T. and Allen, K. [2018]: 'The Nature of Shape Constancy Mechanisms as Revealed by Shape Priming', *Journal of Vision*, **18**, pp. 1-17.
- Quinlan, P. T. and Humphreys, G. W. [1993]: 'Perceptual Frames of Reference and Two-Dimensional Shape Recognition: Further Examination of Internal Axes', *Perception*, **22**, pp. 1343-64.
- Rescorla, M. [2014]: 'Perceptual Constancies and Perceptual Modes of Presentation', *Philosophy and Phenomenological Research*, **88**, pp. 468-76.
- Rock, I. [1983]: *The Logic of Perception*, Cambridge, MA: MIT Press.
- Rudel, R. G. and Teuber, H.-L. [1963]: 'Discrimination of Direction of Line in Children', *Journal of Comparative and Physiological Psychology*, **56**, pp. 892-8.
- Schellenberg, S. [2008]: 'The Situation-Dependency of Perception', *The Journal of Philosophy*, **105**, pp. 55-84.
- Schellenberg, S. [2018]: *The Unity of Perception*, Oxford: Oxford University Press.
- Schmidt, F. and Fleming, R. W. [2016]: 'Visual Perception of Complex Shape-Transforming Processes', *Cognitive Psychology*, **90**, pp. 48-70.
- Schmidt, F. and Fleming, R. W. [2018]: 'Identifying Shape Transformations from Photographs of Real Objects', *PloS One*, **13**, e0202115.
- Schmidt, F., Phillips, F. and Fleming, R. W. [2019]: 'Visual Perception of Shape-Transforming Processes: 'Shape Scission'', *Cognition*, **189**, pp. 167-80.
- Sebastian, T. B. and Kimia, B. B. [2005]: 'Curves vs. Skeletons in Object Recognition', *Signal Processing*, **85**, pp. 247-63.
- Sebastian, T. B., Klein, P. N. and Kimia, B. B. [2004]: 'Recognition of Shapes by Editing Their Shock Graphs', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **26**, pp. 550-71.
- Sekuler, A. B. and Swimmer, M. B. [2000]: 'Interactions Between Symmetry and Elongation in Determining Reference Frames for Object Perception', *Canadian Journal of Experimental Psychology*, **54**, pp. 42-56.
- Shaked, D. and Bruckstein, A. M. [1998]: 'Pruning Medial Axes', *Computer Vision and Image Understanding*, **69**, pp. 156-69.
- Siddiqi, K., Shokoufandeh, A., Dickinson, S. J. and Zucker, S. W. [1999]: 'Shock Graphs and Shape Matching', *International Journal of Computer Vision*, **35**, pp. 13-32.
- Singh, M., Seyranian, G. D. and Hoffman, D. D. [1999]: 'Parsing Silhouettes: The Short-Cut Rule', *Perception and Psychophysics*, **61**, pp. 636-60.
- Slater, A. and Morison, V. [1985]: 'Shape Constancy and Slant Perception at Birth', *Perception*, **14**, pp. 337-44.
- Spröte, P. and Fleming, R. W. [2016]: 'Bent Out of Shape: The Visual Inference of Non-Rigid Shape Transformations Applied to Objects', *Vision Research*, **126**, pp. 330-46.
- Spröte, P., Schmidt, F. and Fleming, R. W. [2016]: 'Visual Perception of Shape Altered by Inferred

- Causal History’, *Scientific Reports*, **6**, pp. 1-11.
- Stankiewicz, B. J. and Hummel, J. E. [1996]: ‘Categorical Relations in Shape Perception’, *Spatial Vision*, **10**, pp. 201-36.
- Stankiewicz, B. J., Hummel, J. E. and Cooper, E. E. [1998]: ‘The Role of Attention in Priming for Left–Right Reflections of Object Images: Evidence for a Dual Representation of Object Shape’, *Journal of Experimental Psychology: Human Perception and Performance*, **24**, pp. 732-44.
- Sun, Z. and Firestone, C. [2021]: ‘Curious Objects: How Visual Complexity Guides Attention and Engagement’, *Cognitive Science*, **45**, e12933.
- Sun, Z. and Firestone, C. [2022]: ‘Beautiful on the Inside: Aesthetic Preferences and the Skeletal Complexity of Shapes’, *Perception*, **51**, pp. 904-18.
- Sutherland, N. S. [1960]: ‘Visual Discrimination of Orientation by *Octopus*: Mirror Images’, *British Journal of Psychology*, **51**, pp. 9-18.
- Thoma, V., Hummel, J. E. and Davidoff, J. [2004]: ‘Evidence for Holistic Representations of Ignored Images and Analytic Representations of Attended Images. *Journal of Experimental Psychology: Human Perception and Performance*, **30**, pp. 257-67.
- Thompson, B. [2010]: ‘The Spatial Content of Experience’, *Philosophy and Phenomenological Research*, **81**, pp. 146-84.
- Tipper, S. P. and Behrmann, M. [1996]: ‘Object-Centered Not Scene-Based Visual Neglect’, *Journal of Experimental Psychology: Human Perception and Performance*, **22**, pp. 1261-78.
- Tittle, J. S., Todd, J. T., Perotti, V. J. and Norman, J. F. [1995]: ‘Systematic Distortion of Perceived Three-Dimensional Structure from Motion and Binocular Stereopsis’, *Journal of Experimental Psychology: Human Perception and Performance*, **21**, pp. 663-78.
- Todd, J. T. [2004]: ‘The Visual Perception of 3D Shape’, *Trends in Cognitive Sciences*, **8**, pp. 115-21.
- Todd, J. T., Chen, L. and Norman, J. F. [1998]: ‘On the Relative Salience of Euclidean, Affine, and Topological Structure for 3D Form Discrimination’, *Perception*, **27**, pp. 273-82.
- Todd, J. T. and Norman, J. F. [2003]: ‘The Visual Perception of 3-D Shape from Multiple Cues: Are Observers Capable of Perceiving Metric Structure?’, *Perception and Psychophysics*, **65**, pp. 31-47.
- Todd, J. T. and Petrov, A. A. [2022]: ‘The Many Facets of Shape’, *Journal of Vision*, **22**, pp. 1-30.
- Todd, J. T., Weismantel, E. and Kallie, C. S. [2014]: ‘On the Relative Detectability of Configurational Properties’, *Journal of Vision*, **14**, pp. 1-8.
- Turnbull, O. H. and McCarthy, R. A. [1996]: ‘Failure to Discriminate Between Mirror-Image Objects: A Case of Viewpoint-Independent Object Recognition?’, *Neurocase*, **2**, pp. 63-72.
- Ullman, S. and Basri, R. [1991]: ‘Recognition by Linear Combinations of Models’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, pp. 992-1006.
- Vannuscorps, G., Galaburda, A. and Caramazza, A. [2022]: ‘Shape-Centered Representations of Bounded Regions of Space Mediate the Perception of Objects’, *Cognitive Neuropsychology*, **39**, pp. 1-50.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M. and von der Heydt, R. [2012]: ‘A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure–Ground Organization’, *Psychological Bulletin*, **138**, pp. 1172-217.
- Wagner, M. [1985]: ‘The Metric of Visual Space’, *Perception and Psychophysics*, **38**, 483-95.
- Wagner, M. and Gambino, A. J. [2016]: ‘Variations in the Anisotropy and Affine Structure of Visual Space: A Geometry of Visibles with a Third Dimension’, *Topoi*, **35**, pp. 583-98.
- Warren, W. H. [2012]: ‘Does this Computational Theory Solve the Right Problem? Marr, Gibson, and the Goal of Vision’, *Perception*, **41**, pp. 1053-60.
- Wilder, J., Feldman, J. and Singh, M. [2016]: ‘The Role of Shape Complexity in the Detection of Closed Contours’, *Vision Research*, **126**, pp. 220-31.
- Wright, W. [2013]: ‘Colour Constancy Reconsidered’, *Acta Analytica*, **28**, pp. 435-55.

- Yu, Y., Todd, J. T. and Petrov, A. A. [2021]: 'Failures of Stereoscopic Shape Constancy over Changes of Viewing Distance and Size for Bilaterally Symmetric Polyhedra', *Journal of Vision*, **21**, pp. 1-19.
- Zahn, C. T. and Roskies, R. Z. [1972]: 'Fourier Descriptors for Plane Closed Curves', *IEEE Transactions on Computers*, **100**, pp. 269-81.
- Zhang, D. and Lu, G. [2004]: 'Review of Shape Representation and Description Techniques', *Pattern Recognition*, **37**, pp. 1-19.