Hummel, J. E. (in press). The complementary properties of holistic and analytic representations of object shape. To appear in G. Rhodes and M. Peterson (Eds.), *Analytic and Holistic Processes in the Perception of Faces, Objects and Scenes*.

The Complementary Properties of Holistic and Analytic Representations of Shape

John E. Hummel
University of California Los Angeles

John E. Hummel
Department of Psychology
UCLA
405 Hilgard Ave.
Los Angeles, CA 90095-1563
(310) 206-9262
jhummel@psych.ucla.edu

Address from August, 2001 to July, 2002:
Psychology Building
1101 E 10th St.
Indiana University
Bloomington, IN. 47405-7007
(812) 855-4642
johummel@indiana.edu

Preparation of this chapter was supported by NSF Grant SBR-9729023, and by grants from the UCLA Academic Senate and HRL Laboratories.

REPRESENTATIONS OF SHAPE

The mental representation of object shape plays a central role in numerous activities, including object recognition (Biederman & Ju, 1988), categorization and category learning (Rosch & Mervis, 1975; Rosch et al, 1976; Saiki & Hummel, 1996, 1998a; Tversky & Hemenway, 1984), visual search (Enns, 1992), scene segmentation (Peterson & Gibson, 1994), and even visual reasoning and analogy (Goldstone, Medin & Gentner, 1991; Hummel, 2000; Hummel & Holyoak, 2001; Pedone, Hummel & Holyoak, 2001). What does it mean to represent an object's shape?

A representation of shape consists of four independent components (cf. Palmer, 1978). The first is a set of <u>primitives</u>—a vocabulary of basic elements that can be put together to describe a shape. Examples include pixels (as in the case of the raw retinal image), simple 2-dimensional (2D) image features such as lines and vertices (Edeman, 1998; Poggio & Edelman, 1990; Riesenhuber & Poggio, 1999; Tarr & Bülthoff, 1995), more complex features such as volumetric (3D) parts (e.g., Marr & Nishihara, 1978), and approximations of volumetric parts (Biederman, 1987; Hummel & Biederman, 1992), among others.

The second component is a <u>reference frame</u>—coordinate system that serves as the basis for specifying the arrangement of an object's features or parts. A reference frame is defined by three independent properties: an <u>origin</u>, which serves as a zero point, relative to which locations are defined; a <u>scale</u>, which maps distances (from the origin) onto numerical values (i.e., coordinates in the reference frame); and an <u>orientation</u>, which specifies the directions of extent of the coordinate system's axes (i.e., the direction in which each axis "points"). These properties are independent in the sense that any or all of them may be defined relative to the viewer, the object or the environment. For example, it is possible to define a reference frame in which the origin and scale are defined relative to the object, but the orientation is defined relative to the viewer (see, e.g., Hummel & Stankiewicz, 1996a).

The third component of a representation of shape is a vocabulary of <u>relations</u> for specifying how an object's features or parts are arranged within the reference frame. The most direct approach is simply to represent the elements in terms of their coordinates in the reference frame: Each coordinate expresses a numerical relation (namely, a distance along one axis) between an element and the origin of the reference frame. Coordinate-based coding is the dominant approach to representing relations among computational models of object recognition, and characterizes representations based both 2D views (Edelman, 1998; Edelman & Poggio, 1991; Poggio & Edelman, 1990; Reisenhuber & Poggio, 1999; Tarr & Bülthoff, 1995) and 3D models (Lowe, 1985; Ullman, 1989, 1995; for reviews, see Hummel & Stankiewicz, 1996b, and Hummel, 2000). An alternative to direct coordinate-based coding is to represent an object's features or parts in terms of their relations, not to the origin of the reference frame, but to one another (e.g., Biederman, 1987; Clowes, 1967; Hummel & Biederman, 1992; Marr & Nishihara, 1978; Sutherland, 1969). The resulting representation is referred to as a "structural description." Representing relations explicitly affords tremendous flexibility in the vocabulary and form of the relations expressed. In addition to expressing relative location, elements can be represented in terms of their relative size, orientation, etc., and these relations may be expressed categorically (e.g., "element A is above element B" or "A is larger than B"), metrically (e.g., "A is 4 units above B"; "A is 2.3 times larger than B"), or both (Hummel & Stankiewicz, 1998). By contrast, coordinate-based representations express only one type of relation (i.e., distance from the origin of the reference frame), and the coordinates are necessarily metrically precise.

The final aspect of a representation of shape concerns the <u>binding</u> of elements to one another and to their locations and/or relations (Hummel & Biederman, 1992). It is one thing to know that an image contains a square and a circle, and that one of these shapes is above the other; it is another to bind the shapes to their relations in order to specify that the circle is above the square rather than vice-versa. Likewise, knowing that an image contains features A, B and C, and that locations d, e and f are occupied is not the same as binding features to locations to specify which feature resides at which location. The issue of binding is closely related to (and easily confused with) the issue of relations, but they are importantly different: The latter refers to the <u>vocabulary</u> of relations used to express the configuration of an object's features or parts; the former refers to the manner in which elements or properties are conjoined with one another and with their locations and/or relations. There are two qualitatively different ways to represent these bindings: statically and dynamically.

A static binding is a conjunctive representation, in which collections of elements or properties in the world map onto individual elements of the representation. For example, each neuron in visual area V1 represents a conjunction (i.e., static binding) of a particular orientation to a particular spatial frequency, a particular location in the visual field, etc. Similarly, a neuron that responded only to, say, blue squares would represent a static binding of the color blue to the shape square (see Figure 1a). Conjunctive bindings are static in the sense that the binding is fixed in the identity of the unit that represents it, with different units for different conjunctions (bindings) of the same properties.

The alternative to conjunctive coding is to represent separate properties with separate units, and bind them together dynamically, using some kind of "tag" that is external to the units themselves. For example, one set of units might represent colors and another shapes; a blue square and a red circle would be represented by "tagging" the unit for blue as bound to the unit for square and the unit for red as bound to the unit for circle (see Figure 1b). In principle, many different binding tags are imaginable, but at present, the only neurally plausible tag is based on the use of time. The basic idea is that units that are bound together fire in synchrony with one another, and units that are not bound together fire out of synchrony (Gray & Singer, 1989; Hummel & Biederman, 1992; Strong & Whitehead, 1989; von der Malsburg, 1981/1994). A red circle and a blue square would be represented by unit(s) for red firing in synchrony with units for circle, while units for blue fire in synchrony with units for square (and out of synchrony with red and circle; Figure 1c). A red square and a blue circle would be represented by the very same units, but the synchrony relations would be reversed (Figure 1d). Numerous artificial neural networks use synchrony for binding, and there is a growing body of evidence for the role of synchrony for binding in biological nervous systems (see Singer & Gray, 1995, and Singer, 2000, for reviews).

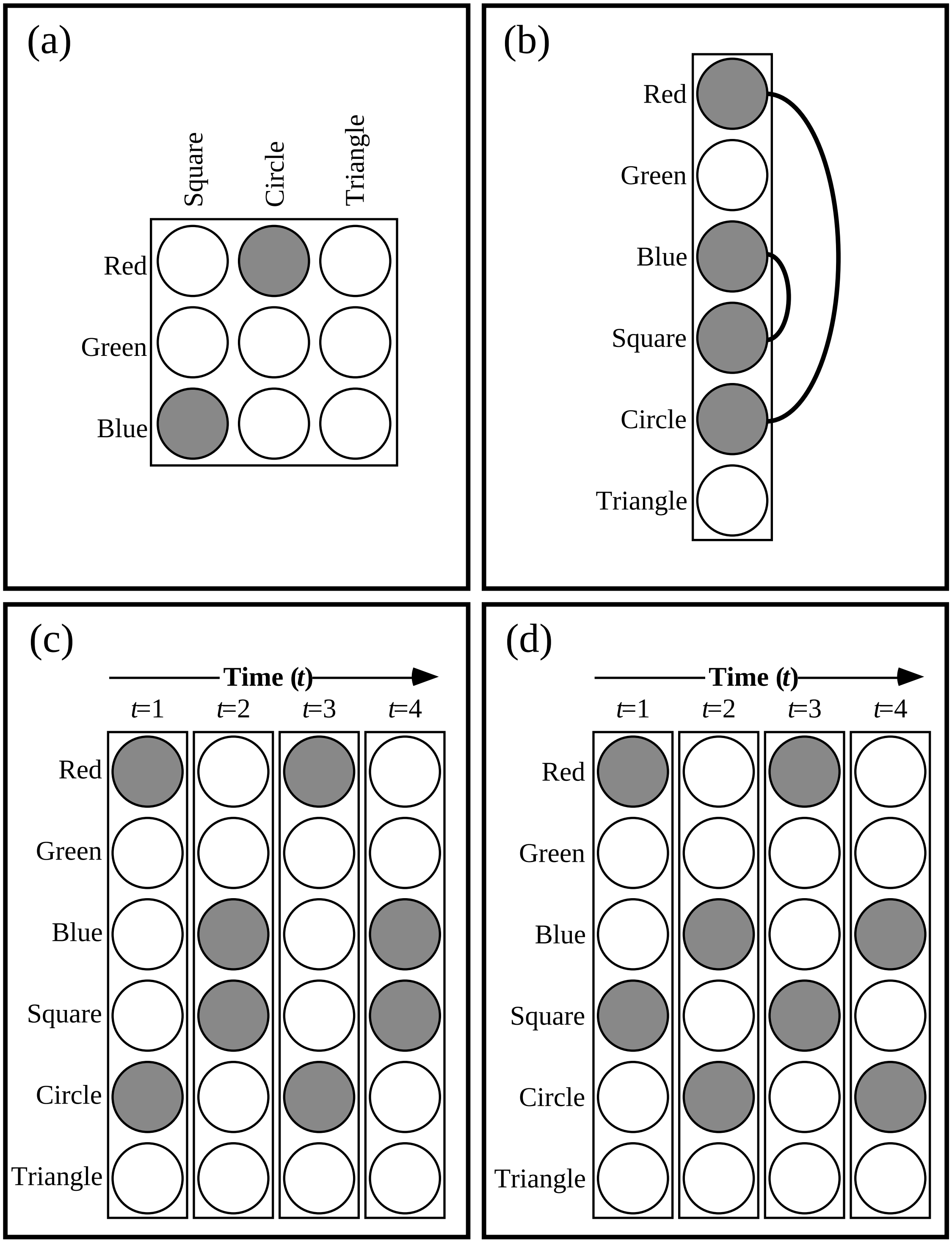

Figure 1. Illustration of static and dynamic binding. Gray circles represent active units, and white circles inactive units. (a) Representation of a red circle and a blue square using a static binding of color and shape. (b) Representation of a red circle and a blue square using a dynamic binding of color and shape, where arcs serve as binding tags. (c) Representation of a red circle and a blue square using synchrony of firing for dynamic binding. Successive panels of circles denote the same units at different points in time. (d) Representation of a red square and a blue circle using synchrony for binding.

The Complementary Properties of Static and Dynamic Binding

As summarized in Table 1, static and dynamic binding have strikingly complementary properties, which as detailed in the next section, map directly onto the properties of holistic and analytic representations, respectively. Perhaps the most obvious difference between them concerns how they pay the computational "cost" of binding. Static binding (i.e., conjunctive coding) pays the cost "up front" at the level of representation: Static binding designates a separate unit for every possible conjunction of properties. For example, representing image edges at eight different orientations, three spatial frequencies and five velocities at each of 100 X 100 locations would require 8 X 3 X 5 X 100 X 100 = 1,200,000 conjunctive units. The same eight orientations, three spatial frequencies, five velocities and 10,000 locations, represented independently and bound together dynamically, would require only 8 + 3 + 5 + 10,000 = 10,016 units. And if the horizontal and vertical dimensions were represented independently of one another, then the number of units would drop to 8 + 3 + 5 + 100 + 100 = 216.

Table 1.

	Static Binding (Holistic)	**Dynamic Binding (Analytic)**
Number of units required	grows geometrically with the number of properties to be bound	grows linearly with the number of properties to be bound
Capacity limits on processing	few or none: permits substantial parallel processing	potentially severe: permits little parallel processing
Role of attention	largely automatic	requires attention
Independence of bound properties	bound properties are represented conjunctively	bound properties are represented independently
Mapping onto Garner's (1974) terminology	dimensions are integral	dimensions are separable

Representing visual properties independently and binding them together dynamically thus requires geometrically fewer units than a static binding of the same properties. When the number of dimensions or properties becomes large, the number of units required to span the space of conjunctions can quickly become prohibitive. For example, the number of units required to statically bind elements to their interrelations grows exponentially with the number of relations (Saiki & Hummel, 1998a), making conjunctive coding completely impractical as an approach to representing the spatial relations among an object's features or parts. It is thus no coincidence that models of shape perception based exclusively on conjunctive coding (e.g., models in the view-based tradition) also represent objects in terms of their features' coordinates rather than their interrelations (for reviews, see Hummel, 2001, Hummel & Stankiewicz, 1996b).

In contrast to static binding, which pays the computational cost of binding in units, dynamic binding pays the cost in processing time. The reason is that dynamic binding is inherently capacity-limited and thus necessarily entails serial processing. For example, in the case of binding by synchrony, it is only possible to have a finite number of groups of neurons simultaneously active and firing out of synchrony with one another. There is no necessary limit on the number of neurons in any one synchronized group; the limit is on the number of groups. The size of this limit is proportional to the length of time between successive bursts of spikes from a given group of neurons divided by the temporal width of each burst (or, equivalently, on the temporal precision of a single neuron's spiking). A reasonable estimate of this limit is four to six separate groups (see Singer, 2000; Singer & Gray, 1995), a number that corresponds very closely to the observed capacity limits of visual attention and visual working memory (Bundesen, 1998; Luck & Beach, 1998; Luck & Vogel, 1997; see Hummel & Holyoak, submitted, for a review).

The limited capacity of dynamic binding implies that some tasks must be performed sequentially. If a given task entails processing, say, 12 bindings, but the visual system can only represent four of those bindings at a time, then the task will require at least three processing steps (see Hummel & Holyoak, 1997). The capacity limit also makes dynamic binding dependent on attention to control which of several potentially conflicting groups are bound at any given time (Hummel & Stankiewicz, 1996a; see also Logan, 1994; Stankiewicz, Hummel & Cooper, 1998; Treisman & Gelade, 1980). By contrast, although static binding requires a large number of units, the fact that separate conjunctions are represented by separate units makes it possible, at least in principle, to process many different conjunctions in parallel: 1,200,000 separate orientation-frequency-motion-location conjunction detectors may seem like a large number of units, but it may be a worthwhile investment if the alternative is to process 10,000 separate locations one at a time.

A second difference between dynamic and static binding—at least as important as their complementary costs—is that dynamic binding makes it possible to represent the bound properties independently of one another, whereas static binding does not. For example, a static binding of shapes to colors in which each unit codes for a specific color shape conjunction (e.g., blue-and-square, blue-and-round, red-and-square, red-and-round, etc.) does not represent color independently of shape. As a result it does not specify what a blue square has in common with a blue circle, or that a blue square has more in common with a blue circle than with a red circle: They are all simply different units (see Figure 1a). Although this example is based on a strictly localist conjunctive code (i.e., with one unit per binding), Hummel and Stevens (in preparation; Holyoak & Hummel, 2000) demonstrate that this property is a mathematical inevitability of any conjunctive representation, including distributed conjunctive codes such as tensor products (Halford et al. 1994; Smolensky, 1990), circular convolutions (Metcalfe, 1990, 1991), holographic reduced representations (Plate, 1993) and recursive auto-associative memories (Pollack, 1990): Conjunctive representations necessarily violate the independence of the bound properties. By contrast, dynamic binding makes it is possible to represent properties independently (i.e., assigning units to individual properties, such as blue and square, rather than conjunctions, such as blue-and-square) and still specify how properties go together. The resulting representation explicitly specifies what a blue square has in common with a red square and how they differ.

Whether preserving independence is desirable or not depends on the goals of the computation. The advantage of violating independence (as in the case of static binding) is that it makes it possible to make decisions: If category A is defined as things that are both blue and square, then a unit that binds blue to square statically can unambiguously discriminate A's from not-A's. An independent representation with color on one unit and shape on another would not discriminate them as cleanly. For example, a blue circle, which is not a member of A, would activate the "blue" part of the representation of A. The advantage of preserving independence is that it makes it possible for a representation to use (and re-use) the same elements in different combinations, and to appreciate what different bindings of the same elements have in common and how they differ. Effectively, the capacity for dynamic binding turns a representation into a symbol system (Hummel & Holyoak, 1997). The difference between symbolic and non-symbolic representational systems is precisely that the former, but not the latter, can compose simple elements (symbols) into complex relational structures without losing the individual symbols in the process (Fodor & Pylyshyn, 1988; Newell, 1980; see also Hummel & Holyoak, 1997, submitted; Hummel & Stevens, in preparation).

The complementary properties of static and dynamic binding suggest a rule of thumb for the design of a visual system: Conjunctions that can be exhaustively enumerated a priori, and those that need to be processed in parallel (such as the kinds of visual properties processed in visual area

V1), should be bound together statically; by contrast, complex properties and properties that need to be represented explicitly—especially those that need to be represented in terms of a complex and potentially open-ended vocabulary of relations—should be represented independently and bound together dynamically. These rules of thumb are apparent in extant models of object recognition. View-based models, whose representations are based exclusively on simple features duplicated at each of many locations in the visual field, use static binding exclusively (i.e., with no use of dynamic binding). By contrast, structural description models, whose representations specify complex interrelations (including relative location, relative size, etc.) among complex features (e.g., the shape attributes of convex parts), must use a combination of both static binding (at the level of local image features) and dynamic binding (at the level of complex features and relations).

Binding and the Holistic/Analytic Distinction

The terms holistic and analytic may mean slightly different things to different people. I will use holistic to refer to representations that are generated and used "all of a piece"—representations that are difficult or impossible to evaluate in terms of their constituent dimensions or parts. An analytic representation is simply the opposite. It is a representation that affords analysis in terms of its constituent parts and their interrelations. The distinction between face recognition on the one hand and common object recognition on the other illustrates the difference between holistic and analytic representations. The visual representations that permit us to discriminate one person's face from another's are holistic, as evidenced by the fact that it is often easy to say who looks like whom, but barring obvious markings such as facial hair and scars, it is usually difficult to say why (Bartlett, Searcy & Abdi, this volume; Cooper & Wojan, 2000; Farah, 1990; Murray, Rhodes, & Schuchinsky, this volume; Palermo & Rhodes, in press; Tanaka & Farah, 1993). By contrast, the visual representations that allow us to recognize objects as members of a general class—including those that allow us discriminate faces from non-face objects (Cooper & Wojan, 2000)—are represented much more analytically, as evidenced by the ease with which we can say what objects have visually in common and how they differ (Hummel, 2000; Saiki & Hummel, 1998a; Thoma & Hummel, submitted).

Another example of holistic and analytic representations can be found in the relationship between color and shape. Color is defined by the physical dimensions of saturation, brightness and hue. These dimensions are physically independent but they are psychologically integral, in that people have great difficulty making judgments about one without interference from the others (Garner, 1974; see also Cheng & Pachella, 1984): To the visual system, saturation, brightness and hue are the single holistic dimension color. By contrast, it is possible to make judgments about an object's shape with little or no interference from its color, and vice versa: Color and shape are psychologically separable.

As these examples illustrate, holistic (i.e., integral) and analytic (i.e., separable) are better described as relations among the dimensions defining a representation than as properties of a representation as a whole. The dimensions defining color are holistic (integral) with respect to one another, but they are analytic (separable) with respect to the dimensions defining shape. Similarly, some dimensions of the visual representation of shape are analytic with respect to one another, whereas others are holistic. Object parts are visually independent of (i.e., analytic with respect to) their spatial relations (Saiki & Hummel, 1998a), but the dimensions that permit us to discriminate one person's face from another's are holistic with respect to one another.

From a computational perspective, the distinction between holistic and analytic dimensions maps directly onto the distinction between static (conjunctive) and dynamic binding. In a given representation, two dimensions will be holistic with respect to one another to the extent that they are bound together statically (i.e., conjunctively) in that representation; they will be analytic to the extent that they are represented independently and bound together dynamically.[1] For example, if every

[1] For the purposes of storage in long-term memory, even independent dimensions must at some point be bound together statically (see Hummel & Biederman, 1992). As such, it is perhaps more accurate to say that two

unit in a given representation codes for a particular feature at a particular location (like the neurons in V1), then to that representation, feature identity is holistic with respect to location. Likewise, if every unit codes for some combination of brightness, saturation and hue, then to that representation, brightness, saturation and hue are holistic with respect to one another.

The Strengths and Limitations of Holistic and Analytic Representations

As approaches to the representation of shape, holistic and analytic representations have complementary strengths and weaknesses that reflect their complementary solutions to the binding problem. In brief, analytic representations are flexible, expressive and representationally efficient (i.e., many conjunctions can be represented with few units), but computationally "expensive" (i.e., limited in the number of conjunctions that can be processed in parallel). Holistic representations have limited flexibility and expressiveness, and they are representationally “expensive,” (requiring a separate unit for each conjunction) but computationally "inexpensive" (permitting the processing of many conjunctions in parallel).

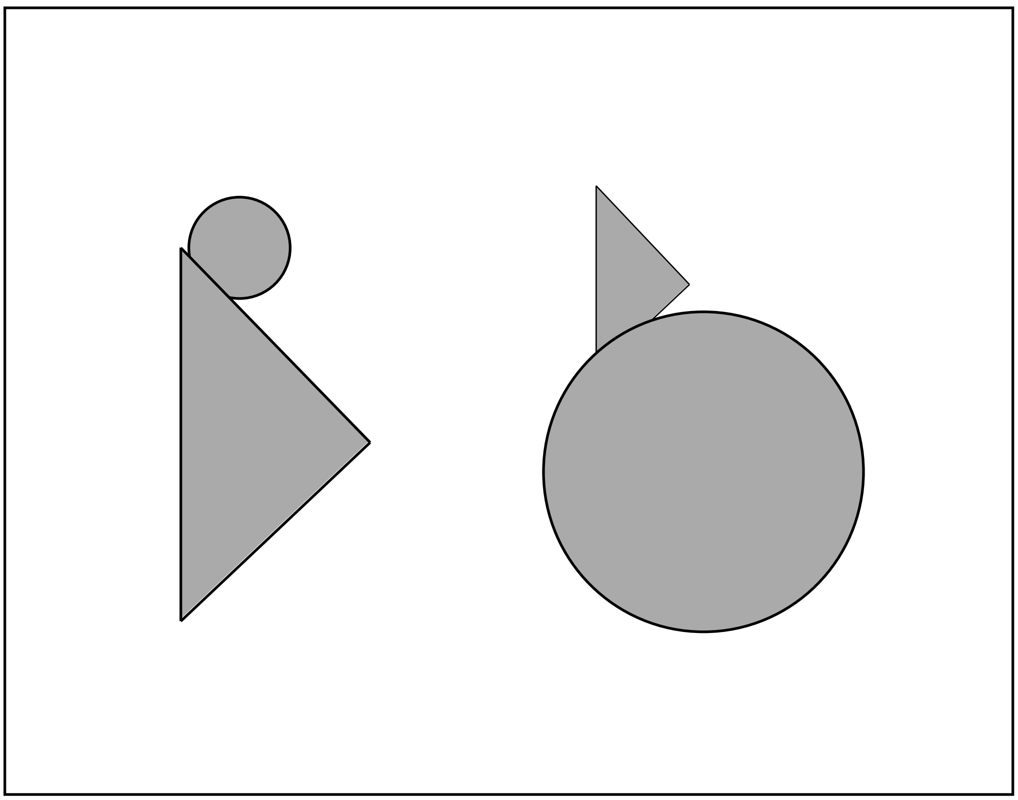

Figure 2. Two simple "objects." It is easy to appreciate that the circle on the left corresponds to the circle on the right in terms of its shape, but to the triangle on the right in terms of its spatial relations. Your ability to appreciate these alternative correspondences is a manifestation of your ability to perceive object parts independently of their spatial relations (Hummel, 2000).

Consider first the case of analytic representations. The flexibility of an analytic representation derives from its ability to represent things independently of one another. An animal

---

dimensions will be psychologically holistic to the extent that the visual system only binds them statically (i.e., never represents them independently), and they will be analytic to the extent that it ever represents them independently.

that can represent object shape independently of location in the visual field can respond to an object in the same way regardless of where its image projects to the retina. Similarly, representing an object's parts independently of their interrelations makes it possible to reason flexibly about those parts and relations. For example, consider the two simple "objects" in Figure 2. Does the circle in the left-hand object correspond to the circle or the triangle in the right-hand object? In terms of shape, the circle on the left corresponds to the circle on the right; but in terms of relative size and location, the circle on the left corresponds to the triangle on the right. Your ability to appreciate these alternative correspondences stems directly from your ability to represent the shapes independently of their spatial relations, and vice-versa (Hummel, 2000).

There is substantial evidence for the role of analytic representations in human shape perception and object recognition. Using methods similar to those Garner (1974) used to demonstrate the perceptual separability of shape and color, Stankiewicz (in press) showed that our visual systems represent the dimensions defining an object's 3D shape (e.g., its aspect ratio and the curvature of its major axis) independently of the dimensions defining the angle from which it is viewed. This result suggests that our ability to recognize objects in novel viewpoints reflects, at least in part, our ability to perceptually separate the dimensions that define object shape from those that define viewpoint (see also Biederman, 1987; Biederman & Cooper, 1991a, 1992; Biederman & Gerhardstein, 1993, 1995; Hummel & Biederman, 1992). Similarly, Saiki and Hummel (1998a) showed that the human visual system represents an object's parts independently of their spatial relations (see also Logan, 1994), and Hummel and Stankiewicz (1996b) showed that, at least for attended stimuli, the visual representations that mediate object recognition, same-different judgments and similarity judgments are based on explicit representations of the categorical relations among an object's parts (see also Saiki & Hummel, 1998b). By demonstrating that we perceive the attributes of object shape—including the spatial relations among an object's parts—both explicitly and independently of one another, these findings constitute direct evidence for the role of analytic representations in shape perception.

Indirect evidence for the role of analytic representations in shape perception comes from a variety of experiments supporting the behavioral predictions of specific structural description theories. These findings provide indirect support in the sense that they are not addressed to the analytic/holistic distinction, per se, but rather to various other predictions (mostly concerning the role of viewpoint in object recognition) of specific models that postulate analytic representations. Specifically, as predicted by Biederman's (1987; Hummel & Biederman, 1992) structural description theory, our ability to recognize an object is insensitive to the location of its image in the visual field (Biederman & Cooper, 1991a), the size of the image (Biederman & Cooper, 1992) and left-right reflection (Biederman & Cooper, 1991a; Stankiewicz et al., 1998). Object recognition (Biederman, 1987) and visual priming for objects (Biederman & Cooper, 1991b) also appear to be mediated by an explicit representation of volumetric parts.

In summary, the strength of an analytic representation is that it provides tremendous flexibility and economy of representation, and serves as the foundation for complex symbolic representations and processes. Theories of shape perception based on analytic representations also provide a simple and strikingly complete account of many aspects of human shape perception, object recognition and categorization. The limitation of analytic representations is that they depend on dynamic binding, and dynamic binding imposes a bottleneck on processing. As noted previously, dynamic binding is time-consuming and capacity-limited, and there is substantial evidence that it requires visual attention (Hummel & Stankiewicz, 1996a; Logan, 1994; Luck & Beach, 1998; Luck & Vogel, 1997; Treisman & Gelade, 1980).

The limitations of dynamic binding are problematic for theories based solely on analytic representations of shape (e.g., those of Biederman, 1987, Hummel & Biederman, 1992, and Marr & Nishihara, 1978). The process of generating an analytic representation cannot be any faster or more automatic than the process of dynamic binding, but object recognition apparently can be. Face recognition in the macaque can be accomplished to a high degree of certainty based on the first set of spikes to reach inferotemporal cortex (at least for over-learned faces; Oram & Perrett, 1992). Clearly, the macaque visual system recognizes faces without waiting around for several sets of desynchronized spikes. Although Oram and Perrett obtained their data using faces (which are known to be processes holistically) as stimuli, there is also behavioral evidence that people can recognize common (non-face) objects very rapidly. Intraub (1981) showed that people can recognize common objects presented at the rate of ten per second (see also Potter, 1976). These findings suggest that object recognition is much too fast to depend exclusively on dynamic binding (Hummel & Stankiewicz, 1996a). Similarly, although dynamic binding—and therefore analytic representations—require visual attention, object recognition apparently does not, as shown by findings of both negative priming (e.g., Tipper, 1985; Treisman & DeSchepper, 1996) and positive priming (Stankiewicz, et al., 1998; Stankiewicz & Hummel, in press) for unattended objects.

Holistic representations (based on static binding) are the complement of analytic representations (based on dynamic binding) in both their strengths and weakness. The strength of a holistic representation is that it carries bindings statically (i.e., conjunctively) in the units of representation, making dynamic binding unnecessary. Holistic representations can therefore be generated automatically and with minimal cost in terms of attention or working memory (Thoma & Hummel, submitted). The limitation of a holistic representation is that it lacks the flexibility and expressiveness of an analytic representation. For example, to a fully holistic representation of shape, the two objects in Figure 2 are simply different. It does not even make sense to ask whether the circle on the left corresponds to the circle on the right: There is no "circle on the left", just the shape as a whole. A related limitation is that, because they bind elements to their relations conjunctively, holistic representations are sharply limited in the ways they can represent relations. As noted previously, the number of pairwise relations among an object's elements grows exponentially with the number of relations. As a result, holistic representations, although not logically constrained, are pragmatically constrained to represent features in terms of their coordinates, rather than their interrelations (see Hummel & Stankiewicz, 1996b). In turn, the exclusive use of coordinates causes holistic representations to resemble metrically precise templates, which are less flexible and less expressive than explicit structural descriptions (Clowes, 1967; Sutherland, 1968).

It is worth noting in this context that the deep limitation of the view-based approach to object recognition (e.g., Edelman, 1998; Poggio & Edelman, 1991; Tarr & Bülthoff, 1995; see also Tarr, this volume, and the references therein) is that it relies exclusively on holistic representations of shape, making it fundamentally inconsistent with the structural aspects of shape perception reviewed earlier (Hummel, 1994, 2000, 2001; Hummel & Stankiewicz, 1996b). Although Tarr's critique of the structural description approach to recognition (which he identifies primarily with Marr & Nishihara, 1978, and Biederman, 1987) emphasizes the role of structural descriptions qua "3D view-invariant representations," the fundamental difference between the view- and structure-based accounts of object recognition concerns, not the role of viewer- vs. object-centered representations, but the role of holistic vs. structured (i.e., explicitly relational) representations of shape (Hummel, 1994; Hummel & Biederman, 1992; Hummel & Stankiewicz, 1996b; see Hummel, 2000, for a review). For example, the structural descriptions generated by the models of Hummel and Biederman (1992) and Hummel and Stankiewicz (1996a, 1998; Hummel, 2001) specify the spatial relations among object parts in viewer-centered coordinates. One reason is that, as detailed by Hummel (1994), representing spatial relations explicitly makes it possible to account for the view-invariances and sensitivities that characterize human object recognition without having to postulate 3D object-centered representations, and without having to postulate complex "alignment," "normalization," or "transformation" operations to bring viewed images into correspondence with stored views.

EXPLOITING THE STRENGTHS OF HOLISTIC AND ANALYTIC REPRESENTATIONS

This survey of the strengths and limitations of holistic and analytic representations suggests that any visual system that relied exclusively on one or the other as a basis for representing shape would have serious limitations. To a visual system that could only represent shape analytically, shape perception would be slow and laborious, always requiring visual attention and working memory. To a visual system that could only represent shape holistically, the visual world would be a mysterious place, full of objects that were either vaguely similar or vaguely dissimilar, but without any basis for saying why things seemed similar or not.[2]

In response to the complementary strengths and limitations of analytic and holistic accounts of shape perception, Hummel and Stankiewicz (1996a; Hummel, 2001) proposed a model of object recognition based on a hybrid analytic+holistic representation of shape. The basic idea starts with the observation that holistic representations based on static binding are not capacity-limited in the way that analytic representations are. The model uses dynamic binding to generate analytic representations whenever possible (e.g., when an object image is attended), but uses static binding to generate useful (if limited) holistic representations when dynamic binding fails (e.g., when an image in not attended). Importantly, the holistic (statically-bound) and analytic (dynamically-bound) components of the representation are not simply "joined at the hip." Rather, the holistic component represents the substructure of whatever subset of the analytic representation happens to be the focus of processing at a given instant. If the "subset" happens to be the entire object—i.e., when the visual system fails to segment the object into independent parts, as when the image is not attended—then the "substructure" will be a holistic representation of the entire object. This holistic representation lacks the flexibility and view-invariance of the analytic representation, but is capable of supporting recognition if the object is depicted in a familiar viewpoint. When the visual system succeeds in segmenting the object into parts, the holistic component represents the substructure (i.e., local features [Hummel & Stankiewicz, 1996a] or surface characteristics [Hummel, 2001]) of each part individually.

The analytic component of the model's representation consists of a collection of units that represent the shape attributes of an object's parts independently of one another, of their locations in the visual field, and of their interrelations. That is, each unit responds to one shape attribute or relation, and will respond to that attribute or relation regardless of the part's (or object's) other attributes and relations. As a result, the representation generated on these units is invariant with translation, scale and left-right reflection (although, like human shape perception, it is sensitive to rotations about the line of sight and some rotations in depth). However, this same independence makes these units heavily dependent on dynamic binding: If more than one of an object's parts fires at a given instant, then the description generated on the independent units will be a meaningless jumble of unbound features (see Hummel & Stankiewicz, 1996a). The holistic component of the representation specifies the local features of each object part (or, when segmentation fails, the entire object) separately at each of several locations in a circular reference frame. That is, each unit responds to a specific feature at a specific location in the reference frame (i.e., a feature-location conjunction). The origin and scale of this reference frame are defined on whatever part (or object) it happens to be representing at the time, but its orientation is defined relative to the viewer. That is, although it is holistic, this representation is not completely viewer-centered: As a result, it is sensitive to rotation and left-right reflection but it is invariant with translation and scale.

The resulting model accounts for a large body of existing findings on human object recognition (including its view-invariances and sensitivities, as well as the role of parts [Biederman & Cooper, 1991b] and explicit spatial relations [Hummel & Stankiewicz, 1996b] in shape perception) and generates several novel predictions. Most generally, it predicts that when the visual system succeeds in segmenting an object into its parts (i.e., when it succeeds in generating an analytic representation of the parts and their interrelations), shape perception will have the characteristics of a structural description: Recognition will be largely invariant with variations in viewpoint, and part attributes will be represented independently of one another and their

---

[2] There is some evidence that this is the visual world experienced by pigeons (Peissig, et al., 1999; Peissig, et al., in press; Sekuler, Lee & Shuttleworth, 1996). It is not, however, the world experienced by people. In this context, it is interesting to speculate that analytic representations may be a late evolutionary development. Holistic representations are easy to learn: It is only necessary to learn conjunctions of more primitive elements. By contrast, analytic representations, which treat dimensions independently of one another, are very difficult to learn unless you start out with a representation that can already treat the relevant dimensions independently—i.e., unless you start with a representation that is already analytic (Goldstone, 2001; Kellman, Burke & Hummel, 1999).

interrelations. When the visual system fails to segment an image into its parts, shape perception will have the characteristics of the holistic representation: Recognition will be more sensitive to variations in viewpoint, and part attributes will not be represented independently of their locations.

These general predictions can be translated into the following more specific predictions about patterns of visual priming: (1) Attending to an object's image will visually prime (1a) that image, (1b) translated versions of that image (as when the image is moved from the left-hand side of the visual field to the right-hand side), (1c) scaled (larger or smaller) versions of that image, (1d) a left-right reflection of that image, and (1e) split versions of the image (as when the image is cut horizontally and the top half is placed below the bottom half, or cut vertically, and the left half placed to the right of the right half). Predictions 1a - 1d correspond directly to the view-invariances of the analytic representation: Visual priming is expected to show the same invariances as the representations primed (see, e.g., Biederman & Cooper, 1991a). Prediction 1e follows from the fact that the units in the analytic representation are independent of their locations in the visual field and relations to one another. Imagine a split image of, say, a horse, with the left half of the image moved to the right, so that the horse is split in half, facing its own hind quarters. According to the model, attending to this image will prime the units representing the shape attributes of the horse's parts. Since these units are indifferent to the parts' locations, this priming will transfer to an intact version of the same image (in which the locations and interrelations are changed relative to their locations in the split image). Hence, for attended images, priming is predicted to transfer from a split image to its intact counterpart.

(2) Viewing an image <u>without</u> attending to it will visually prime (2a) that image, (2b) a translated version of that image, and (2c) a scaled version of that image, but it will <u>not</u> prime (2d) a left-right reflection. Moreover, (2e) an unattended split image will not prime its intact counterpart. Predictions 2a - 2d correspond to the view-invariances and -sensitivities of the holistic representation. (Recall that the origin and scale of the holistic representation are defined relative to the object [or, more accurately, whatever subset of the object happens to be firing at a given instant], so the representation generated there is invariant with translation and scale.) Prediction 2e follows from the fact that each unit in the holistic representation responds to a particular feature at a particular location. Thus, because splitting an image changes the locations of its features, both in the image as a whole and relative to one another, the theory predicts that priming an ignored split image will not transfer to its intact counterpart. Hummel (2001) reports simulation results corresponding to each of these predictions.

Brian Stankiewicz and his colleagues tested predictions 1a - 1d and 2a - 2d, and the results were exactly as predicted. Attended images visually primed themselves (Stankiewicz et al., 1998; Stankiewicz & Hummel, in press), translated and scaled versions of themselves (Stankiewicz & Hummel, in press) and their left-right reflections (Stankiewicz et al., 1998). Images viewed without attention visually primed themselves (Stankiewicz et al., 1998; Stankiewicz & Hummel, in press) and translated and scaled versions of themselves (Stankiewicz & Hummel, in press), but not left-right reflection of themselves (Stankiewicz et al., 1998). Moreover, the effects of viewpoint (same view vs. translated, scaled or reflected) and attention (attended vs. ignored) were strictly additive in all these cases: The advantage in priming enjoyed by identical images over non-identical images was the same for both attended and ignored images (about 50 ms in the case of left-right reflection and 0 ms in the case of translation and scale changes). This additivity is consistent with the model's claim that different representation mediate priming in the attended and ignored cases (see Stankiewicz et al., 1998).

Thoma and Hummel (submitted) tested predictions 1e and 2e and the results were again exactly as predicted. Both intact and split attended images primed intact versions of themselves, as did ignored intact images. However, ignored split images did not prime intact versions of themselves: Consistent with the hypothesis that ignored images are represented holistically, priming for ignored images did not transfer across deviations in the (object-relative) locations of the objects' features. As in the case of left right reflection, these effects were strictly additive, with identical images enjoying about 50 ms more priming than split primes in both the attended and ignored conditions. These findings are particularly interesting in the context of the distinction between analytic and holistic representations. The discussion in this chapter of the relation between analytic representations, dynamic binding and visual attention predicts that attended images should be represented analytically, and ignored images should be represented holistically. The Thoma and Hummel findings directly support this prediction.

A Hierarchy of Shape Representations

Up to this point, this chapter has focused on the differences between holistic and analytic representations of shape, and their relationship to the differences between static and dynamic solutions to the binding problem. One general way to summarize these differences is to say that analytic representations are structurally richer than holistic representations: Dynamic binding makes it possible to represent visual properties, features, parts and relations independently of one another and still specify how they are bound together; in turn, this capacity makes it possible to represent arbitrarily complex relations (including hierarchical relations), and ultimately serves as the foundation for symbolic representations and processes. Holistic representations, lacking any basis for dynamic binding, lack this kind of flexibility and structural sophistication.

But this does not imply that holistic representations are devoid of relational information. Relational structures are present in holistic representations, but they are implicit: After all, the visual machinery that generates analytic representations of object shape must start with a holistic representation (such as the pattern of activation in V1) as input. Thus, although analytic representations are structurally “less sophisticated” than analytic representations, they are nonetheless more sophisticated than representations that are completely devoid of relational information (such as a simple list of features bound to neither relations nor locations).

Consistent with this observation, Palermo and Rhodes (in press) showed that attention plays an important role in generating the holistic representations that allow us to discriminate one face from another. Specifically, they showed that subjects were better able to recognize an isolated feature of a target face (either the eyes, nose or mouth) when that feature was embedded in the context of the complete face (the context in which it was presented) than when it was presented in isolation, but that this beneficial effect of context obtained only when the target face was attended. The beneficial effect of the context (i.e., the complete face) in the attended condition is a replication of Tanaka and Farah (1993), and is evidence for holistic processing in that condition. Accordingly, the lack of any context effect in the unattended condition is evidence of a lack of holistic processing in that condition. These data suggest that attention is important for the holistic processing of faces, a finding that appears to at first blush be at odds with the findings of Thoma and Hummel and those of Stankiewicz and colleagues.

One potential explanation for the apparent discrepancy between the findings (and conclusions) of Palermo and Rhodes and those of Thoma, Stankiewicz and colleagues concerns the differences between the task of discriminating faces and the task of recognizing objects. Faces are structurally very similar, so the (putatively) holistic representations we use to discriminate them must be metrically very precise. By contrast, the holistic representations postulated by the Hummel and Stankiewicz (1996a; Hummel, 2001) model are metrically very coarse, distinguishing only 17 different locations in a circular reference frame (the center, plus two distances from the center at each of eight orientations). It is entirely possible that the metric precision required for face discrimination requires visual attention (see La Berge & Brown, 1989, and Hummel & Stankiewicz, 1998, for discussions of the role of visual attention in generating metrically precise representations of shape).

Another potential explanation for the apparent discrepancy between the findings of Palermo and Rhodes and those of Thoma, Stankiewicz and colleagues—not mutually exclusive with the first—concerns the difference between holistic representations on the one hand and collections of “free floating” features (i.e., features bound to neither locations or relations) on the other. As noted previously, holistic representations carry structural information, albeit implicitly. Free floating features do not. From an information theoretic perspective, holistic representations are therefore a great deal richer than collections of free floating features, so from a sampling perspective, they should be proportionally more difficult to generate. It is conceivable that Palermo and Rhodes’ subjects represented the target faces more holistically in the attended conditions, and more as free-floating features in the unattended conditions. But whatever the explanation for the apparent discrepancy between the Palermo and Rhodes findings and those of Thoma, Stankiewicz and colleagues, it seems clear that, all other things being equal, analytic representations are more attention-demanding than holistic representations, and holistic representations may be more demanding than collections of free-floating features.

Other Factors Affecting Holistic and Analytic Representations of Shape

This chapter has focused on the role of attention in generating analytic representations of object shape. However, numerous other factors are likely to affect whether the visual system will generate an analytic representation of an object's shape, and some of these deserve mention. One is time. Representing an object's shape analytically entails segmenting the object's image into its constituent features or parts. The operations that perform this kind of segmentation depend on lateral communication between neurons representing local features (such as lines, vertices, etc.) in an object's image (Hummel & Biederman, 1992), and probably also exploit feedback from representations of known objects to representations of local image features (Peterson & Gibson, 1994). Such interactions take time, and imply that analytic representations cannot be generated in any strictly feed-forward manner (Hummel & Stankiewicz, 1996a). Behaviorally, these considerations predict that recognition based on an analytic representation of shape should take longer than recognition based on a holistic representation (at least for the simple kinds of holistic representations postulated by Hummel & Stankiewicz, 1996a; holistic representations like those studied by Palmero & Rhodes, in press, may, like analytic representations, be time-consuming to generate). To my knowledge, this prediction remains to be tested.

A related factor that stands to affect whether the visual system will succeed in generating an analytic representation of an object's shape is whether the object's image presents the local cues (e.g., matched concavities; Hoffman & Richards, 1985) necessary for segmentation (Hummel & Biederman, 1992). Very "blooby" or irregular figures (such as lumps of clay) may afford few such cues, and may therefore be represented more holistically than figures that afford natural segmentation into features or parts. Finally, as suggested by the case of face recognition, there is reason to believe that discriminating one member of a class from other qualitatively similar instances may profit from metrically-precise holistic coding (but see Hummel & Stankiewicz, 1998, for an alternative account of some kinds of subordinate-level recognition). But whether the goal is to permit recognition of unattended images, to permit rapid (potentially feed-forward) recognition of objects in familiar views, to permit recognition of objects that do not afford part-based analysis, or to permit recognition of individuals based on fine metric differences, holistic coding affords numerous benefits as a complement to analytic coding.

SUMMARY AND CONCLUSION

A representation of object shape is analytic (i.e., structured) to the extent that it represents the attributes of that shape (e.g., features or part attributes) independently of one another and of their interrelations. A representation is holistic to the extent that it codes attributes and relations conjunctively, rather than independently. Analytic and holistic representations have strikingly complementary strengths and limitations as representations of object shape. Representing the attributes of shape independently of one another affords tremendous flexibility and expressiveness, but due to the resulting need for dynamic binding, incurs a cost in terms of attention, working memory and processing time. Binding the dimensions of a stimulus conjunctively, as in a holistic representation, does not incur the computational costs of an analytic representation, but neither does it afford the same flexibility or expressiveness. These considerations suggest that a "well designed" visual system would be configured to exploit the strengths of both analytic and holistic representations in order to minimize the deleterious effects of the weakness of either. The experimental findings reviewed here suggest that the human visual system is indeed configured to exploit the strengths of an analytic representation when it attends to a stimulus, and simultaneously enjoy the advantages of a holistic representation when it does not.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-147.

Biederman, I., & Cooper, E. E. (1991a). Evidence for complete translational and reflectional invariance in visual object priming. Perception, 20, 585-593.

Biederman, I. & Cooper, E. E. (1991b). Priming contour deleted images: Evidence for intermediate representations in visual object recognition. Cognitive Psychology, 23, 393-419.

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. Journal of Experimental Psychology: Human Perception and Performance, 18, 121-133.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for 3-dimensional viewpoint invariance. Journal of Experimental Psychology: Human Perception and Performance, 19, 1162-1182.

Biederman, I. & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: A critical analysis. Journal of Experimental Psychology: Human Perception and Performance, 21, 1506-1514.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. Cognitive Psychology, 20, 38-64.

Bundesen, C. (1998). Visual selective attention: Outlines of a choice model, a race model and a computational theory. Visual Cognition, 5, 287-309.

Cheng, P. W., & Pachella, R. G. (1984). A psychophysical approach to dimensional separability. Cognitive Psychology, 16, 279-304.

Clowes, M. B. (1967). Perception, picture processing and computers. In N.L. Collins & D. Michie (Eds.), Machine Intelligence, (Vol 1, pp. 181-197). Edinburgh, Scotland: Oliver & Boyd.

Cooper, E. E., & Wojan, T. J. (2000). Differences in the coding of spatial relations in face identification and basic-level object recognition. Journal of Experimental Psychology: Learning, Memory, & Cognition, 26, 470-488.

Edelman, S. (1998). Representation is representation of similarities. Behavioral & Brain Sciences, 21, 449-498.

Edelman, S., & Poggio, T. (1991). Bringing the grandmother back into the picture: A memory-based view of object recognition. MIT A.I. Memo No. 1181. April.

Enns, J. T. (1992). Sensitivity of early human vision to 3-D orientation in line-drawings. Canadian Journal of Psychology, 46, 143-169.

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

Fodor, J. A., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. Cognition, 28, 3-71.

Garner, W. R. (1974). The processing of information and structure. Hillsdale, NJ: Erlbaum.

Goldstone, R. L. (2001). The sensitization and differentiation of dimensions during category learning. Journal of Experimental Psychology: General, 130, 116-139.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. Cognitive Psychology, 23, 222-262.

Gray, C. M., & Singer, W. (1989). Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. Proceedings of the National academy of Sciences, USA 86, 1698-1702.

Halford, G. S., Wilson, W. H., Guo, J., Gayler, R. W., Wiles, J., & Stewart, J. E. M. (1994). Connectionist implications for processing capacity limitations in analogies. In K. J. Holyoak & J. A. Barnden (Eds.), Advances in connectionist and neural computation theory, Vol. 2: Analogical connections (pp. 363-415). Norwood, NJ: Ablex.

Hoffman, D. D., & Richards, W. A. (1985). Parts of recognition. Cognition, 18, 65-96.

Holyoak, K. J., & Hummel, J. E. (2000). The proper treatment of symbols in a connectionist architecture. In E. Dietrich and A. Markman (Eds.). Cognitive Dynamics: Conceptual Change in Humans and Machines (pp. 229 - 264). Hillsdale, NJ: Erlbaum.

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 (pp. 157 - 185). Hillsdale, NJ: Erlbaum.

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. Visual Cognition, 8, 489 - 517.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. Psychological Review, 99, 480-517.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. Psychological Review, 104, 427-466.

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In M. L. Gattis (Ed.), Spatial schemas in abstract thought (pp. 279-305). Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (submitted). A symbolic-connectionist theory of relational inference and generalization.

Hummel, J. E., & Stankiewicz, B. J. (1996a). An architecture for rapid, hierarchical structural description. In T. Inui and J. McClelland (Eds.). Attention and Performance XVI: Information Integration in Perception and Communication (pp. 93-121). Cambridge, MA: MIT Press.

Hummel, J. E., & Stankiewicz, B. J. (1996b). Categorical relations in shape perception. Spatial Vision, 10, 201-236.

Hummel, J. E., & Stankiewicz, B. J. (1998). Two roles for attention in shape perception: A structural description model of visual scrutiny. Visual Cognition, 5, 49-79.

Hummel, J. E.. & Stevens, G. (in preparation). The proper treatment of symbols in a neural architecture.

Intraub, H. (1981). Identification and processing of briefly glimpsed visual scenes. In D. Fisher, R. A. Monty, and J. W. Sender (Eds.). Eye movements: Cognition and Visual Perception (pp. 181-190). Hillside, NJ: Erlbaum.

Kellman, P. J., Burke, T. & Hummel, J. E. (1999). Modeling perceptual learning of abstract invariants. Proceedings of the Twenty First Annual Conference of the Cognitive Science Society (pp. 264-269). Mahwah, NJ: Erlbaum.

LaBerge, D. & Brown, V. (1989). Theory of attentional operations in shape identification. Psychological Review, 96, 101-124.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. Journal of Experimental Psychology: Human Perception and Performance, 20 (5), 1015-1036.

Lowe, D. (1985). Perceptual Organization and Visual Recognition. Kluwer Academic Publishers, Boston.

Luck, S. J., & Beach, N. J. (1998). Visual attention and the binding problem: A neurophysiological perspective. In R. D. Wright (Ed.), Visual Attention (pp. 455-478). New York: Oxford University Press.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390, 279-281.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of three dimensional shapes. Proceedings of the Royal Society of London, Series B. 200, 269-294.

Metcalfe, J. (1990). Composite Holographic Associative Recall Model (CHARM) and blended memories in eyewitness testimony. Journal of Experimental Psychology: General, 119, 145-160.

Metcalfe, J. (1991). Recognition failure and the composite memory trace in CHARM. Psychological Review, 98, 529-553.

Newell, A. (1980). Physical symbol systems. Cognitive Science, 4, 135-183.

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

Palermo, R., & Rhodes, G. (in press). The influence of divided attention on holistic face perception. Cognition.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch and B. B. Lloyd (Eds.) Cognition and Categorization. (pp. 259-303). Hillsdale, NJ: Lawrence Erlbaum.

Pedone, R., Hummel, J. E., & Holyoak, K. J. (2001). The use of diagrams in analogical problem solving. Memory and Cognition, 29, 214-221.

Peissig, J. J., Young, M. E., Wasserman, E. A., & Biederman, I. (1999). The pigeon's perception of depth-rotated shapes. Cahiers de Psychologie Cognitive (Current Psychology of Cognition), 18, 657-690.

Peissig, J. J., Young, M. E., Wasserman, E. A., & Biederman, I. (in press). Seeing things from a different angle: The pigeon's recognition of single geons rotated in depth. Journal of Experimental Psychology: Animal Behavior Processes.

Peterson, M. A., & Gibson, B. S. (1994). Must figure-ground organization precede object recognition? An assumption in peril. Psychological Science, 5, 253-259.

Plate, T. (1991). Holographic reduced representations: Convolution algebra for compositional distributed representations. In J. Mylopoulos and R. Reiter (Eds.), Proceedings of the 12th International Joint Conference on Artificial Intelligence (pp. 30-35). San Mateo: Morgan Kaufmann.

Poggio, T. & Edelman, S. (1990). A neural network that learns to recognize three-dimensional objects. Nature, 343, 263-266.

Pollack, J. B. (1990). Recursive distributed representations. Artificial Intelligence, 46, 77-106.

Potter, M. C. (1976). Short-term conceptual memory for pictures. Journal of Experimental Psychology: Human Learning and Memory, 2, 509-522.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nature Neuroscience, 11, 1019-1025.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7, 573-605.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. Cognitive Psychology, 8, 382-439.

Saiki, J., & Hummel, J. E. (1996). Attribute conjunctions and the part configuration advantage in object category learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 1002-1019.

Saiki, J., & Hummel, J. E. (1998a). Connectedness and part-relation integration in shape category learning. Memory and Cognition, 26, 1138 - 1156.

Saiki, J. & Hummel, J. E. (1998b). Connectedness and the integration of parts with relations in shape perception. Journal of Experimental Psychology: Human Perception and Performance, 24, 227-251.

Sekuler, A. B., Lee, J. A. J., & Shettleworth, S. J. (1996). Pigeons do not complete party occluded figures. Perception, 25, 1109-1120.

Singer, W. (2000). Response synchronization, a universal coding strategy for the definition of relations. In M. S. Gazzaniga (Ed.), The new cognitive neurosciences (2nd edition). Cambridge, MA: MIT Press.

Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. Annual Review of Neuroscience, 18, 555-586.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. Artificial Intelligence, 46, 159-216.

Stankiewicz, B. J. (in press). Evidence for separable dimensions in the perception of object shape. Journal of Experimental Psychology: Human Perception and Performance.

Stankiewicz, B. J., & Hummel, J. E. (in press). Automatic priming for translation- and scale-invariant representations of object shape. Visual Cognition.

Stankiewicz, B. J., Hummel, J. E., & 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, 732-744.

Strong, G. W., & Whitehead, B. A. (1989). A solution to the tag-assignment problem for neural networks. Behavioral and Brain Sciences, 12, 381-433.

Sutherland, N. S. (1968). Outlines of a theory of visual pattern recognition in animals and man. Proceedings of the Royal Society of London (Series B), *171*, 95-103.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 146A, 225-245.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). Journal of Experimental Psychology: Human Perception and Performance, 21, 1494-1505.

Thoma, V. & Hummel, J. E. (submitted). Direct evidence for holistic representations of object shape in the absence of visual attention.

Tipper, S. P. (1985). The negative priming effect: Inhibitory effects of ignored primes. Quarterly Journal of Experimental Psychology, 37A, 571-590.

Treisman, A. & DeSchepper, B. (1996). Object tokens, attention, and visual memory. In T. Inui and J. McClelland (Eds.), Attention and Performance XVI: Information Integration in Perception and Communication. (pp. 15-46). Cambridge, MA: MIT Press.

Treisman, A. & Gelade, G. (1980). A feature integration theory of attention. Cognitive Psychology, 12, 97-136.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. Journal of Experimental Psychology: General, 113, 169-193.

Ullman, S. (1989). Aligning pictoral descriptions: An approach to object recognition. Cognition, 32, 193-254.,

Ullman, S. (1995). High-level Vision: Object Recognition and Visual Cognition. Cambridge MA: MIT Press.

von der Malsburg, C. (1981/1994). The correlation theory of brain function (1994 reprint of a report originally published in 1981). In E. Domany, J. L. van Hemmen, & K. Schulten (Eds.), Models of neural networks II (pp. 95-119). Berlin: Springer.