# Categorization and Object Shape

**Markus Graf**

**Abstract** Categorization is essential for perception and provides an important foundation for higher cognitive functions. In this review, I focus on perceptual aspects of categorization, especially related to object shape. In order to visually categorize an object, the visual system has to solve two basic problems. The first one is how to recognize objects after spatial transformations like rotations and size-scalings. The second problem is how to categorize objects with different shapes as members of the same category. I review the literature related to these two problems against the background of the hierarchy of transformation groups specified in Felix Klein's *Erlanger Programm*. The *Erlanger Programm* provides a general framework for the understanding of object shape, and may allow integrating object recognition and categorization literatures.

## 1   Introduction

Categorization is regarded as one of the most important abilities of our cognitive system, and is a basis for thinking and higher cognitive functions. An organism without such abilities would be continually confronted with an ever-changing array of seemingly meaningless and unrelated impressions. The categorization of environmental experiences is a basic process that must be in place before any organism can engage in other intellectual endeavors. In the words of the cognitive linguist George Lakoff: "There is nothing more basic than categorization to our thought, perception, action, and speech. Every time we see something as a kind of thing, for

M. Graf
Department of Psychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103, Leipzig, Germany and Max Planck Institute for Biological Cybernetics, Tübingen, Germany
e-mail: markus.graf@cbs.mpg.de

example, a tree, we are categorizing. Whenever we reason about kinds of things – chairs, nations, illnesses, emotions, any kind of thing at all – we are employing categories. (…) Without the ability to categorize, we could not function at all, either in the physical world or in our social and intellectual lives. An understanding of how we categorize is central to any understanding of how we think and how we function …" (Lakoff 1987, pp. 5–6).

A fundamental aspect of categorization is to visually recognize and categorize objects. Probably the most important feature for object categorization is object shape (e.g., Biederman and Ju 1988). Usually we are able to visually recognize objects although we see them from different points of view, in different sizes (due to changes in distance), and in different positions in the environment. Even young children recognize objects so immediately and effortlessly that it seems to be a rather ordinary and simple task. However, changes in the spatial relation between the observer and the object lead to immense changes of the image that is projected onto the retina. Hence, to recognize objects regardless of orientation, size, and position is not a trivial problem, and no computational system proposed so far can successfully recognize objects over a wide range of categories and contexts. The question about how we recognize objects despite spatial transformations is usually referred to as the first basic problem of object recognition. Moreover, we are not only able to recognize identical objects, after spatial transformations, but we can also effortlessly categorize an unfamiliar object, for instance a dog, a bird, or a butterfly, despite, sometimes, large shape variations within basic categories (e.g., Rosch et al. 1976). How do we generalize over different instances of an object class? This ability for class-level recognition or categorization is considered as the second basic problem of recognition. The main difficulty in classification arises from the variability in shape within natural classes of objects (e.g., Ullman 2007).

Objects can be recognized or categorized on different levels. For example, a specific object can be categorized as an *animal*, as a *dog*, as a *beagle*, or as my dog Snoopy. One of these levels has perceptual priority, and is called the basic level of categorization (Rosch et al. 1976; for a review see e.g., Murphy 2002). The basic level is usually also the entry level of categorization (Jolicoeur et al. 1984). Thus, we tend to recognize or name objects at the basic level, i.e., we see or name something as *dog*, *cat*, *car*, *table*, *chair*, etc. The level above the basic level is called superordinate level (e.g., vehicle, animal), while the level below the basic level is called subordinate level (limousine, van, hatchback or collie, dachshund, beagle, etc.).

The basic level is the most inclusive level at which members of this category have a high degree of visual similarity, and at which observers can still recognize an average shape, created from the shapes of several category members. (Rosch et al. 1976).[1] Therefore the basic level is the highest level of abstraction at which it is possible to form a mental image which is isomorphic to an average member of the class

---

[1] In addition, the basic level is the most inclusive level at which we tend to interact with objects in a similar way (Rosch et al. 1976), indicating the importance of knowledge about motor interactions for categorization (see Helbig et al. 2006).

and, thus, the most abstract level at which it is possible to have a relatively concrete image of a category (Rosch et al. 1976, Exp. 3 and 4; see also Diamond and Carey 1986). The pictorial nature of category representations up to the basic level has been confirmed in further experiments. In a signal detection experiment, subjects were better in detecting a masked object when the basic level name of the object was given prior to each trial. Superordinate level names did not aid in detection, which suggests that superordinate level categories are not represented in a pictorial code. Moreover, a priming experiment demonstrated that the basic level is the most abstract level at which preceding exposure of the category name affected same responses under physical identity instructions (Rosch et al. 1976, Exp. 5 and 6). These findings suggest that the basic level is the highest level at which category representations are image-based or pictorial.

Evidence for image-based or pictorial representations has been found also in the object recognition literature. The majority of findings indicate that recognition performance depends systematically on the amount of transformation (rotation, size-scaling, and shift in position) to align input and memory representations (for review see Graf 2006; for details see Sect. 3). This dependency suggests that representations are in a similar format as the visual input, whereas abstract representations should – by definition – be independent of image transformations.

A number of different accounts of recognition and categorization have been proposed, differing in the abstractness of the postulated representations (for reviews see Edelman 1997, 1999; Graf 2006; Murphy 2002; Palmeri and Gauthier 2004; Ullman 1996). Several models rely on relatively abstract representations. Models from the categorization literature are usually based on abstract features or properties (e.g., Nosofsky 1986; Cohen and Nosofsky 2000; Maddox and Ashby 1996; Markman 2001), but abstract features seem to be limited in their capacity to describe complex shapes. Structural description models from the recognition literature involve a decomposition into elementary parts and categorical spatial relations between these parts (like above, below, side-of), and thus are based on abstract propositional representations (e.g., Biederman 1987; Hummel and Biederman 1992; Hummel and Stankiewicz 1998). However, models that rely on abstract representations are difficult to reconcile with strong evidence for a systematic dependency on image transformations, like rotations and size-scalings (for review see, e.g., Graf 2006).

Several different image-based approaches have been proposed, which are better suited to account for the dependency on transformations (for reviews see Jolicoeur and Humphrey 1998; Tarr 2003). Early alignment models relied on transformational compensation processes, like mental rotation, in order to align stimulus representations and memory representations (e.g., Jolicoeur 1985, 1990a; Ullman 1989, 1996). As evidence has accumulated against mental rotations in object recognition (Jolicoeur et al. 1998; Willems and Wagemans 2001; Farah and Hammond 1988; Gauthier et al. 2002; for a review see Graf 2006), later image based approaches avoided the notion of transformation processes (e.g., Edelman 1997, 1998; Edelman and Intrator 2000, 2001; Perrett et al. 1998; Riesenhuber and Poggio 1999; Ullman 2007).

Moreover, hybrid models have been proposed in an attempt to combine structural and image-based approaches (Foster and Gilson 2002; Hayward 2003). Some hybrid

models have been derived from structural description models (Hummel and Stankiewicz 1998; Thoma et al. 2004), while others are extensions of image-based models (Edelman and Intrator 2000, 2001). The notion of structured representations is implicit also in the structural alignment approach brought forward in the literatures on similarity, analogy, and categorization (e.g., Medin et al. 1993; Gentner and Markman 1994, 1995; Goldstone and Medin 1994; Goldstone 1994a, 1994b, 1996; Markman 2001). This approach combines structured representations with the notion of alignment, the latter being used also in image-based approaches of recognition (Ullman 1989, 1996; Lowe 1985, 1987). More recently, a transformational framework of recognition involving alignment has been proposed, now relying not on mental rotations, but on an alignment based on coordinate transformations (Graf 2006; Graf et al. 2005; Salinas and Sejnowski 2001; Salinas and Abbott 2001; for further transformational approaches see Hahn et al. 2003; Leech et al. 2009).

Although the terms *recognition* and *categorization* are often used synonymously, they are investigated in two separate research communities. Traditionally the term *recognition* is more associated with perception and high-level vision, while *categorization* is more associated with cognition (e.g., Palmeri and Gauthier 2004). Members of both communities attend to different conferences, with relatively little overlap (see Farah 2000, p. 252). Surprisingly, relatively little research in the field of categorization is related to object shape (for reviews see Murphy 2002; Ashby and Maddox 2005). Consequently, relatively few attempts were made in order to come to an integrative approach of recognition and categorization (for exceptions see Edelman 1998, 1999; Nosofsky 1986; for an integrative review see Palmeri and Gauthier 2004). The aims of this article are related to this shortcoming. First, I will propose that Felix Klein's hierarchy of transformation groups can be regarded as a framework to conceptualize object shape and shape variability within categories up to the basic level. Second, given the commonalities between object recognition and categorization, I will lay out the foundations for an integrative transformational framework of recognition and categorization.

## 2   Form and Space

A prevalent idea in present cognitive neuroscience is that shape information and spatial information are processed in different visual streams, and therefore more or less dissociated. While shape processing for object recognition is postulated to occur exclusively in the ventral stream, the dorsal visual stream is involved in spatial tasks (Ungerleider and Mishkin 1982; Ungerleider and Haxby 1994), or perception for action (Milner and Goodale 1995; Goodale and Milner 2004). However, from a logical or geometrical point of view, shape and space cannot be strictly separated, but are closely related. As Stephen Kosslyn (1994, p. 277) argued, a shape is equivalent to a pattern formed by placing points (or pixels) at specific locations in space; a close look at any television screen is sufficient to convince anyone of this observation. Thus, shape is nothing more than a set of locations occupied by an object (Farah 2000, p. 71). Moreover, observers need to recognize shapes independent of the spatial

relation between observer and object, and thus compensate or account for spatial transformations in object recognition. Given this tight connection between form and space, it seems reasonable to investigate whether a geometrical (spatial) theory of shape is feasible. There is a growing body of evidence suggesting that areas in the parietal cortex – that is, areas usually associated with spatial or visuomotor processing – are involved in the recognition of disoriented objects (Eacott and Gaffan 1991; Faillenot et al. 1997, 1999; Kosslyn et al. 1994; Sugio et al. 1999; Vuilleumier et al. 2002; Warrington and Taylor 1973, 1978). A recent experiment using transcranial magnetic stimulation confirmed that the parietal cortex is involved in object recognition (Harris et al. 2009). Interestingly, the categorization of distorted dot pattern prototypes (in which different dot patterns were created by shifting the dots in space) involves not only typical shape-related areas like lateral occipital cortex, but also parietal areas (Seger et al. 2000; Vogels et al. 2002).

In accordance with the close connection between shape and space, I will argue that shape and shape variability can be conceptualized in terms of geometrical transformations. The organization of these different types of transformations can be described by *Felix Klein's (*1872*/1893) Erlanger Programm*, in which Klein proposed a nested hierarchy of geometrical transformation groups to provide an integrative framework for different geometries. This hierarchy of transformation groups ranges from simple transformations like rotations, translations (shifts in position), reflections, and dilations (size-scalings) – which make up the so-called Euclidean similarity group – to higher (and more embracing) transformation groups, namely affine, projective and topological transformations. I will explain and illustrate these transformations below, but before I will provide a brief historical survey to shed some light on the importance of the *Erlanger Programm* for geometry.

In the nineteenth century, geometry was in danger of falling apart into several separate areas, because different non-Euclidean geometries have been developed by mathematicians like Gauß, Lobatschewsky, and Bolayi, and later elaborated by Riemann. In 1872, Felix Klein was appointed as an ordinary professor of mathematics in Erlangen. In his inaugural address he proposed that different geometries can be integrated into one general framework – a project which was later called the *Erlanger Programm*. Klein argued that geometrical properties and objects are not absolute, but are relative to transformation groups.[2] For instance, a circle and an ellipse are different objects in Euclidean geometry, but in projective geometry all conic sections are equivalent. Regarding the projective group, a circle can be easily transformed into an ellipse or any other conic section.

Based on the idea that a geometry is defined relative to a transformation group, it was possible to integrate the different geometries by postulating a nested hierarchy of transformation groups. Euclidean geometry, projective geometry and space-curving geometries simply refer to different geometrical transformation groups in Klein's

---

[2] In mathematics, a group is a set that has rules for combining any pair of elements in the set, and that obeys four properties: closure, associativity, existence of an identity element and an inverse element. The mathematical concept of group has been used in cognitive psychology (e.g., Bedford 2001; Chen 2005; Dodwell 1983; Leyton 1992; Palmer 1983, 1989; Shepard 1994).

hierarchy. Note that these transformation groups are also important for the under-
standing of object shape.

The first important group in the hierarchy of transformations here is the so-
called *Euclidean similarity group*, which is made up of rotations, translations
(shifts in position), reflections, and size-scalings. The Euclidean similarity group
can be regarded as the basis of Euclidean geometry (see Ihmig 1997).[3] Mainly these
transformations need to be compensated after changes in the spatial relation
between observer and object.

The next higher transformation group is the group of *affine* transformations,
which includes transformations like linear stretchings or compressions in one
dimension, and also linear shear transformations, which change the angle of the
coordinate system. In short, affine transformations are linear transformations that
conserve parallelism, i.e., in which parallel lines remain parallel. A simple affine
stretching transformation occurs when TV programs in the usual 4:3 format are
viewed on the new 16:9 TV sets. Affine transformations are nicely illustrated by
Albrecht Dürer (1528)/(1996), who was probably the first who systematically
investigated the influence of geometrical transformations on object shape, focusing
on human bodies and faces (see Fig. 1). As every higher transformation group
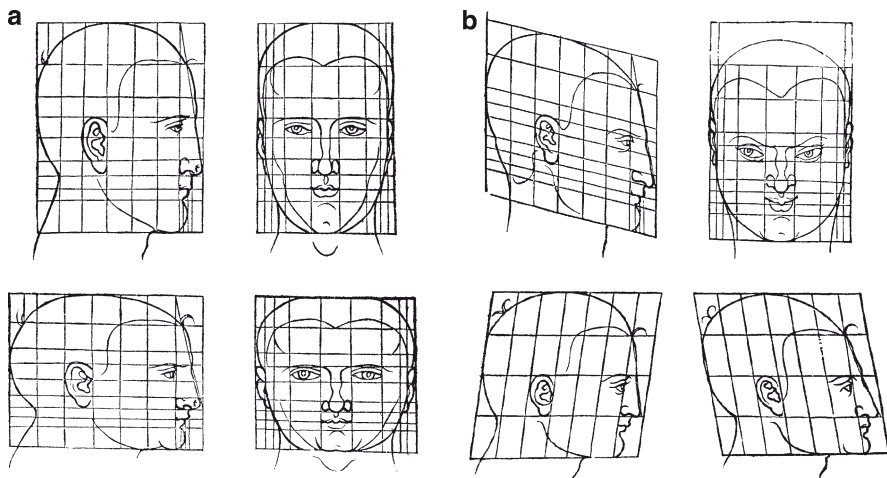


**Fig. 1** As demonstrated by Albrecht Dürer (1528/1996), affine transformations (i.e., linear trans-
formations that conserve parallelism) provide a way to account for some of the shape differences
between different heads. (**a**) *On the left side* affine compression and stretching transformations are
depicted. (**b**) *On the right* affine shear transformations are shown, which can include a transforma-
tion of the angle of the coordinate system (while parallels still remain parallel). *Note*: Drawings
by Albrecht Dürer 1528, State Library Bamberg, Germany, signature L.art.f.8a. Copyright by
State Library Bamberg, Germany. Adapted with permission

---

[3] The Euclidean similarity group can be further subdivided (e.g., Bedford 2001), but this is not
important for present purposes. A more detailed description of the hierarchy of transformation
groups can be found in Michaels and Carello (1981, p. 30–37), Palmer (1983), or Cutting (1986).

includes the lower group (nested hierarchy), the Euclidean similarity group is included in the group of affine transformations.

The next group in the hierarchy is the group of *projective* transformations. Projective transformations are linear transformations which do not necessarily conserve the parallelism of lines (and therefore violate Euclid's parallelity axiom). Central perspective in Renaissance paintings is constructed on the basis of projective geometry, as parallel lines intersect (in the vanishing point). Moreover, projective transformations allow describing yet further systematic changes of object shape beyond affine transformations (see Fig. 2).

Finally, the highest (or most basic) group in the hierarchy of point transformations is the group of *topological* transformations.[4] Topological transformations allow for nonlinear transformations, so that straight lines can be transformed into curved lines. Topological transformations can be illustrated by deforming a rubber sheet without ripping it apart. For this reason, topological geometry is often called
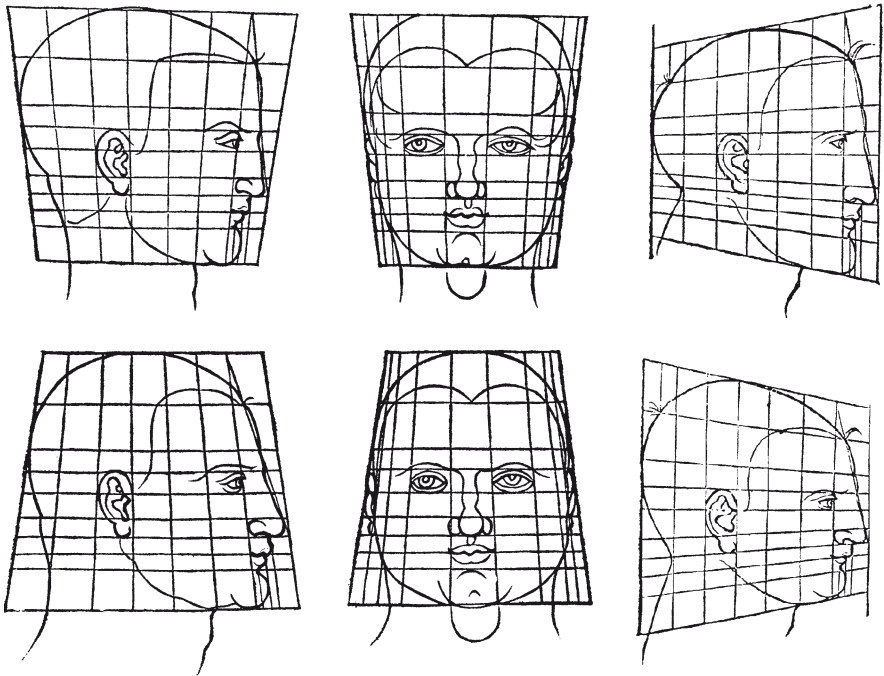


**Fig. 2** Projective transformations of shapes, which allow for linear transformations that violate parallelism, can account for a still larger range of shape variations between different heads. *Note*: Drawings by Albrecht Dürer 1528, State Library Bamberg, Germany, signature L.art.f.8a. Copyright by State Library Bamberg, Germany. Adapted with permission

---

[4] There are also transformations which go beyond point transformations (see Ihmig 1997). However, these do not seem to be of primary importance for an understanding of object shape.
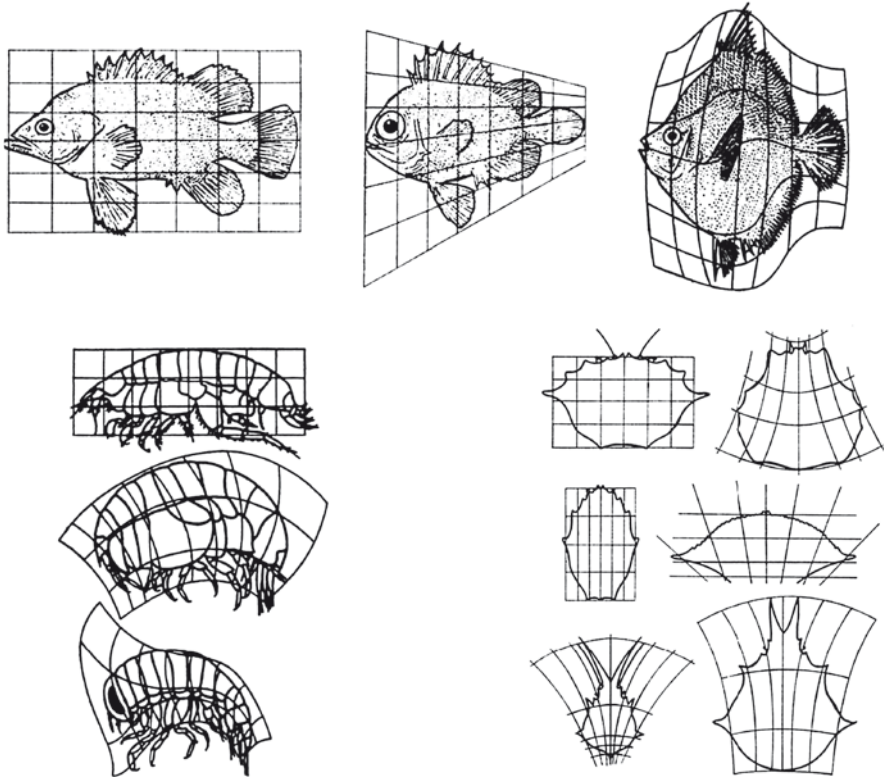
**Fig. 3** Topological transformations, which include also nonlinear (deforming) transformations, describe variations of the shapes of closely related animals, like different types of fish. *Note*: Drawings by Thompson 1917. Copyright by Cambridge University Press. Adapted with permission

rubber sheet geometry. As the hierarchy of transformation groups is a nested hierarchy, the group of topological transformations includes not only space-curving transformations, but all transformations that were described above. Topological geometry was first employed within a scientific theory in 1915 in Einstein's General theory of relativity. Only 2 years later, Thompson (1917)/(1942) used topological transformations to account for differences in the shapes of closely related animals (Fig. 3).

These different geometrical transformations provide a systematic way to describe the shape variability of biological objects, like facial profiles. A specific type of topological transformation is suited to characterize the remodeling of facial profiles by growth (Pittenger and Shaw 1975; Shaw and Pittenger 1977; see Fig. 4). This transformation has been dubbed "cardioidal" transformation, because it changes a circle into a heart shape. Cardioidal transformations can be described mathematically by using rather simple trigonometric functions. A matrix of head profiles was created by applying the cardioidal and an affine shear transformation
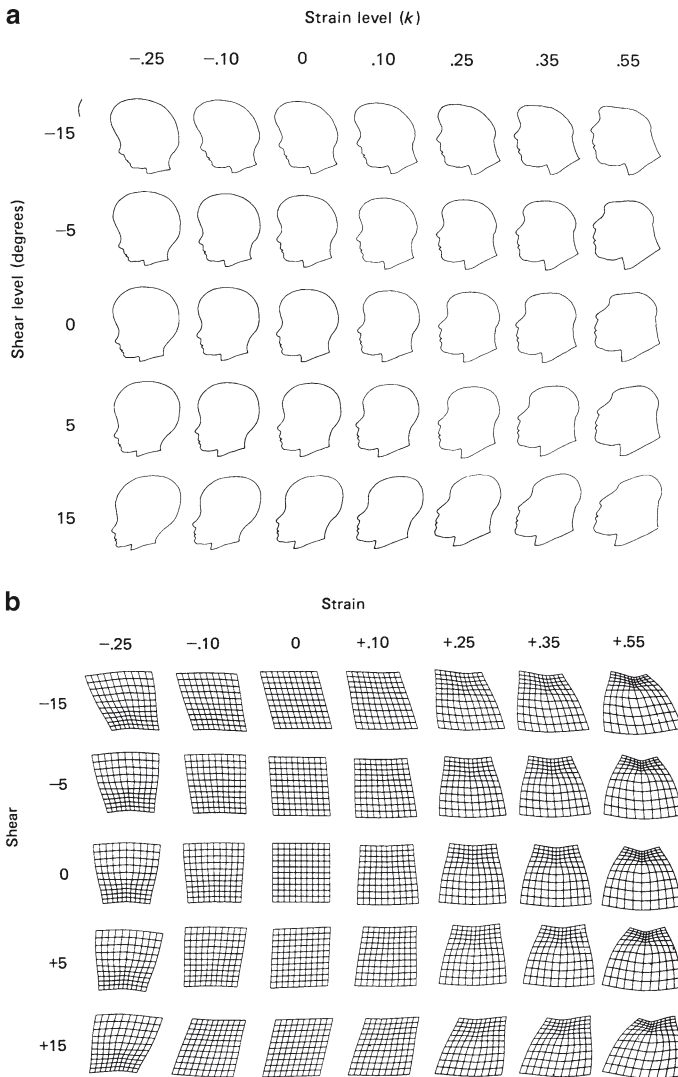
**Fig. 4** Affine shear and topological transformations provide a principled description of the shape space of the category *head*. (**a**) A set of facial profiles was created with topological transformations (*within rows*) and affine shear transformations (*within columns*), in order to investigate age perception (Shaw and Pittenger 1977). (**b**) The geometrical transformations can be conceptualized as transformations of the underlying coordinate system, i.e., as transformations of space. The standard grid is at shear = 0, strain = 0. Note that the deformations underlying the shape changes are rather simple. *Note*: Figures by Shaw and Pittenger 1977. Copyright by Robert Shaw. Adapted with permission

to the profile of a 10-year-old boy. The cardioidal (nonlinear) transformation allows deforming the head into the profile of a baby, into the profile of a grown-up man, or into some Neanderthal-man-like profile (see Fig. 4a, horizontal rows). Within

each column, the level of affine shear is modified. The underlying transformations that correspond to these shape changes are visualized by the deformation of the corresponding coordinate systems (Fig. 4b).[5] Thus, the shape changes are created by a deformation of the underlying space.

More recently, these ideas have been extended by proposing that the shape variability of members of a given basic-level category can be described by topological transformations (Graf 2002). The different facial profiles in Fig. 4a can be considered not only as phases in a growth process, but also as different members of the category *head*. By allowing for topological transformations, shape variability within categories can be accounted for – up to the basic level of categorization (Graf 2002). This proposal is consistent with the finding that the basic level is the highest level at which members of a category have similar shapes (Rosch et al. 1976), and it provides a systematic way to deal with these shape differences. This approach works well for biological objects, and also for many artifact categories (for constraints see Graf 2002). Topological transformations correspond to morphing in computer graphics. With morphing, the shape of one object is transformed into another, based on an alignment of corresponding points or parts. The use of topological (morphing) transformations has significant advantages: First, morphing offers the possibility to create highly realistic exemplars of familiar categories, moving beyond the artificial stimuli previously used in visual categorization tasks (for review see Ashby and Maddox 2005). Second, morphing is an image transformation which aligns corresponding features or parts. Thus, morphing is both image-based and structural. Third, the method allows to vary the shape of familiar objects in a parametric way, and thus permits a systematic investigation of shape processing.

The framework of the *Erlanger Programm* has been employed before in theories in perceptual and cognitive psychology. The *Erlanger Programm* and mathematical group theory proved to be useful to account for perceptual constancy (e.g., Wagemans et al. 1997; Cassirer 1944), perceptual organization (Chen 2001, 2005; Palmer 1983, 1989, 1999), the perception of motion and apparent motion (e.g., Chen 1985; Shepard 1994; Palmer 1983; Foster 1973, 1978), the perception of age (e.g., Shaw and Pittenger 1977), event perception (Warren and Shaw 1985), object identity decisions (Bedford 2001), and object categorization (Graf 2002; Shepard 1994). The *Erlanger Programm* may provide a basis for a broad framework of visual perception, which covers not only recognition and categorization, but also perceptual organization. Thus, the *Erlanger Programm* seems to be a promising theoretical framework, considering that an integral theory of perceptual organization and categorization is necessary (Schyns 1997).

In previous approaches the transformation groups of the Erlanger Programm have been typically used as a basis to define invariants, i.e., formless mathematical properties which remain unchanged despite spatial transformations (e.g., Gibson 1950; Todd et al. 1998; Van Gool et al. 1994; Wagemans et al. 1996; see already

---

[5]Nonlinear transformations seem to play a role also within the visual system. The retina is not flat but curved, and projections onto the retina are therefore distorted in a nonlinear way. Moreover, the projection from the retina to the primary visual cortex is highly nonlinear.

Cassirer 1944; Pitts and McCulloch 1947; for a review see Ullman 1996). For instance, the cross ratio is a frequently used invariant of the projective group (e.g., Cutting 1986; a description of the cross ratio can be found in Michaels and Carello 1981, pp. 35–36).[6] Invariant property approaches may be mathematically appealing but have at least two severe problems: The higher the relevant transformation group is in Klein's hierarchy, the more difficult it gets to find mathematical invariants (e.g., Palmer 1983). And, more important, the invariants that were postulated to underlie object constancy in the visual system could often not be empirically confirmed (e.g., Niall and Macnamara 1990; Niall 1992; but see Chen 2005). In the next section I will review evidence demonstrating that recognition and categorization performance is not invariant, but depends on the amount of geometrical transformation.

## 3 Recognition and Categorization Performance Depend on Spatial Transformations

As I argued in Sect. 2, Felix Klein's hierarchy of transformation groups offers a general way to conceptualize shape and shape variability. It provides an excellent framework for reviewing the existing studies both on shape recognition and categorization. Most studies in the object recognition literature are related to transformations of the Euclidean similarity group, especially to *rotations*, *dilations* (size-scalings) and *translations* (shifts in position). These transformations are the ones most relevant for recognizing a specific object, because the *same* object may be encountered in different orientations, positions, and sizes. Although we are able to recognize objects after spatial transformations, reaction times (RTs) typically increase with increasing transformational distance. This has been shown extensively for orientation, both in the picture plane (e.g., Jolicoeur 1985, 1988, 1990a; Lawson and Jolicoeur 1998, 1999) and in depth (e.g., Lawson and Humphreys 1998; Palmer et al. 1981; Srinivas 1993; Tarr et al. 1998; Lawson et al. 2000; for reviews see e.g., Graf 2006; Tarr 2003). There is also plenty of evidence that recognition performance depends on size (e.g., Bundesen and Larsen 1975; Bundesen et al. 1981; Cave and Kosslyn 1989; Jolicoeur 1987; Larsen and Bundesen 1978; Milliken and Jolicoeur 1992; for a review see Ashbridge and Perrett 1998). Moreover, an increasing number of studies show position dependency (Dill and Edelman 2001; Dill and Fahle 1998; Foster and Kahn 1985; Nazir and O'Regan 1990; Cave et al. 1994). Neurophysiological studies show a similar dependency on orientation, size and position (for a review see Graf 2006).

---

[6] Note that invariants need not necessarily be defined in relation to mathematical groups; invariants can be defined also regarding perspective transformations (e.g., Pizlo 1994), which do not fulfill the requirements of a mathematical group.

What about the higher transformation groups in Klein's hierarchy? Also for specific *affine transformations*, like stretching or compressing in one dimension, a monotonic relation between the extent of transformation and performance was found. The dependency on the amount of affine transformations has been demonstrated for simple shapes like ellipses (Dixon and Just 1978). Two ellipses were presented simultaneously, varying in shape by an affine stretching or compression. Subjects were instructed to judge whether the two ellipses were identical either regarding height or width (the relevant dimension was indicated before each trial). RTs deteriorated systematically with increasing affine stretching or compression of the ellipses – even though just the irrelevant dimension has been transformed. Dixon and Just argued that the stimuli were compared via a normalization process analogous to mental rotation and size scaling. Using more realistic stimuli, William Labov (1973) presented line drawings of cup-like objects that were created by changing the ratio of width to height (an affine stretching or compression). This manipulation changed the shape of the cup, and made it more mug-like, vase-like or bowl-like. The likelihood of assigning the objects into these categories varied with context. For instance, the likelihood of categorizing a cup-like object as a vase, for example, increased in the context "flower" (as compared to the "coffee" context), indicating that category boundaries are at least to some degree vague and context-dependent. More interesting here, Labov's findings also indicate that the likelihood of assigning an object to a category changes with the ratio of width to height – which may be regarded as first evidence that categorization is influenced by affine transformations. Further evidence that categorization performance depends on affine transformations can be found in a study by Cooper and Biederman (1993), although it was not designed to investigate this issue.

Up to today, relatively little research has been done regarding *projective transformations*, i.e., linear transformations which do not necessarily conserve the parallelism of lines. Evidence for a monotonic relation between the extent of projective transformation and task performance was found in an experiment which was designed to investigate whether the visual system distinguishes between different types of projective transformations (Wagemans et al. 1997; see also Niall 2000). Wagemans et al. presented three objects simultaneously on a computer screen, one on top as a reference stimulus and two below it. Subjects were instructed to determine which of the two patterns best matched the reference pattern. Recognition accuracy deteriorated with increasing amount of projective transformation. Thus, there is some provisional evidence that recognition performance depends also on the amount of projective transformations. A dependency on affine and projective transformations has also been demonstrated in neurophysiological experiments. The neural response of shape-tuned neurons in IT depends systematically on the amount of affine and projective transformations (Kayaert et al. 2005).

Furthermore, *topological (i.e., space-curving) transformations* play a role in visual perception. For instance, topological shape transformations were investigated in age perception (e.g., Pittenger and Shaw 1975; Shaw and Pittenger 1977; Pittenger et al. 1979; Mark and Todd 1985). Subjects had to age-rank facial profiles that were subjected to different amounts of affine shear or topological (cardioidal)

transformation (see Fig. 4). The results indicated that the age-rankings increased monotonically with increasing amount of topological (cardioidal) transformation. In addition, profiles transformed by the cardioidal transformation elicited more reliable rank-order judgments than those transformed by affine shear transformation. These experiments supplied evidence that age perception involves topological transformations. These transformations work not only for 2D shapes, but similar results have been found for 3D heads (e.g., Bruce et al. 1989).

Also object categorization is systematically related to the amount of topological transformation. A series of experiments with dot patterns, that were conducted to investigate the formation of abstractions or prototypes (e.g., Posner and Keele 1968, 1970; Posner et al. 1967; Homa et al. 1973), show a systematic dependency on nonlinear deforming transformations. In these experiments random dot patterns were defined as prototypes and distorted with statistical methods in order to produce different exemplars of the same category. The results indicate a monotonic relation between the extent of distortion and dependent variables like RT and error rate. These statistical distortions can be understood as topological transformations, if one assumes that the space between the dots is distorted: Imagine that the dots are glued onto a rubber sheet which can be stretched or compressed in a locally variable way. Thus, the dot pattern experiments fit nicely within an account suggesting that deforming transformations are involved in object categorization.

Further evidence for a monotonic relation between the RT and the amount of topological transformations can be derived from experiments by Edelman (1995) and Cutzu and Edelman (1996, 1998), using animal-like novel objects. A number of shapes were created by varying shape parameters that lead to nonrigid (morphing) transformations of the objects. The shape parameters were selected so that the animal-like objects in the distal shape space corresponded to a specific configuration in proximal shape (parameter) space, e.g., a cross, a square, a star or a triangle (see Fig. 5a). These configurations in proximal shape space could be recovered by multidimensional scaling (MDS) of subject data, using RT-data from a delayed matching to sample task, or similarity ratings (Fig. 5b).[7]

Additional evidence comes from studies on facial expression. Facial movements, such as smiles and frowns, can be described as topological transformations of the face. The latencies for the recognition of the emotional facial expressions increase with increasing topological transformation of facial expression prototypes (Young et al. 1997). A monotonic relation between the amount of topological transformation and RTs was also demonstrated in a task in which two faces were morphed, and the morphs had to be classified as either person A or B: The latencies for the classification of the faces increased with increasing distance from

---

[7] However, Edelman (1998) and Cutzu and Edelman (1996, 1998) do not interpret these results in terms of nonrigid transformations of pictorial representations, but regard the data as evidence for the existence of a low-dimensional monotonic psychological space, in which the similarity relations of the high-dimensional distal shape-space are represented. For a discussion of Edelman's account of recognition and categorization see Graf (2002).
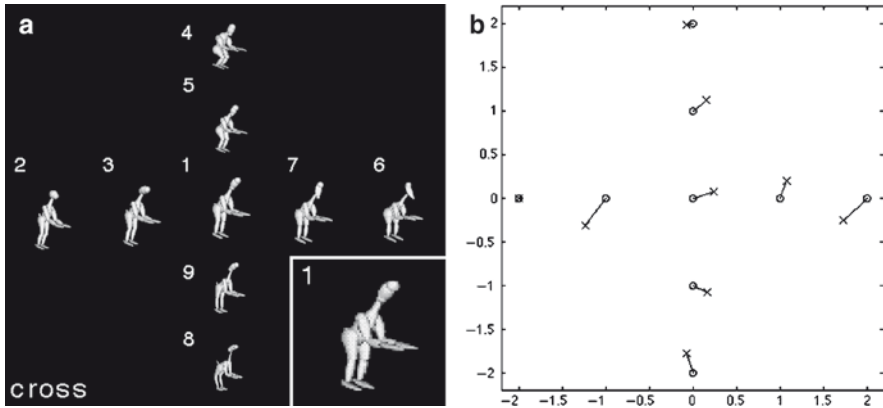
**Fig. 5** A systematic dependency between recognition performance and morph transformation has been demonstrated with animal-like novel objects (Cutzu and Edelman 1996). (**a**) Subjects were confronted with several classes of computer-rendered 3D animal-like shapes, arranged in a complex pattern (here: *cross*) in a common parameter space. (**b**) Response time and error rate data were combined into a measure of perceived pairwise shape similarities, and the object to object proximity matrix was submitted to nonmetric MDS. In the resulting solution, the relative geometrical arrangement of the points corresponding to the different objects reflected the complex low-dimensional structure in parameter space that defined the relationships between the stimulus classes. *Note*: Figures by Cutzu and Edelman 1996. Copyright by Cambridge University Press. Adapted with permission

the reference exemplar (Schweinberger et al. 1999, Exp. 1a). Moreover, the study by Schweinberger et al. indicated that the increase of identification latencies is not just due to an unspecific effect of the morphing procedure, like a loss of stimulus quality, because a morphing along the emotion dimension (happy vs. angry), which was irrelevant for this task, did not affect the latencies for face identification.

More recently, topological transformations in categorization have been studied in a more direct and systematic way, using various recognition and categorization tasks (Graf 2002; Graf et al. 2008; Graf and Bülthoff 2003). Graf and collaborators systematically varied the shapes of familiar biological and artifact categories by morphing between two exemplars from the same basic category (Fig. 6). Their results indicate that categorization performance depends on the magnitude of the topological transformation between two test stimuli (Graf 2002; Graf et al. 2008). That is, when two objects were presented sequentially, performance deteriorated systematically with increasing topological (morph) distance between the two category members. These results could not be reduced to alternative accounts, such as to affine transformations, or to changes in the configuration of parts (Graf 2002; Graf et al. 2008). First, the systematic dependency was also found for categories whose exemplars differed by only small affine changes, i.e., by only small changes of the aspect ratio of the objects. Second, the dependency appeared for categories whose members have very similar part configurations. Moreover,
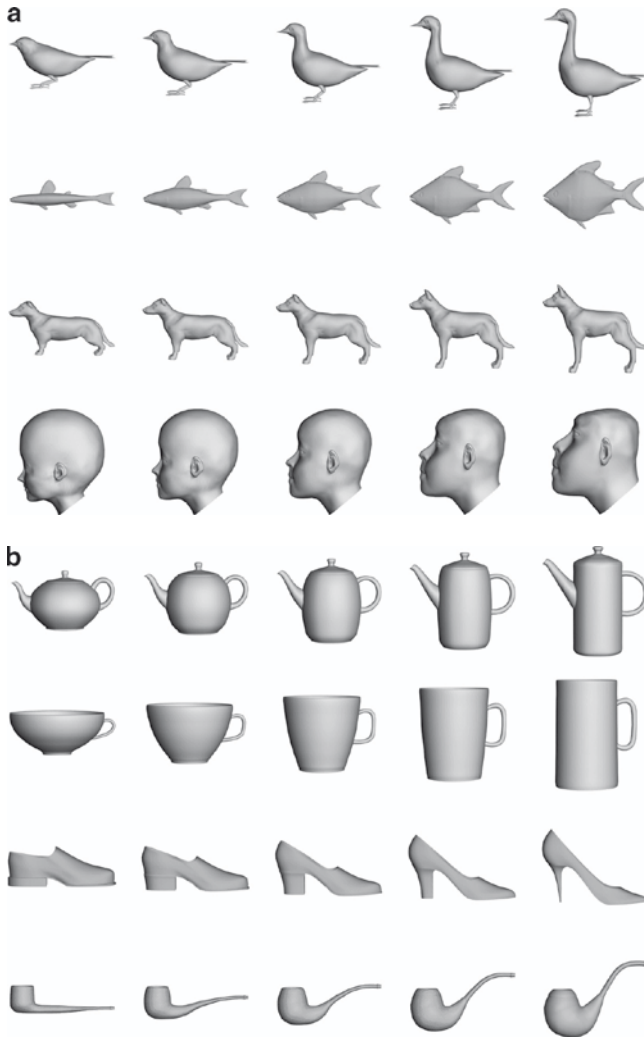
**Fig. 6** Shape variability within basic-level categories can be described well with nonlinear (topological) transformations. Intermediate category members are created by morphing between two exemplars from the same basic level category. Morphing describes within-category shape variability well for biological objects and for many (but not all) artifacts (see Graf 2002, for constraints)

transformation times in the categorization task were sequentially additive, suggesting that categorization relied on analog deforming transformations, i.e., on transformations passing through intermediate points on the transformational path (Graf 2002). The effect of topological distance holds for objects that were rotated in the picture-plane, and for objects that were shifted in position (Graf and

Bülthoff 2003). A systematic dependency on the degree of topological transformation has been demonstrated also in (nonspeeded) similarity- and typicality-rating tasks (Graf 2002; Graf et al. 2008; Hahn et al. 2009). Moreover, similarity judgments were biased by the direction of a morph animation which directly preceded the similarity decision (Hahn et al. 2009). Participants were shown short animations morphing one object into another from the same basic category. They were then asked to make directional similarity judgments – "how similar is object A to object B?" for two stationary images drawn from the morph continuum. Similarity ratings for identical comparisons were higher when reference object B had appeared before object A in the preceding morph sequence. Together, these results indicate that categorization performance depends systematically both on morph distance and on the direction of transformation. These findings are consistent with deformable template matching models of categorization (Basri et al. 1998; Belongie et al. 2002).

There is also neuropsychological evidence for deforming shape transformations in the visual system. A patient with intermetamorphosis reported that animals and objects she owned took the form of another animal or object. She also experienced changes in her husband's appearance, which could become exactly like that of a neighbor or could rapidly transform to look larger or smaller or younger (Courbon and Tusques 1932; see also Ellis and Young 1990).

To summarize, recognition and categorization performance depends on rotations, size-scalings, translations, affine transformations, projective transformations, and topological transformations. Thus, performance in recognition and categorization tasks depends systematically on the amount of geometrical transformation, for almost all transformation groups of the *Erlanger Programm*. The only exception are mirror reflections, for which the amount of transformation is not defined, and therefore a systematic dependency cannot be expected. Given these similar patterns of performance in recognition and categorization tasks it seems reasonable and parsimonious to assume that both basic problems of recognition – the problem of shape constancy and the problem of class level recognition – rely on similar processing principles, and just involve different transformation groups. The recognition of individual objects after spatial transformations is mostly related to Euclidean transformations, like rotations, size-scalings and translations. In some cases, like a projection of a shape onto a slanted surface, also affine and projective transformations may be necessary. For object categorization the higher (deforming) transformation groups, especially nonlinear transformations, seem most important. Shape variability within basic and subordinate level categories usually involves space-curving (topological) transformations. However, simpler transformations may describe some of the shape variability within object categories, like affine stretching transformations for cups and containers (see Labov 1973), or size-scaling transformations if some category members simply differ in size. Thus, recognition and categorization can be coarsely assigned to involve different transformation groups in Klein's hierarchy, although there is no clear-cut mapping.

# 4    Integrative Transformational Framework of Recognition and Categorization

The general dependency on the amount of transformation for almost all transformation groups is difficult to reconcile with the notion of invariance, or with invariant properties. In principle, invariant property approaches predict that recognition performance does not depend on the amount of transformation, as invariants are by definition unaffected by transformations (e.g., Van Gool et al. 1994; Chen 1982, 1985; Palmer 1983, 1989, 1999). Hence, a systematic relation between recognition or categorization performance and the amount of geometrical transformation would not be predicted.[8] Thus, there are reasons to doubt that invariant property models equally account for object recognition and categorization. This extends also to other models which rely on abstract representations and predict that recognition and categorization performance is basically independent of geometrical transformations, like structural description models from the recognition literature (e.g., Biederman 1987; Hummel and Biederman 1992; Hummel and Stankiewicz 1998).

Most of the existing image-based models account for the systematic dependency of recognition and categorization performance on the amount of spatial transformations – even though many models do not involve explicit transformation processes to compensate for image transformations (e.g., Edelman 1998, 1999; Perrett et al. 1998; Riesenhuber and Poggio 1999, 2002; Wallis and Bülthoff 1999). However, there are two further important classes of findings which need to be accounted for by any model of object recognition (Graf 2006). One class relates to the notion that the visual system carries out analog transformation processes in object recognition, i.e., continuous or incremental transformation processes. Evidence for analog transformations comes from a study showing that rotation times in a sequential picture–picture matching task are sequentially additive (Bundesen et al. 1981; for review see Graf 2006). In other words, the transformations in the visual system seem to traverse intermediate points on the transformational path. Sequential additivity was demonstrated not only for rotations, but also for morph transformations in a categorization task (Graf 2002). Several further studies suggest analog transformation processes (Kourtzi and Shiffrar 2001; Georgopoulos 2000; Georgopoulos et al. 1989; Lurito et al. 1991; Wang et al. 1998), but the evidence is not yet conclusive (for discussion see Graf 2006).

Another important class of findings is related to congruency effects in object recognition, suggesting that object recognition involves the adjustment of a perceptual frame of reference, or coordinate system. For instance, misoriented objects are recognized better when a different object has been presented immediately before in the

---

[8] In a variation of this approach, Shepard (1994) claimed that the linearity of transformation time (in mental rotation tasks and apparent motion tasks) is an invariant.

same orientation (Gauthier and Tarr 1997; Graf et al. 2005; Jolicoeur 1990b, 1992; Tarr and Gauthier 1998). This orientation congruency effect appeared when the two objects were similar or dissimilar, when they belonged to the same or to different superordinate categories, and when the objects had the same or a different (horizontal vs. vertical) main axis of elongation (Graf et al. 2005). Congruency effects have been found also for size (e.g., Larsen and Bundesen 1978; Cave and Kosslyn 1989). These congruency effects suggest that recognition involves the adjustment of a perceptual reference frame or coordinate system, because performance is improved when the coordinate system is adjusted to the right orientation or size.

   The existing image-based models cannot account for congruency effects in object recognition, because they are based on units that are simultaneously tuned to shape and orientation. Therefore, they do not predict a facilitation effect for the recognition of dissimilar shapes in the same orientation or size. Also current hybrid models which integrate image-based and structural representations (e.g., Edelman and Intrator 2000, 2001; Foster and Gilson 2002; Thoma et al. 2004) do not account for congruency effects. These models may account for congruency effects with similar parts, or similar structures, but not for dissimilar objects (see Graf 2006).

   In order to integrate this large body of findings, Graf (2002, 2006) proposed a transformational framework of recognition and categorization. According to this transformational framework, the transformation groups of the *Erlanger Programm* describe *time-consuming (and error-prone) transformation processes* in the visual system. During recognition, differences in the spatial relation between memory representation and stimulus representation are compensated by a transformation of a perceptual coordinate system which brings memory and stimulus representations into correspondence. When both are aligned, a matching can be performed in a simpler way. Note that these transformations are not transformations of mental images as in mental rotation, but coordinate transformations (transformations of a perceptual coordinate system), and thus similar to transformations involved in visuomotor control tasks (Graf 2006; Salinas and Sejnowski 2001; Salinas and Abbott 2001). Transformations of the Euclidean similarity group (especially rotations, size-scalings and shifts in position) are usually sufficient to compensate for spatial transformations in object recognition (due to changes in the spatial relation between observer and object).

   The transformational framework can be extended to account for categorization up to the basic level and compensate for shape differences – simply by allowing for nonlinear (deforming) transformations. Categorization is achieved by a deforming transformation which aligns memory and stimulus representations. For instance, Snoopy can be categorized as a dog by a topological (morphing) transformation which aligns the pictorial representation of the category dog and Snoopy's shape, until both can be matched. This approach provides an integrative framework of recognition and categorization up to the basic level, based on Klein's hierarchy of transformation groups. Recognition and categorization rely on similar processing principles, and differ mainly by involving different transformations (see Sect. 3). The transformational framework explains why recognition and categorization latencies depend in a systematic way on the amount of transformation which is

necessary for an alignment of memory representation and stimulus representation.[9] Moreover, the transformational framework is in accordance with findings showing that dynamic transformation processes are involved in categorization (Zaki and Homa 1999; see also Barsalou 1999), and with evidence for a transformational model of similarity (Hahn et al. 2003), including effects of the direction of a preceding morph transformation on subsequent similarity judgments (Hahn et al. 2009). Given the evidence that perceptual space is deformable and non-Euclidean (e.g., Hatfield 2003; Luneburg 1947; Suppes 1977; Watson 1978), the present framework suggests that physical space, perceptual space, and representational (categorical) space are endowed with a topological structure, leading to a unitary concept of space for these domains.

As morphing relies on an alignment of corresponding object parts or features, morphing can be regarded as an image-based structural alignment process. Thus, morphing may be an image-based instantiation of the structural alignment approach, which is prominent in the categorization literature (e.g., Gentner and Markman 1994, 1995; Goldstone and Medin 1994; Goldstone 1994a, 1994b, 1996; Markman 2001; Markman and Gentner 1993a, 1993b, 1997; Markman and Wisniewski 1997; Medin et al. 1993). The concept of image-based deforming transformations, or elastic matching, fits nicely with the idea of structured representations (Basri et al. 1998). Knowledge about the hierarchical organization of an object may also be important within an alignment approach, because this knowledge can guide the alignment process (Basri 1996), e.g., by facilitating the assignment of correspondent points or regions. This framework is compatible with evidence suggesting that object representations are structural or part-based (Biederman and Cooper 1991; Tversky and Hemenway 1984; Goldstone 1996; Newell et al. 2005; but see Cave and Kosslyn 1993; Murphy 1991), without having to assume abstract propositional representations.

According to the transformational framework, transformation processes are of primary importance – and not the search for invariant properties or features. This process-based view does not coincide with the invariants-based and static mathematical interpretation of the *Erlanger Programm* (as e.g., suggested by Niall 2000 or Cutting 1986, pp. 67–68). It seems more appropriate to assume a transformational framework of recognition and categorization, i.e., to focus on transformations and not on invariants. Nevertheless, invariants or features may be useful in a preselection process, or may play a role in solving the correspondence problem (e.g., Chen 2001; Carlsson 1999). Invariants may be involved in a fast feedforward sweep in visual processing which does not lead to a conscious percept, while conscious object perception seems to require recurrent processes (Lamme 2003; Lamme and Roelfsema 2000), potentially including transformational processes (Graf 2006).

---

[9] Not only the extent but also the type of topological distortion might be relevant (in analogy to the affine transformation group, cp. Wagemans et al. 1996). However, possible influences of the type of topological transformations are not in the focus of this work (for a discussion of the types of transformation see Bedford 2001).

To conclude, the transformational framework seems to be a highly parsimonious approach, because it is integrative in several different ways.

First, the transformational framework provides an integrative framework of recognition and categorization, based on Klein's hierarchy of transformation groups. Second, the transformational framework has the capacity to integrate image-based and structured (part-based) representations (see also Hahn et al. 2003), and can be regarded as an image-based extension of the structural alignment approach (e.g., Medin et al. 1993; Markman 2001). Third, according to the transformational framework, object recognition involves the adjustment of a perceptual coordinate system, i.e., involves coordinate transformations (Graf 2006). As coordinate transformations are fundamental also for visuomotor control, similar processing principles seem to be involved in object perception and perception for action (e.g., Graf 2006; Salinas and Abbott 2001; Salinas and Sejnowski 2001). This is compatible with the proposal that perception and action planning are coded in a common representational medium (e.g., Prinz 1990, 1997; Hommel et al. 2001). In accordance with this integrative approach, the recognition of manipulable objects can benefit from knowledge about typical motor interactions with the objects (Helbig et al. 2006). Finally, a framework of categorization based on pictorial representations and transformation processes is closely related to embodied approaches of cognition, like Larry Barsalou's (1999) framework of perceptual symbol systems, and embodied approaches to conceptual systems (Lakoff and Johnson 1999). The transformational framework does not invoke abstract propositional representations, but proposes that conceptual representations have a similar format as the perceptual input (see Graf 2002, 2006).

## 5   Open Questions and Outlook

The dependency of performance on the amount of transformation provides suggestive evidence that both object recognition and object categorization up to the basic level can be described by geometrical transformation processes. A process-based interpretation of the transformations of the *Erlanger Programm* offers the foundation for an integrative framework of object recognition and categorization. In any case, the *Erlanger Programm* provides a useful scheme to understand object shape, and to review the literature on object recognition and categorization.

Clearly, there are still open questions in the transformational framework. First, topological transformations are very powerful and can cross category boundaries. The question arises why, for instance, the dog template is aligned with a dog, but not with a cat, a cow, or a fish. Thus, constraints are necessary in order to avoid categorization errors. One important constraint is the transformational distance, which tends to be shorter within the same category than between categories. However, transformational distance alone might not be sufficient in all cases. Additional constraints may be provided by information about tolerable transformations (Bruce et al. 1991; see also Bruce 1994; Zaki and Homa 1999; see already

Murphy and Medin 1985; Landau 1994). The stored category exemplars may span some kind of space of tolerable topological transformations for each object category (Vernon 1952; Cootes et al. 1992; Baumberg and Hogg 1994; for a more detailed discussion see Graf 2002). In accordance, children at a certain age tend to overgeneralize and, for example, categorize many quadrupeds as dogs (e.g., Clark and Clark 1977; Waxman 1990).

A second problem is the so-called alignment paradox (e.g., Corballis 1988). It can be argued that an alignment through the shortest transformational path can only be achieved if the object is already identified or categorized. However, Ullman (1989, pp. 224–227) demonstrated that an alignment can be based on information which is available before identification or categorization (e.g., dominant orientation or anchor points), so that the paradox does not arise. This method can be used even for nonrigid transformations, if flexible objects are treated as locally rigid and planar. Further solutions to this correspondence problem have been proposed in the computer vision literature (e.g., Belongie et al. 2002; Carlsson 1999; Sclaroff 1997; Sclaroff and Liu 2001; Witkin et al. 1987; see also Ullman 1996). It should be noted that the problem is not fully solved yet. A potential resolution is that information necessary to perform an alignment is processed in a fast and unconscious feedforward sweep, while conscious recognition and categorization require recurrent processes (Lamme 2003; Lamme and Roelfsema 2000), like transformation processes.

Third, the neuronal implementation of a transformational framework of recognition and categorization is an open issue. It has been proposed that coordinate transformation processes in recognition and categorization are based on neuronal gain modulation (Graf 2006). Current approaches suggest that recognition and categorization are limited to the ventral visual stream (for reviews see e.g., Grill-Spector 2003; Grill-Spector and Sayres 2008; Malach et al. 2002). However, it seems possible that spatial transformation (and morphing) processes in recognition and categorization also involve the dorsal pathway, which is traditionally associated with spatial processing and coordinate transformations. There is suggestive evidence that the dorsal stream is involved in the recognition of objects that are rotated or size scaled (Eacott and Gaffan 1991; Faillenot et al. 1997, 1999; Gauthier et al. 2002; Harris et al. 2009; Kosslyn et al. 1994; Sugio et al. 1999; Vuilleumier et al. 2002; Warrington and Taylor 1973, 1978), and in the categorization of distorted dot patterns (Seger et al. 2000; Vogels et al. 2002).

Despite these open questions, the transformational framework seems promising due to its integrative potential. Moreover, if topological transformations are included into a framework of categorization, a number of further interesting issues can be tackled. First, deformations of objects due to nonrigid motion can be described by topological transformations. Many biological objects, including humans, deform when they move. In order to perceive an organism that is moving or adopting a new posture, the representation has to be updated, and deformations have to be compensated. Similarly, emotional facial expressions can be described by topological transformations (e.g., Knappmeyer et al. 2003). Second, the recognition of articulated objects can be covered by models that allow for deforming

transformations (e.g., Basri et al. 1998). Third, shape changes related to biological growth can be accounted for (Pittenger and Shaw 1975; Shaw and Pittenger 1977). Fourth, high-level adaptation phenomena, as reported in the face recognition literature (e.g., Leopold et al. 2001), can be described with deforming transformations. These high-level adaptation phenomena appear to involve deformations of the representational space. And fifth, the recognition of deformable objects, like for instance a rucksack, or cloths, etc., seems to require deformable transformations.

# References

Ashbridge E, Perrett DI (1998) Generalizing across object orientation and size. In: Walsh V, Kulikowski J (eds) Perceptual constancy. Why things look as they do. Cambridge University Press, Cambridge, pp 192–209

Ashby FG, Maddox WT (2005) Human category learning. Annu Rev Psychol 56:149–178

Barsalou LW (1999) Perceptual symbol systems. Behav Brain Sci 22:577–660

Basri R (1996) Recognition by prototypes. Int J Comput Vis 19:147–167

Basri R, Costa L, Geiger D, Jacobs D (1998) Determining the similarity of deformable shapes. Vis Res 38:2365–2385

Baumberg A, Hogg D (1994) Learning flexible models from image sequences. Proceedings of the Third European Conference on Computer Vision 1994. Springer, Berlin, pp 299–308

Bedford F (2001) Towards a general law of numerical/object identity. Curr Psychol Cogn 20:113–176

Belongie S, Malik J, Puzicha J (2002) Shape matching and object recognition using shape contexts. IEEE Trans Pattern Anal Mach Intell 24:509–522

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Biederman I, Cooper EE (1991) Priming contour-deleted images: evidence for intermediate representations in visual object recognition. Cogn Psychol 23:393–419

Biederman I, Ju G (1988) Surface vs. edge-based determinants of visual recognition. Cogn Psychol 20:38–64

Bruce V (1994) Stability from variation: the case of face recognition. The M.D. Vernon memorial lecture. Q J Exp Psychol 47A:5–28

Bruce V, Burton M, Doyle T, Dench N (1989) Further experiments on the perception of growth in three dimensions. Percept Psychophys 46:528–536

Bruce V, Doyle T, Dench N, Burton M (1991) Remembering facial configurations. Cognition 38:109–144

Bundesen C, Larsen A (1975) Visual transformation of size. J Exp Psychol Hum Percept Perform 1:214–220

Bundesen C, Larsen A, Farrell JE (1981) Mental transformations of size and orientation. In: Long J, Baddeley A (eds) Attention and Performance, IX. Erlbaum, Hillsdale, NJ, pp 279–294

Carlsson S (1999) Order structure, correspondence and shape based categories. In Forsyth DA, Mundy JL, di Gesú V, Cipolla R (eds) Shape, contour and grouping in computer vision. Lecture Notes in Computer Science 1681. Springer, Berlin, pp 58–71

Cassirer E (1944) The concept of group and the theory of perception. Philos Phenomenol Res 5:1–35

Cave KR, Kosslyn SM (1989) Varieties of size-specific visual selection. J Exp Psychol Gen 118:148–164

Cave CB, Kosslyn SM (1993) The role of parts and spatial relations in object identification. Perception 22:229–248

Cave KR, Pinker S, Giorgi L, Thomas CE, Heller LM, Wolfe JM, Lin H (1994) The representation of location in visual images. Cogn Psychol 26:1–32

Chen L (1982) Topological structure in visual perception. Science 218:699–700

Chen L (1985) Topological structure in the perception of apparent motion. Perception 14:197–208

Chen L (2001) Perceptual organization: to reverse back the inverted (upside-down) question of feature binding. Vis Cogn 8:287–303

Chen L (2005) The topological approach to perceptual organization. Vis Cogn 12:553–637

Clark HH, Clark EV (1977) Psychology and language. Hartcourt Brace Jovanovich, New York

Cohen AL, Nosofsky RM (2000) An exemplar-retrieval model of speeded same-different judgments. J Exp Psychol Hum Percept Perform 26:1549–1569

Cooper EE, Biederman I (1993) Geon differences during object recognition are more salient than metric differences. Poster presented at the annual meeting of the Psychonomic Society. Washington DC

Cootes TF, Taylor CJ, Cooper DH, Graham J (1992) Training models of shape from sets of examples. Proceedings of the British Machine Vision Conference 1992. Springer, Berlin, pp 9–18

Corballis MC (1988) Recognition of disoriented shapes. Psychol Rev 95:115–123

Courbon P, Tusques J (1932) Illusions d'intermetamorphose et de charme. Annales Medico-Psychologiques 14:401–406

Cutting JE (1986) Perception with an eye for motion. MIT Univ Press, Cambridge, MA

Cutzu F, Edelman S (1996) Faithful representation of similarities among three-dimensional shapes in human vision. Proc Natl Acad Sci 93:12046–12050

Cutzu F, Edelman S (1998) Representation of object similarity in human vision: psychophysics and a computational model. Vis Res 38:2229–2257

Diamond R, Carey S (1986) Why faces are and are not special: an effect of expertise. J Exp Psychol Gen 115:107–117

Dill M, Edelman S (2001) Imperfect invariance to object translation in the discrimination of complex shapes. Perception 30:707–724

Dill M, Fahle M (1998) Limited translation invariance of human visual pattern recognition. Percept Psychophys 60:65–81

Dixon P, Just MA (1978) Normalization of irrelevant dimensions in stimulus comparisons. J Exp Psychol Hum Percept Perform 4:36–46

Dodwell PC (1983) The Lie transformation group model of visual perception. Percept Psychophys 34:1–16

Dürer A (1528) Vier Bücher von menschlicher proportion. Reprint on the basis of the original edition, 3rd edn (1996). Verlag Dr. Alfons Uhl, Nördlingen, Germany

Eacott MJ, Gaffan D (1991) The role of monkey inferior parietal cortex in visual discrimination of identity and orientation of shapes. Behav Brain Res 46:95–98

Edelman S (1995) Representation of similarity in three-dimensional object discrimination. Neural Comput 7:408–423

Edelman S (1997) Computational theories of object recognition. Trends Cogn Sci 1:296–304

Edelman S (1998) Representation is representation of similarities. Behav Brain Sci 21:449–498

Edelman S (1999) Representation and recognition in vision. MIT Univ Press, Cambridge, MA

Edelman S, Intrator N (2000) (Coarse coding of shape fragments) + (retinotopy) ≈ representation of structure. Spatial Vis 13:255–264

Edelman S, Intrator N (2001) A productive, systematic framework for the representation of visual structure. In: Lean TK, Dietterich TG, Tresp V (eds) Advances in neural information processing systems 13. MIT Univ Press, Cambridge, MA, pp 10–16

Ellis HD, Young AW (1990) Accounting for delusional misidentifications. Br J Psychiatry 157:239–248

Faillenot I, Toni I, Decety J, Grégoire M-C, Jeannerod M (1997) Visual pathways for object-oriented action and object recognition. Functional anatomy with PET. Cereb Cortex 7:77–85

Faillenot I, Decety J, Jeannerod M (1999) Human brain activity related to the perception of spatial features of objects. NeuroImage 10:114–124

Farah MJ (2000) The cognitive neuroscience of vision. Blackwell, Oxford, UK

Farah MJ, Hammond KM (1988) Mental rotation and orientation-invariant object recognition: dissociable processes. Cognition 29:29–46

Foster DH (1973) A hypothesis connecting visual pattern recognition and apparent motion. Kybernetik 13:151–154

Foster DH (1978) Visual apparent motion and the calculus of variations. In: Leeuwenberg ELJ, Buffart HFJM (eds) Formal theories of visual perception. Wiley, New York, pp 67–82

Foster DH, Gilson SJ (2002) Recognizing novel three-dimensional objects by summing signals from parts and views. Proc R Soc Lond B 269:1939–1947

Foster DH, Kahn JI (1985) Internal representations and operations in visual comparison of transformed patterns: effects of pattern point-inversion, positional symmetry, and separation. Biol Cybern 51:305–312

Gauthier I, Tarr MJ (1997) Orientation priming of novel shapes in the context of viewpoint-dependent recognition. Perception 26:51–73

Gauthier I, Hayward WG, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (2002) BOLD activity during mental rotation and viewpoint-dependent object recognition. Neuron 34:161–171

Gentner D, Markman AB (1994) Structural alignment in comparison: no difference without similarity. Psychol Sci 5:152–158

Gentner D, Markman AB (1995) Similarity is like analogy. In: Cacciari C (ed) Similarity in language, thought, and perception. Brepols, Brussels, Belgium, pp 111–148

Georgopoulos AP (2000) Neural mechanisms of motor cognitive processes: functional MIT Univ Press and neurophysiological studies. In: Gazzaniga MS (ed) The new cognitive neurosciences. MIT Univ Press, Cambridge, MA, pp 525–538

Georgopoulos AP, Lurito JT, Petrides M, Schwartz AB, Massey JT (1989) Mental rotation of the neuronal population vector. Science 243:234–236

Gibson JJ (1950) The perception of the visual world. Houghton Mifflin, Boston

Goldstone RL (1994a) Similarity, interactive activation, and mapping. J Exp Psychol Learn Memory Cogn 20:3–28

Goldstone RL (1994b) The role of similarity in categorization: providing a groundwork. Cognition 52:125–157

Goldstone RL (1996) Alignment-based nonmonotonicities in similarity. J Exp Psychol Learn Memory Cogn 22:988–1001

Goldstone RL, Medin DL (1994) Time course of comparison. J Exp Psychol Learn Memory Cogn 20:29–50

Goodale MA, Milner AD (2004) Sight unseen. Oxford University Press, Oxford

Graf M (2002) Form, space and object. Geometrical transformations in object recognition and categorization. Wissenschaftlicher Verlag, Berlin

Graf M (2006) Coordinate transformations in object recognition. Psychol Bull 132:920–945

Graf M, Bülthoff HH (2003) Object shape in basic level categorisation. In Schmalhofer F, Young RM, Katz G (eds) Proceedings of the European Cognitive Science Conference. Lawrence Erlbaum, Mahwah, NJ, pp 390

Graf M, Kaping D, Bülthoff HH (2005) Orientation congruency effects for familiar objects: coordinate transformations in object recognition. Psychol Sci 16:214–221

Graf M, Bundesen C, Schneider WX (2009) Topological transformations in basic level object categorization. Manuscript submitted for publication

Grill-Spector K (2003) The neural basis of object perception. Curr Opin Neurobiol 13:1–8

Grill-Spector K, Sayres R (2008) Object recognition: insights from advances in fMRI methods. Curr Dir Psychol Sci 17:73–79

Hahn U, Chater N, Richardson LB (2003) Similarity as transformation. Cognition 87:1–32

Hahn U, Close J, Graf M (2009) Transformation direction influences shape-similarity judgments. Psychological Science, 20:447–454. DOI: 10.1111/j.1467-9280.2009.02310.x

Harris IM, Benito CT, Ruzzoli M, Miniussi C (2008) Effects of right parietal transcranial magnetic stimulation on object identification and orientation judgments. Journal of Cognitive Neuroscience, 20:916–926

Hatfield G (2003) Representation and constraints: the inverse problem and the structure of visual space. Acta Psychol 114:355–378

Hayward WG (2003) After the viewpoint debate: where next in object recognition? Trends Cogn Sci 7:425–427

Helbig HB, Graf M, Kiefer M (2006) The role of action representations in visual object recognition. Exp Brain Res 174:221–228

Homa D, Cross J, Cornell D, Goldman D, Shwartz S (1973) Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. J Exp Psychol 101:116–122

Hommel B, Müsseler J, Aschersleben G, Prinz W (2001) The theory of event coding (TEC): a framework for perception and action planning. Behav Brain Sci 24:849–937

Hummel JE, Biederman I (1992) Dynamic binding in a neural network for shape recognition. Psychol Rev 99:480–517

Hummel JE, Stankiewicz BJ (1998) Two roles for attention in shape perception: a structural description model of visual scrutiny. Vis Cogn 5:49–79

Ihmig K-N (1997) Cassirers Invariantentheorie der Erfahrung und seine Rezeption des "Erlanger Programms". Cassirer Forschungen, Band 2. Felix Meiner Verlag, Hamburg

Jolicoeur P (1985) The time to name disoriented natural objects. Memory Cogn 13:289–303

Jolicoeur P (1987) A size-congruency effect in memory for visual shape. Memory Cogn 15:531–543

Jolicoeur P (1988) Mental rotation and the identification of disoriented objects. Can J Psychol 42:461–478

Jolicoeur P (1990a) Identification of disoriented objects: a dual-systems theory. Mind Lang 5:387–410

Jolicoeur P (1990b) Orientation congruency effects on the identification of disoriented shapes. J Exp Psychol Hum Percept Perform 16:351–364

Jolicoeur P (1992) Orientation congruency effects in visual search. Can J Psychol 46:280–305

Jolicoeur P, Humphrey GK (1998) Perception of rotated two-dimensional and three-dimensional objects and visual shapes. In Walsh V, Kulikowski J (eds) Perceptual constancy. Why things look as they do. Cambridge University Press, Cambridge, pp 69–123

Jolicoeur P, Gluck MA, Kosslyn SM (1984) Pictures and names: making the connection. Cogn Psychol 16:243–275

Jolicoeur P, Corballis MC, Lawson R (1998) The influence of perceived rotary motion on the recognition of rotated objects. Psychon Bull Rev 5:140–146

Kayaert G, Biederman I, Op de Beeck H, Vogels R (2005) Tuning for shape dimensions in macaque inferior temporal cortex. Eur J Neurosci 22:212–224

Klein F (1872/1893) Vergleichende Betrachtungen über neuere geometrische Forschungen. Mathematische Annalen 43:63–100

Knappmeyer B, Thornton IM, Bülthoff HH (2003) The use of facial motion and facial form during the processing of identity. Vis Res 43:1921–1936

Kosslyn SM (1994) Image and brain. MIT Univ Press, Cambridge, MA

Kosslyn SM, Alpert NM, Thompson WL, Chabris CF, Rauch SL, Anderson AK (1994) Identifying objects seen from different viewpoints. A PET investigation. Brain 117:1055–1071

Kourtzi Z, Shiffrar M (2001) Visual representation of malleable and rigid objects that deform as they rotate. J Exp Psychol Hum Percept Perform 27:335–355

Labov W (1973) The boundaries of words and their meanings. In: Bailey C-JN, Shuy RW (eds) New ways of analyzing variations in English. Georgetown University Press, Washington, D.C

Lakoff G (1987) Women, fire, and dangerous things. What categories reveal about the mind. University of Chicago Press, Chicago

Lakoff G, Johnson M (1999) Philosophy in the flesh: the embodied mind and Its challenge to western thought. Basic Books, New York, NY

Lamme VAF (2003) Why visual attention and awareness are different. Trends Cogn Sci 7:12–18

Lamme VAF, Roelfsema PR (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23:571–579

Landau B (1994) Object shape, object name, and object kind: representation and development. In: Medin DL (ed) The psychology of learning and motivation 31. Academic Press, New York, pp 253–304

Larsen A, Bundesen C (1978) Size scaling in visual pattern recognition. J Exp Psychol Hum Percept Perform 4:1–20

Lawson R, Humphreys GW (1998) View-specific effects of depth rotation and foreshortening on the initial recognition and priming of familiar objects. Percept Psychophys 60:1052–1066

Lawson R, Jolicoeur P (1998) The effects of plane rotation on the recognition of brief masked pictures of familiar objects. Memory Cogn 26:791–803

Lawson R, Jolicoeur P (1999) The effect of prior experience on recognition thresholds for plane-disoriented pictures of familiar objects. Memory Cogn 27:751–758

Lawson R, Humphreys GW, Jolicoeur P (2000) The combined effects of plane disorientation and foreshortening on picture naming: one manipulation are two? J Exp Psychol Hum Percept Perform 26:568–581

Leech R, Mareschal D, Cooper RP (2008) Analogy as relational priming: A developmental and computational perspective on the origins of a complex cognitive skill. Behavioral and Brain Sciences 31:357–378

Leopold DA, O'Toole AJ, Vetter T, Blanz V (2001) Prototype-referenced shape encoding revealed by high-level aftereffects. Nat Neurosci 4:89–94

Leyton M (1992) Symmetry, causality, mind. MIT Univ Press, Cambridge, MA

Lowe DG (1985) Perceptual organization and visual recognition. Kluwer, Boston, MA

Lowe DG (1987) Three-dimensional object recognition from single two-dimensional images. Artif Intell 31:355–395

Luneburg RK (1947) Mathematical analysis of binocular vision. Princeton University Press, Princeton, NJ

Lurito T, Georgakopoulos T, Georgopoulos AP (1991) Cognitive spatial-motor processes: 7. The making of movements at an angle from a stimulus direction: studies of motor cortical activity at the single cell and population levels. Exp Brain Res 87:562–580

Maddox WT, Ashby FG (1996) Perceptual separability, and the identification-speeded classification relationship. J Exp Psychol Hum Percept Perform 22:795–817

Malach R, Levy I, Hasson U (2002) The topography of high-order human object areas. Trends Cogn Sci 6:176–184

Mark LS, Todd JT (1985) Describing perceptual information about human growth in terms of geometric invariants. Percept Psychophys 37:249–256

Markman AB (2001) Structural alignment, similarity, and the internal structure of category representations. In: Hahn U, Ramscar M (eds) Similarity and categorization. Oxford University Press, Oxford, pp 109–130

Markman AB, Gentner D (1993a) Splitting the differences: a structural alignment view of similarity. J Memory Lang 32:517–535

Markman AB, Gentner D (1993b) Structural alignment during similarity comparisons. Cogn Psychol 25:431–467

Markman AB, Gentner D (1997) The effects of alignability on memory. Psychol Sci 8:363–367

Markman AB, Wisniewski EJ (1997) Similar and different: the differentiation of basic-level categories. J Exp Psychol Learn Memory Cogn 23:54–70

Medin DL, Goldstone RL, Gentner D (1993) Respects for similarity. Psychol Rev 100:254–278

Michaels CF, Carello C (1981) Direct perception. Prenctice-Hall, Englewood Cliffs, NJ

Milliken B, Jolicoeur P (1992) Size effects in visual recognition memory are determined by perceived size. Memory Cogn 20:83–95

Milner AD, Goodale MA (1995) The visual brain in action. Oxford University Press, Oxford, England

Murphy GL (1991) Parts in object concepts: experiments with artificial categories. Memory Cogn 19:423–438

Murphy GL (2002) The big book of concepts. MIT Univ Press, Cambridge, MA

Murphy GL, Medin DL (1985) The role of theories in conceptual coherence. Psychol Rev 92:289–316

Nazir TA, O'Regan JK (1990) Some results on translation invariance in the human visual system. Spat Vis 5:81–100

Newell FN, Sheppard DM, Edelman S, Shapiro KL (2005) The interaction of shape- and location-based priming in object categorisation: evidence for a hybrid "what+where" representation stage. Vis Res 45:2065–2080

Niall KK (1992) Projective invariance and the kinetic depth effect. Acta Psychol 81:127–168

Niall KK (2000) Some plane truths about pictures: notes on Wagemans, Lamote, and van Gool (1997). Spat Vis 13:1–24

Niall KK, Macnamara J (1990) Projective invariance and picture perception. Perception 19:637–660

Nosofsky RM (1986) Attention, similarity, and the identification-categorization-relationship. J Exp Psychol Gen 115:39–57

Palmer SE (1983) The psychology of perceptual organization: a transformational approach. In: Beck J, Hope B, Rosenfeld A (eds) Human and machine vision. Academic Press, New York, pp 269–339

Palmer SE (1989) Reference frames in the perception of shape and orientation. In: Shepp BE, Ballesteros S (eds) Object perception: structure and process. Erlbaum, Hillsdale, NJ, pp 121–163

Palmer SE (1999) Vision science. Photons to phenomenology. MIT Univ Press, Cambridge, MA

Palmer SE, Rosch E, Chase P (1981) Canonical perspective and the perception of objects. In: Long J, Baddeley A (eds) Attention and performance IX. Erlbaum, Hillsdale, NJ, pp 135–151

Palmeri TJ, Gauthier I (2004) Visual object understanding. Nat Rev Neurosci 5:1–13

Perrett DI, Oram WM, Ashbridge E (1998) Evidence accumulation in cell populations responsive to faces: an account of generalization of recognition without mental transformations. Cognition 67:111–145

Pittenger JB, Shaw RE (1975) Aging faces as viscal-elastic events: implications for a theory of nonrigid shape perception. J Exp Psychol Hum Percept Perform 1:374–382

Pittenger JB, Shaw RE, Mark LS (1979) Perceptual information for the age level of faces as a higher order invariant of growth. J Exp Psychol Hum Percept Perform 5:478–493

Pitts W, McCulloch WS (1947) How we know universals: the perception of auditory and visual forms. Bull Math Biophys 9:127–147

Pizlo Z (1994) A theory of shape constancy based on perspective invariants. Vis Res 34:1637–1658

Posner MI, Keele S (1968) On the genesis of abstract ideas. J Exp Psychol 77:353–363

Posner MI, Keele S (1970) Retention of abstract ideas. J Exp Psychol 83:304–308

Posner MI, Goldsmith R, Welton KE (1967) Perceived distance and the classification of distorted patterns. J Exp Psychol 73:28–38

Prinz W (1990) A common coding approach to perception and action. In: Neumann O, Prinz W (eds) Relationships between perception and action: current approaches. Springer, Berlin, New York, pp 167–201

Prinz W (1997) Perception and action planning. Eur J Cogn Psychol 9:129–154

Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. Nat Neurosci 2:1019–1025

Riesenhuber M, Poggio T (2002) Neural mechanisms of object recognition. Curr Opin Neurobiol 12:162–168

Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P (1976) Basic objects in natural categories. Cogn Psychol 8:382–439

Salinas E, Abbott LF (2001) Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. In Nicolelis MAL (ed) Advances in neural population coding. Progress in brain research, vol 130. Elsevier, Amsterdam, pp 175–190

Salinas E, Sejnowski TJ (2001) Gain modulation in the central nervous system: where behavior, neurophysiology, and computation meet. Neuroscientist 7:430–440

Schweinberger SR, Burton AM, Kelly SW (1999) Asymmetric dependencies in perceiving identity and emotion: experiments with morphed faces. Percept Psychophys 61:1102–1115

Schyns PG (1997) Categories and concepts: a bi-directional framework for categorization. Trends Cogn Sci 1:183–189

Sclaroff S (1997) Deformable prototypes for encoding shape categories in image databases. Pattern Recogn 30:627–642

Sclaroff S, Liu L (2001) Deformable shape detection and description via model-based region grouping. IEEE Trans Pattern Anal Mach Intell 23:475–489

Seger CA, Poldrack RA, Prabhakaran V, Zhao M, Glover G, Gabrieli JDE (2000) Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. Neuropsychologia 38:1316–1324

Shaw R, Pittenger J (1977) Perceiving the face of change in changing faces: implications for a theory of object perception. In Shaw R, Bransford J (eds) Perceiving, acting, and knowing. Toward an ecological psychology. Erlbaum, Hillsdale, NJ pp. 103–132

Shepard RN (1994) Perceptual-cognitive universals as reflections of the world. Psychon Bull Rev 1:2–28

Srinivas K (1993) Perceptual specificity in nonverbal priming. J Exp Psychol Learn Memory Cogn 19:582–602

Sugio T, Inui T, Matsuo K, Matsuzawa M, Glover GH, Nakai T (1999) The role of the posterior parietal cortex in human object recognition: a functional magnetic resonance imaging study. Neurosci Lett 276:45–48

Suppes P (1977) Is visual space Euclidean? Synthese 35:397–421

Tarr MJ (2003) Visual object recognition: can a single mechanism suffice? In Peterson MA, Rhodes G (eds) Perception of faces, objects, and scenes. Analytic and holistic processes. Oxford University Press, Oxford, England, pp 177–211

Tarr MJ, Gauthier I (1998) Do viewpoint-dependent mechanisms generalize across members of a class? Cognition 67:71–109

Tarr MJ, Williams P, Hayward WG, Gauthier I (1998) Three-dimensional object recognition is viewpoint-dependent. Nat Neurosci 1:275–277

Thoma V, Hummel JE, Davidoff J (2004) Evidence for holistic representations of ignored images and analytic representations of attended images. J Exp Psychol Hum Percept Perform 30:257–267

Thompson D'AW (1917) On growth and form, 2nd edn. (1942). Cambridge University Press, Cambridge

Todd JT, Chen L, Norman JF (1998) On the relative salience of Euclidean, affine, and topological structure for 3-D form discrimination. Perception 27:273–282

Tversky B, Hemenway K (1984) Objects, parts, and categories. J Exp Psychol Gen 113:169–193

Ullman S (1989) Aligning pictorial descriptions: an approach to object recognition. Cognition 32:193–254

Ullman S (1996) High-level vision. Object recognition and visual cognition. MIT Univ Press, Cambridge, MA

Ullman S (2007) Object recognition and segmentation by a fragment-based hierarchy. Trends Cogn Sci 11:58–64

Ungerleider LG, Haxby JV (1994) 'What' and 'where' in the human brain. Curr Opin Neurobiol 4:157–165

Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW (eds) Analysis of visual behavior. MIT Univ Press, Cambridge, MA, pp 549–586

Van Gool LJ, Moons T, Pauwels E, Wagemans J (1994) Invariance from the Euclidean geometer's perspective. Perception 23:547–561

Vernon MD (1952) A further study of visual perception. Cambridge University Press, London

Vogels R, Sary G, Dupont P, Orban GA (2002) Human brain regions involved in visual categorization. NeuroImage 16:401–414

Vuilleumier P, Henson RN, Driver J, Dolan RJ (2002) Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat Neurosci 5:491–499

Wagemans J, Van Gool L, Lamote C (1996) The visual systems measurement of invariants need not itself be invariant. Psychol Sci 7:232–236

Wagemans J, Lamote C, Van Gool L (1997) Shape equivalence under perspective and projective transformations. Psychon Bull Rev 4:248–253

Wallis G, Bülthoff H (1999) Learning to recognize objects. Trends Cogn Sci 3:22–31

Wang G, Tanifuji M, Tanaka K (1998) Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. Neurosci Res 32:33–46

Warren WH, Shaw RE (1985) Events and encounters as units of analysis for ecological psychology. In Warren WH, Shaw RE (eds) Persistence and change. Proceedings of the first international conference on event perception. Erlbaum, Hillsdale, NJ, pp 1–27

Warrington EK, Taylor AM (1973) The contribution of the right parietal lobe to object recognition. Cortex 9:152–164

Warrington EK, Taylor AM (1978) Two categorical stages of object recognition. Perception 7:695–705

Watson AB (1978) A Riemann geometric explanation of the visual illusions and figural aftereffects. In: Leeuwenberg E, Buffart H (eds) Formal theories of visual perception. Wiley, New York, NY, pp 139–169

Waxman SR (1990) Linguistic biases and the establishment of conceptual hierarchies: evidence from preschool children. Cogn Dev 5:123–150

Willems B, Wagemans J (2001) Matching multicomponent objects from different viewpoints: mental rotation as normalization? J Exp Psychol Hum Percept Perform 27:1090–1115

Witkin A, Terzopoulos D, Kass M (1987) Signal matching through scale space. Int J Comput Vis 2:133–144

Young AW, Rowland D, Calder AJ, Etcoff NL, Seth A, Perrett DI (1997) Facial expression megamix: tests of dimensional and category accounts of emotion recognition. Cognition 63:271–313

Zaki SR, Homa D (1999) Concepts and transformational knowledge. Cogn Psychol 39:69–115