Object Recognition in Picasso’s Abstract Art

Aaron Kozbelt

The visual arts have for hundreds of years presented viewers with images that have been more or less representational of the external world. In fact, many of these images have been such convincing representations of the world that they can fool the viewer into thinking that the representation is in fact reality or a photograph, and not paint on a canvas. Examples of such trompe l’oeil (fool the eye) paintings are some of the works of Jan Vermeer, or Fra Andrea Pozzo’s painting on the ceiling of the Church of Sant’ Ignazio in Rome (Rock, 1984). There are certain limitations on being fooled, however. For example, spectators must view the painting from the same angle from which the artist viewed it. Characteristics of paintings that would destroy the illusion, such as brush strokes, must also be concealed (Rock, 1984).

It is not just in trompe l’oeil paintings that viewers experience a perception of what is represented in the picture. In viewing representational art, the spectator perceives both the scene depicted and the markings and brushstrokes that are characteristics of the surface of the painting. The philosopher Michael Polanyi refers to the latter as our “subsidiary awareness” and the former as our “focal awareness.” (Rock, 1984) Moreover, objects are recognizably specific objects in many styles of painting. With trompe l’oeil paintings, the same visual processes, including those dealing with object recognition, would seem to operate as with objects in the real world. What about paintings that are not strictly representational? Everyone knows that paintings that are not exceptionally well-painted and are not entirely convincing representations still convey a sense that certain objects are represented. What about other styles of paintings? If objects are recognizable in paintings that are not exact representations of the world, then objects should also be recognizable (to some degree) in art that is more abstract. How readily an object in a painting is recognized should depend on how closely it resembles the actual object. This pattern should break down at some point, when the object is abstracted beyond recognition. The great twentieth-century artist Pablo Picasso (1881–1973) is still thought of as notorious by many for his very free abstractions and the distortions of objects in his work. Despite this notoriety, most people have little difficulty in recognizing specific objects in many of his paintings, drawings and sculptures. The enormous stylistic variety of Picasso’s abstract work can provide an excellent basis for examining exactly how
objects are perceived and categorized in terms of various object recognition theories, and at what point and why the breakdown of object recognition occurs.

Some of the pictorial depth cues which function in realistic painting also operate in Picasso’s abstract art. Among these cues are interposition, linear perspective, size constancy, and aerial perspective. Interposition is a very important principle which in many works Picasso maintained despite distorting individual objects. Both size constancy, the knowledge that objects farther away are smaller, as well as interposition are evident in a 1968 drawing on the theme of *The Turkish Bath.* Other pictorial depth cues, including linear perspective, are not common in Picasso’s work. His invention of Cubism around 1909 effectively renounced the Western artistic tradition of creating depth by perspective. Cubism attempted to use not one, two, or three vanishing points, but rather an infinite number, which led to objects being viewed from multiple positions simultaneously. When Picasso does use some aspects of linear perspective, the effect does not necessarily create depth in the traditional sense, often due to other flattened elements in the work which function to create a more abstract pattern, as in *L’Aubade* of 1942 (§370) or *Child with Pigeons* of the following year (§374). Aerial perspective, in which more distant objects appear hazier and bluer, is used even less, as usually Picasso’s concern is in how objects are constructed and less in how subtle blending or color effects can convey a sense of distance. Although some pictorial cues are found in some works by Picasso, in other works he violates all of these principles, by flattening the picture space into abstract geometrical shapes, distorting the size and orientation of parts of objects, and representing objects from multiple points of view simultaneously. Therefore, these simple pictorial cues to depth cannot by any means explain all of the object recognition processes viewers experience when looking at Picasso’s work. Other factors must contribute to object recognition in both realistic and abstract art.

Marr and Nishihara (1978) present a view of object recognition that is based on several properties of the object in question: an axis of the object’s overall shape and decomposition of the object into parts. This model is object-centered and uses an internal three-dimensional object representation. This allows for greater economy in encoding information about objects, but is not directly applicable to two-dimensional paintings, since the objects in the paintings are generally viewed from one perspective only. Paintings can depict objects from any point of view, however, so any recognizable view of an object should activate its three-dimensional object representation. A special class of shapes called generalized cylinders (or generalized cones) give both the properties of axis and decomposition (Humphreys and Bruce, 1989). It is easy to determine the axis of the overall shape, and many complex objects are easily reducible to generalized cylinders. If this theory of object recognition is correct, then objects that are built up only of these generalized cylinders (or something close to them) should be recognizable in paintings in which the artist abstracts objects by reducing them to components corre-
sponding to these cylinders. Picasso presents many pieces where objects are represented by or are decomposable into these generalized cylinders. In works such as *Les Demoiselles d'Avignon* (1907, *99), *Centaures, femme, et oiseaux* (1946)*, or The Bathers* (1956, *427), where objects are represented as broad geometrical shapes resembling generalized cylinders, the abstracted figures are still recognizable as human (or birds or centaurs). The fact that the proportions and articulations of the generalized cylinders of the painted figures roughly correspond to those of real people, together with the presence of facial features in some of the figures, also contribute to their recognizability.

The proportions of the cylinders can be exaggerated to a high degree without losing the ability to recognize the object, particularly when other information is present. In *La Baignade* (1937)*, many of the figures' proportions (e.g., of the necks) appear to be quite exaggerated, yet viewers tend to have no problem identifying the subject matter of the painting. The presence of other factors, including facial features, some three-dimensional modeling, color, and the appearance of some type of social interaction also probably plays a role in this recognition. It is not difficult to imagine, however, removing these confounds and still being able to recognize the figures based solely on the generalized cylinders. Object recognition, then, is still operating under these conditions of geometrical abstraction.

When a figure is presented in such a way that the proportions or orientation of the geometrical patterns do not follow the standard grouping of generalized cylinders, but emphasize some other aspect of the object, the recognizability of the object can be drastically decreased. Many drawings and paintings from Picasso's Analytic Cubist period, for example a 1910 drawing of a *Female Nude* (*141), break down the form of the figures in ways that do not correspond well to generalized cylinders. In many works of this period, there also seems to be a severe limitation of Gestalt grouping processes, making it difficult to determine which features or locations belong together. This highly abstract drawing seems to be straining the limits of human recognition ability. The Marr-Nishihara model would explain this failure in terms of the viewer's inability to apply generalized cylinders to these cases. It is also possible that the viewer could not extract cylinders because all of the low-level Gestalt groupings have been disrupted.

There are other limits, however, to the Marr-Nishihara of object recognition. The first is that no neurons that code for general cylinders have been discovered in the brain. The second deals with the limitations of generalized cylinders themselves. There are many objects (for instance, an apple) that are not easily described in terms of these generalized cylinders. The grapes in the still-life at the bottom of *Les Demoiselles d'Avignon*, for example, are recognizable as grapes, but it is difficult to see how generalized cylinders might account for this perception. Clearly, some other process of object recognition is occurring.

Gestalt principles of organization seem necessary for the process of object recognition, for example, as an alternative explanation to generalized
cylinders for the failure to recognize the Female Nude. Among the Gestalt principles of organization are good continuation, similarity, proximity, common fate, closure, and figure/ground; these principles explain how the visual system groups and interprets patterns so that the resulting structure is as simple as possible (Wertheimer, 1974). The 1968 drawing on the theme of The Turkish Bath\(^4\) demonstrates many of these principles; most evident are good continuation and closure, which define and separate the figures in the drawing.

Other combinations of Gestalt principles are necessary to understand more abstract works. Seated Woman\(^5\), a painting which looks at first glance to be simply a collection of tilted geometric and organic shapes, is one such work. Rudolf Arnheim interprets the work in the following way: “The similarity of the geometric shapes used throughout the picture emphasizes the unity of the whole and understates the distinction between the woman and the screenlike background. The distinction is made clear by other means. Essentially a left slant is used for the figure, a right slant for the ground.” (Arnheim, 1954) Therefore, the axes (and color) of the shapes function as the primary organizational principles, according to Arnheim. While this interpretation is useful in distinguishing the otherwise obscured figure-ground elements of the painting (if the viewer makes that distinction), it is still quite conceivable the the viewer will not be able to recognize the subject matter of the painting. Even though Gestalt principles are important in grouping elements in a painting, the difficulty in recognizing a seated woman suggests that by themselves, Gestalt principles are inadequate for object recognition. In general, Gestalt principles tend to explain, rather than predict, how the visual system operates, since different grouping patterns can be generated by pitting different Gestalt principles against each other (Wertheimer, 1974).

A 1917 work, Harlequin and Woman with Necklace (*206), looks at first glance to be utterly dominated by Gestalt principles, particularly similarity, proximity, and perhaps most noticeably, figure-ground relationships (e.g., the famous face-vase illusion comes to mind in the profiles of the figures’ heads). If a purely Gestalt-based system is inadequate to recognize human figures (despite Arnheim’s interpretation of Seated Woman), how are we able to recognize the people in this painting? Unlike the 1910 Female Nude, in Harlequin, Gestalt principles operate, but the Marr-Nishihara generalized cylinders do not appear to be particularly salient. Also unlike the Female Nude, the figures in this work seem more recognizable as such. The Gestalt principles seem to work to organize the space into several parts, but recognition is perhaps best explained by a featural approach, rooted in Gestalt elements and most evident in the facial features, the hat, and the necklace. Therefore, here as in the still-life in Les Demoiselles d’Avignon, generalized cylinders do not appear to be absolutely necessary for object recognition. The difference is that in Harlequin, the generalized cylinders don’t account for the recognition of objects (human figures) that can typically be understood in terms of cylinders, while in Les Demoiselles, it is difficult to see how the grapes in the still-
life are ever broken into cylinders. Since neither Gestalt principles nor generalized cylinders can fully account for object recognition, some other process or processes are operating in object recognition.

Pentland’s system of object recognition attempts to overcome the limits of generalized cylinders by the use of superquadratic components, which include basic shapes (spheres, wedges, etc.) and deformations which preserve the smooth forms and do not introduce concavities (Humphreys and Bruce, 1989). These smooth outlines are important, Pentland argues, because contour information is so important in specifying an object’s three dimensional shape and in determining where one object occludes another. Pentland’s ideas provided the framework for another theory of object recognition, recognition-by-components (RBC), which has been proposed by Irving Biederman (1987). While his theory, like the Marr-Nishihara model, does not have any physiological basis, it is partly able to overcome the limitations of generalized cylinders. The model does this by incorporating a number of different basic units from which objects are built and recognized. These basic volumetric primitives are called geons (geometric icons). Each geon is distinguished from other geons by a set of defining characteristics, called non-accidental properties (Biederman, 1990). These non-accidental properties are derived ultimately from Gestalt principles of organization.

Among the important non-accidental properties are smooth continuation (either curved or straight), cotermination (the intersection of lines at vertices), parallelism (parallel or non-parallel), and symmetry (symmetrical or not symmetrical). Many non-accidental properties have to do with line intersections. The non-accidental properties of cotermination, parallelism, and symmetry can explain the illusion of the Ames chair, in which sticks hung in the air can be seen as making a chair when viewed from a particular peephole (Biederman, 1990). In reality, the sticks do not actually touch one another, but the grouping effects of these non-accidental properties appears to account for this illusion.

Biederman argues that about 24 geons exist, three of which are sufficient to describe any object at the entry-level classification (Biederman, 1990). Biederman also uses Marr’s term “generalized cone” to describe the set of shapes made up of a volume swept out by a cross-section moving along an axis, of which geons are a subset. Edges represent the primary information extracted from objects, according to Biederman. A view of an object that preserves the non-accidental properties of the geons will allow activation of the object model and therefore identification of the object, given only a few of the geons. Essentially, Biederman’s RBC theory is a feature theory of object recognition (Matlin, 1994).

Biederman has done much empirical work on object recognition. Most of these studies have involved presenting an object to subjects for around 100 milliseconds. One of his most convincing demonstrations involves the recognition of degraded line-drawings of objects, that is, how much and what type of information can be deleted from an object before the object is no longer
recognizable. The intersections of lines seem to be the most important regions in distinguishing geons from one another. In one study, Biederman deleted parts of line drawings of objects. The missing information consisted either of vertices or the contours in the middle of segments. When the same amount of visual information was removed from figures in the two conditions, subjects were still able to identify the object (such as a flashlight) in the condition where midsegment contours were removed, but were significantly impaired at object identification when vertex information was removed. This makes sense according to Biederman because most of the important geon information is in the vertices (Biederman, 1990). Curiously, a technique comparable to this kind of selective contour deletion is rare in Picasso, but Biederman’s results demonstrate the power of the right kind of visual information to describe objects in an extremely economical fashion.

A fundamental concern with features is overwhelmingly evident in Picasso’s œuvre. The concerns of Cubism to show objects from multiple directions simultaneously as the only “true” way to represent an object relate to being able to integrate many types of features without relying on a literal depiction of the object. In many cases, the salient characteristics of many of these objects give a clue to the identity of the object in the painting, even when fragmentation of the objects limits the applicability of generalized cylinders or Gestalt principles. *Violin and Grapes* (1912, *159) presents a complex set of fragmented shapes which seem to represent several objects from several points of view. Some salient characteristics seem to emerge which allow for object recognition. In this case, the curved outlines, indication of a bow, the / shapes, and the curved shape on the right give some clue that this is a violin, even though they are presented in a fragmented way and from multiple points of view. Despite this fragmentation, a featural approach (as advocated by Biederman) will still explain the recognition. Neither generalized cylinders nor an overwhelming Gestalt organization are particularly evident in this painting, but unlike the 1910 *Female Nude*, featural information and recognition are preserved. There is also some preservation of the placement of features with relation to each other, even though the connections between features are not articulated.

While Biederman’s system seems to provide a reasonable explanation of many phenomena in object recognition, it suffers from several important flaws. First, it is difficult either to prove or disprove the existence of geons. One must be able to disprove a theory’s postulates in order for it to be considered scientifically sound. Recognition-by-components is not well-defined enough to be considered a theory. Biederman simply assumes that geons exist, and although much of his experimental evidence is consistent with his ideas, he never succeeds absolutely in proving the existence of geons. Perhaps more importantly, as stated above, there is no physiological evidence for geons in the anatomy of the brain. Moreover, other evidence suggests that there may be multiple object recognition pathways in the brain in which different objects are recognized in different ways.
Such an alternative approach has been proposed by Martha Farah (1992). In contrast to the models discussed thus far, Farah's work investigates the possibility of specialized subsystems in object recognition. Farah reviewed cases of 99 patients with associative visual agnosia, in which there is considerable basic perceptual ability intact, but also problems in recognizing various types of objects. This corresponds to the highest levels of visual object recognition (as in Marr's 3D model). "Associative visual agnosia does not always affect the recognition of all types of stimuli equally . . . [This] suggests that there may be some division of labor within the visual recognition system." (Farah, 1992) The most common dissociations observed in these patients were pure alexia (extreme difficulty reading words) and prosopagnosia (the inability to recognize faces). Farah devised a model explaining the relative importance of the two hypothesized types of visual recognition ability for recognizing faces, common objects, and printed words. This model predicts that patients will either have difficulty in recognizing faces or written words, with the possibility of some impairment of other objects in either case. Out of the 99 patients, Farah found only two whose impairments violated her model, but in both cases there were important confounds.

What could be behind this pattern? Some of the most specialized visual stimuli are letters and faces. The processing patterns of objects and faces was also studied by Farah. Words appear to be decomposed into parts (letters); that is, one has to recognize letters before one can recognize the word. Faces, on the other hand, do not undergo decomposition into parts, but rather are recognized "holistically." (Bruce, 1988) Farah and Tanaka taught subjects to recognize a set of faces, along with a set of nonface objects, and then assessed the subject's ability to recognize both the whole patterns and their parts. The recognition of intact faces showed a greater disadvantage of parts to wholes. That is, subjects were better at recognizing a face's nose when it was on the face rather than by itself (Farah, 1992). The face inversion effect, in which it is extremely difficult to recognize pictures of even famous people when they are presented upside down, also suggests a holistic recognition process for faces.

There is other evidence that faces are recognized in a fundamentally different manner than other visual stimuli. Single-cell recordings in the temporal cortex of monkeys have shown a selective response to faces, with stronger reactions to whole faces in a particular orientation (Desimone, 1991). While this experiment was not performed on humans, it is not unreasonable to assume that a similar process operates in human vision. Studies on the innateness of face recognition abilities in humans give also suggest that faces are dealt with in a special way. For example, work by Morton and Johnson (1991) demonstrates that newborn infants have a special interest in face-like patterns, as they will track them to a greater extent than "scrambled" stimuli with the same amount of visual information in a different configuration. Research also suggests that older infants may be perceiving faces as wholes rather than as individual features (Goldstein, 1989).
This concept of holistic recognition of faces is very important in dealing with abstract art. The human face was abstracted in numerous ways by Picasso as much as the human figure and other objects, yet the vast majority of these abstracted faces present no recognition problems. Why? First, faces are not typically presented in exceedingly complex positions. Full frontal, profile, or three-quarter views account for the orientation of most faces in paintings; only rarely does one see extreme foreshortening or very unusual angles in paintings of faces. This small number of positions cuts down on the number of patterns of facial features one needs to be familiar with. Bottom-up information is preserved in these cases; Picasso doesn’t violate Gestalt properties or non-accidental properties in these paintings. The apparent innateness (or at least very easy acquisition) of facial recognition ability in infants implies that humans are easily equipped to perform this task. Finally, the fact that faces are recognized holistically means that less emphasis needs to be placed on individual facial features than on the configuration of the face as a whole. The implication of all of these factors is that the individual features may be quite abstracted, as long as their relation, and hence the general shape of the face, is not altered too drastically. For most abstract faces in Picasso, the abstraction is precisely at the level of features, and not at the arrangement of features. Despite the astonishing variety of Picasso’s distortions of faces (as any quick look at a selection of his works will demonstrate), in almost every case, the features roughly correspond to where they should be on an actual face. For example, the eyes are not placed in random locations on the face; they are almost always near the center or top of the head, roughly where eyes on a real face would be. [As a sample of the variety of these distortions, see Weeping Woman (1937, *344), Man with a Pipe (1915, *189), Woman with Pears (1909, *133), Seated Woman (1962, *443), and Woman with a Pillow (1969, *449).] Only in the most savage distortions, as in the Head of a Woman with Two Profiles (1939, *362), does holistic recognition of faces break down, at least in my perception. At this level of abstraction, however, I can still recognize individual facial features. In a work such as Head (1913, *171), even this level of recognition can be eliminated.

Given Farah’s distinction between faces and other type of objects, what is the relation of human figures to faces? Since faces are perceived holistically, the configuration of features is much more important than the individual features themselves. For words, which are recognized in a piecemeal manner, both features (letters) and their configuration (letter order, i.e., words) seem necessary; however, many features (letters) of words can be dropped without affecting how well the whole is recognized. “You probably could read most sentences fairly well even if only half of the letters were present. F-r-x—p-l-, t-s e-s- t-r—d t—s s—t-n—.” (Matlin, 1994) These letters and words are usually presented in a certain context, however, which influences how they are perceived (see below for discussion of context). Because recognizing familiar words in everyday situations (with context) is an overlearned activity for most people, is it only in alexics that word
Object Recognition in Picasso’s Abstract Art

recognition is piecemeal? Without a context, that is, in the case of very difficult, unfamiliar words, a piecemeal process (e.g., sounding out the word) also goes on, even in “normal” individuals. In most everyday contextual situations, then, configuration may be more important than individual features for both extremes of Farah’s model. This is true even if for words, features are recognized before the whole (since many features may be deleted without drastically impairing word recognition ability).

If we assume, however, that out of context words are read piecemeal and faces are seen holistically, the position of objects in Farah’s model is somewhat ambiguous. Is the recognition holistic or piecemeal, or does it contain elements of both? A certain percentage of both alexics and prosopagnosics also suffer from some object recognition impairment, which would support this idea; however, a better-defined concept of where a human figure would fit into Farah’s model seems necessary to resolve this issue.

The recognition of figures in Picasso’s work seems dependent on minimal information: generalized cylinders, as in Les Demoiselles d’Avignon and Centaures, femme, et oiseaux, or only a few features, as in Harlequin or Violin and Grapes. The configuration of features is also important; according to Biederman, geon arrangement distinguishes different objects (Biederman, 1990). Consider Bather with a Sand Shovel (1960, *439). Here the articulation of cylinders or features does not correspond to that of real people. Certain features are detectable, however, such as a face and perhaps a navel or a few limbs. Personally, however, I only recognize body parts, and not an integral figure, due to the articulation of the features. Is articulation of features more important with human figures than with other objects, such as the Violin of 1912? Answering questions such as this could provide some insight into the relation of objects to faces and words in Farah’s model.

Most of the discussion thus far has dealt with bottom-up influences in object recognition, that is, how we recognize elements in a painting given the visual information on the retina. With abstract paintings, there will be an enormous role of top-down factors. Top-down processing uses knowledge the viewer already has to help give meaning to the visual stimulus that is presented, and has been demonstrated in a wide variety of situations. The top-down influence is experimentally evident in reaction times and error rates. Palmer (1975) found that people were more likely to recognize an ambiguous figure in an appropriate context. In a kitchen scene, therefore, a loaf of bread was recognized more readily than a mailbox, which is visually similar in many ways to a loaf of bread (Matlin, 1994). Biederman (1981) flashed a city scene on a screen and asked subjects to locate the fire hydrant. He found that observers made more errors when the fire hydrant was, for example, placed on top of a mailbox than on the sidewalk. The viewer’s knowledge of the usual position of fire hydrants influenced their ability to recognize them (Goldstein, 1989). Vecera (1993) showed that image segmentation is performed faster on upright letters than inverted letters. This
difference in reaction time again indicates that knowledge is influencing some aspects of visual processing. The word superiority effect, in which subjects are better at identifying letters placed in the context of a word than letters that are not, is another example of the influence of top-down processing on perception (Matlin, 1994).

Context effects and top-down processing are extremely important in recognizing objects and determining meaning in abstract art. Picasso's work often presents ambiguous geometric patterns which appear to have some type of contextual meaning. How can a context be generated from only a few shapes, however? His 1921 painting, *Three Musicians* (*231)*, is a very interesting example of the context phenomenon. Most of the painting consists of flat geometric shapes. There does not seem to be anything inherent in most of these shapes that would tell the viewer the subject matter of the work. Two sets of elements are recognizable, however, and determine the painting's meaning. First, there appear to be three faces near the top of the piece. The recognizability of these shapes as faces can be traced to the abstraction being present primarily at the level of facial features, not in feature configuration. The faces are also in proportion to the rest of the (seated) figures. The second set of recognizable elements deals with the elements of the painting that give it a musical connotation: the shape with dots resembling a wind-instrument on the left-hand side, the shape resembling a guitar in the center, and the lines and dots resembling written musical patterns on the right side of the painting. Why does the viewer see these as dealing with music? Familiarity with written musical notation would seem to be an important element in making this association; however, here there are three lines per stave instead of the five in normal written music. The “notes” resemble eighth notes, but with their exaggerated size, the presence of only three lines per stave, and the lack of other musical elements (e.g., time signature, key signature, tempo) these elements do not exactly correspond to written music as most people know it. This pattern shows a remarkable parallel to the relationship between objects in the real world and objects in paintings. It would be reasonable to assume the same processes are operating in both conditions. The primary reason, then, that most people familiar with music notation read this element of the painting as such has to do with the physical features of the painted elements. The painted notes have enough in common with real music notation, and don’t overlap with any other type of stimuli, that a musical identity is paired with the signs. This identity is then related to the figures in the painting, and they become *Three Musicians*. Both bottom-up (shapes going from the retina to the brain that are related to previously stored information) and top-down (previously-stored information giving meaning to the picture as a whole, not just the recognizable parts) processes seem to be working in recognizing objects in this painting.

Titles can also influence perception. Consider again the 1913 *Head*. I myself cannot recognize the subject matter without the title. With that cue, however, a few features may be thought of as resembling facial features. The
circle near the top of the dark diagonal strip may signify an eye, the curved line connected to a long straight line below it, representing perhaps a nose and slightly smiling mouth. This influence of the title, another top-down phenomenon, is related to the Rat-Man Illusion, in which what you are told about an ambiguous shape ("This is a rat" versus "This is a man") influences your perception. For Head, it is not difficult to imagine an alternative interpretation, given another title.

Because there is really nothing inherent about the shapes in a painting to determine what they represent, other factors (including other shapes and features) can also contribute to the mind’s construction of what the shapes represent. Consider the 1930 painting, Bather (*279). If only the top third or so of the painting (the area above the water line) is viewed, the shapes that seem to wrap around near the bottom can be given various interpretations, which can vary from individual to individual. Given the title of the painting and the presence of what seems to be a face (features resembling eyes added to an overall shape and articulation which match up fairly closely to real faces), it is quite possible to interpret these shapes as representing the arms of the woman.

When the painting is seen in its entirety, however, the same shapes can be interpreted quite differently. In looking at the whole painting, there is a conflict between the initial interpretation and the second interpretation which can not coexist well (if one makes the assumption that the figure can only have one set of arms.) The rest of the figure is well-defined in terms of cylinders (even though they seem to have been hollowed out in places), and the articulations and proportions are relatively intact. The same shapes, then, can lead to vastly different interpretations given the context and active processing of the visual information. This phenomenon is akin to reading T/AE C/AT, in which the context affects perceptions of identical but ambiguous shapes to derive some meaning (THE CAT). As with letters in the previous example, this is a difficult phenomenon to explain in painting. No single theory or set of phenomena, including pictorial depth cues, Gestalt organizational principles, the Marr-Nishihara model, Farah’s model, and Biederman’s RBC theory, can fully account for this context effect.

In considering these various models, we have overlooked an important issue: What was Picasso’s own view of how objects could be recognized in his own work? He may have given us a hint when he said, “I don’t draw what I see. I draw what I know.” What could Picasso have meant by this statement? In The Artist and his Model (1964), Marie-Laure Bernadac interprets the recognizability of the objects in terms of distinguishing features: “the artist’s face is suggested by a sort of X which links the eyes, the nose, and the mouth, and the model is reduced to her basic contours.” (Bernadac, 1988) Picasso himself said, “A dot for the breast, a line for the artist, five spots of paint for the foot, a few pink and green lines—that’s enough, isn’t it? What more need I do? What can I add to that? It’s all been said.” These simplifications correspond to Picasso’s desire to “adopt codified signs that summarize and signify each
part of the body, and whose total image expresses the nude.” (Bernadac, 1988)

Supporting this is Picasso’s statement, “Things have got to be named . . . I want to say the nude; I don’t want just to make a nude like a nude; I just want to say breast, to say foot, to say hand, belly—find a way to say it and that’s enough.” (Bernadac, 1988) The implication of these remarks is that Picasso consciously focused on distinguishing features when representing objects.

In examining a large number of pieces from his late period, when he distorted the proportions of the human figure in extreme ways, there is an almost obsessive concern with featural ideas. In the drawing on the theme of The Turkish Bath, a cursory evaluation of the figures in terms of distinguishing features reveals some important consistencies in the depiction of these features. Most noticeable is the distinct presence in many of the figures of features such as navels, nipples, genitals, well-defined facial features (even including eyelashes), and five digits (often including nails) per hand or foot. The detail of such features is even more evident in other drawings and etchings from Picasso’s late period. It is as though their absence would severely hinder the viewer’s ability to recognize the figures as such, especially given the extreme distortion in the proportions of many of these figures from this period. While one could argue that this characteristic is a peculiarity of Picasso’s late style, the importance of featural information throughout his long and highly varied career, as well as his own comments, suggest that a featural approach can describe his method of maintaining recognizability of objects despite extreme distortions. It is reasonable to assume the same processes or components that made the objects Picasso was painting recognizable to him are the same processes we use to recognize the objects in the finished work. Whether these distinguishing features correspond, for example, to Biederman’s geons is another matter.

How these features are incorporated into the human object recognition system remains a mystery. In certain works, featural information is seemingly minimized, but the figures are still recognizable. (This is also perhaps due to Gestalt principles and generalized cylinders, or perhaps generalized cylinders can be considered a subclass of features or geons.) Also, how do features account for the perception of multiple objects simultaneously? Picasso’s Woman-Flower (1946, *394) possesses attributes of both a woman and a plant. How would Biederman’s RBC theory reconcile the distinctiveness of geons for different objects with this case, where two different objects are synthesized?

The enormous number of works by Picasso provide an extraordinarily varied testing ground for object recognition theories. Pictorial depth cues, Gestalt principles of organization, and the Marr-Nishihara model of object recognition provide fairly good explanations of how objects are recognized, when these models are applicable. Picasso’s own comments and inferences one can make from his work suggest that a feature approach may explain most object recognition processes. Among other problems, however, a feature theory such as Biederman’s RBC theory may apply differently to faces
Object Recognition in Picasso’s Abstract Art

and other objects, as demonstrated by Farah. Farah’s model, though, is unclear as to the exact nature of how nonface and nonword objects are recognized by object recognition subsystems. Moreover, none of the models are able to fully explain context effects. Perhaps some other theories which have not yet been examined in detail will better explain these phenomena. For example, a prototype explanation of pattern recognition (advocated by Posner and Keele (1968) and Franks and Bransford (1971)) has not been directly applied to object recognition. Klatzky (1980) has pointed out that a feature approach is compatible with a prototype approach. Future research may shed more light on the relation of prototypes and features to the other theories. As all the approaches examined here have some applicability in different situations, it seems that any comprehensive theory of object recognition will incorporate some elements from each of these theories. Pablo Picasso’s work should prove to be a rich testing ground for that theory as well.

References


Notes

1. Because of copyright difficulties, works by Picasso could not be reproduced here. The necessity of seeing the relevant work by Picasso in order to understand and apply the theories of object recognition is obvious, however. To circumvent this problem, relevant works by Picasso will be listed in the text, and the interested reader is encouraged to consult the proper reference to see them. All but a few of the examples may be found in Rubin’s excellent and thorough survey of the artist’s output, *Pablo Picasso: A Retrospective*. These are marked in the text with (*page number in Rubin). Works not found in Rubin will be referenced with the appropriate plate information. *The Turkish Bath* is plate 367 in Picasso’s 1969 book, *Picasso: His Recent Drawings, 1966–1968*. The motivated reader may wish to examine further works in Picasso’s oeuvre from Rubin or other retrospectives to extend the ideas presented here to other works (or other artists).

2. Richet, page 400, plate 1271.


5. Arnheim, Plate 1, opposite page 58.

6. This two-line feature is also noticeable in other works of this period, such as the *Head of a Man* (*173) and *Head of a Harlequin* (*169), both of 1913. These examples hint at the effect of learning an artist’s visual “vocabulary,” which aids object recognition in an “expert” familiar with the artist’s style. This phenomenon is directly related to top-down processing.

7. Interestingly, this statement is contrary to the general advice given to novice artists (that is, to try to forget what you know and draw only what you see) when trying to get them to produce realistic (and presumably recognizable) renderings of objects in the visual world. This standard advice is related to Robert Thouless’ 1931 study on shape constancy, which demonstrated that most people combine the standard shape of an object (e.g., the circularity of a coin) with the perceived shape (the elliptical appearance of a coin on a table, viewed from above) to produce partial shape constancy. (Goldstein, 1989)