Look around the area where you are now located. If you are inside, you are likely to see furniture, windows, walls, and a floor; if you are outside, you are likely to see plants, the sky, and perhaps a road or a sidewalk. Of course, you do not really directly see these objects. You actually detect details like edges, movement, and spatial clues about distances, then these features and properties somehow combine into meaningful wholes. How does simple feature detection become perception of multidimensional, complete objects? This perceptual task, which humans do effortlessly, has turned out to be remarkably difficult for computers to accomplish, much more so than complex tasks such as chess or solving mathematical problems.

Treisman has described some general conclusions about visual perception that are emerging from research done in her lab and elsewhere. The visual perception of an object appears to involve three steps. The first step is coding simple features with reference to maps that preserve spatial (i.e., location) information. The second step requires focused attention (i.e., paying attention to the object being perceived) and results in a temporary representation of the object. The third step is to compare the temporary representation with descriptions of objects stored in memory. If a match is found between the temporary representation and a stored description, the object has been recognized, completing the process of visual perception.

Why should an article on vision written for a general scientific magazine be included in a book of readings in cognitive psychology? For one consideration, its author is one of the foremost contributors in the world to the field of vision. This article gives readers a chance to learn about her work in a forum that is more accessible than some of the original journal articles that have published the work. For another, this article gives a broader account of vision than would many individual journal articles. In this article, Treisman displays an important skill that many cognitive psychologists lack: the ability to describe complicated concepts and findings from their work and that of others in a way that is accessible to the general reader. Such a task, easy as it sounds, in fact is quite a difficult challenge. One has to make sure that the ideas and findings are presented correctly and

clearly and in a way that general readers can understand. One has to know not only what details to include, but what details to omit. Also, one has to know how to relate one’s own work to that of others in a way that is informative and not self-serving. Treisman is among the best writers of such articles.

The seemingly effortless ability to perceive meaningful wholes in the visual world depends on complex processes. The features automatically extracted from a scene are assembled into objects.

If you were magically deposited in an unknown city, your first impression would be of recognizable objects organized coherently in a meaningful framework. You would see buildings, people, cars, and trees. You would not be aware of detecting colors, edges, movements, and distances, and of assembling them into multidimensional wholes for which you could retrieve identities and labels from memory. In short, meaningful wholes seem to precede parts and properties, as the Gestalt psychologists emphasized many years ago.

This apparently effortless achievement, which you repeat innumerable times throughout your waking hours, is proving very difficult to understand or to simulate on a computer—much more difficult, in fact, than the understanding and simulation of tasks that most people find quite challenging, such as playing chess or solving problems in logic. The perception of meaningful wholes in the visual world apparently depends on complex operations to which a person has no conscious access, operations that can be inferred only on the basis of indirect evidence.

Nevertheless, some simple generalizations about visual information processing are beginning to emerge. One of them is a distinction between two levels of processing. Certain aspects of visual processing seem to be accomplished simultaneously (that is, for the entire visual field at once) and automatically (that is, without attention being focused on any one part of the visual field). Other aspects of visual processing seem to depend on focused attention and are done serially, or one at a time, as if a mental spotlight were being moved from one location to another.

In 1967, Ulric Neisser, then at the University of Pennsylvania, suggested that a “preattentive” level of visual processing segregates regions of a scene into figures and ground so that a subsequent, attentive level can identify particular objects. More recently, David C. Marr, investigating computer simulation of vision at the Massachusetts Institute of Technology, found it necessary to establish a “primal sketch”: a first stage of processing, in which the pattern of light reaching an array of receptors is converted into a coded description of lines, spots, or edges and their locations, orientations, and colors. The representation of surfaces and volumes and finally the identification of objects could begin only after this initial coding.

In brief, a model with two or more stages is gaining acceptance among psychologists, physiologists, and computer scientists working in artificial intelligence. Its first stage might be described as the extraction of features from patterns of light; later stages are concerned with the identification of objects and their settings. The phrase
"features and objects" is therefore a three-word characterization of the emerging hypothesis about the early stages of vision.

I think there are many reasons to agree that vision indeed applies specialized analyzers to decompose stimuli into parts and properties, and that extra operations are needed to specify their recombination into the correct wholes. In part the evidence is physiological and anatomical. In particular, the effort to trace what happens to sensory data suggests that the data are processed in different areas of considerable specialization. One area concerns itself mainly with the orientation of lines and edges, another with color, still another with directions of movement. Only after processing in these areas do data reach areas that appear to discriminate between complex natural objects.

Some further evidence is behavioral. For example, it seems that visual adaptation (the visual system's tendency to become unresponsive to a sustained stimulus) occurs separately for different properties of a scene. If you stare at a waterfall for a few minutes and then look at the bank of the river, the bank will appear to flow in the opposite direction. It is as if the visual detectors had selectively adapted to a particular direction of motion independent of whatever is moving. The bank looks very different from the water, but it nonetheless shows the aftereffects of the adaptation process.

How can the preattentive aspect of visual processing be further subjected to laboratory examination? One strategy is suggested by the obvious fact that in the real world parts that belong to the same object tend to share properties: they have the same color and texture, their boundaries show a continuity of lines or curves, they move together, they are at roughly the same distance from the eye. Accordingly the investigator can ask subjects to locate the boundaries between regions in various visual displays and thus can learn what properties make a boundary immediately salient—make it "pop out" of a scene. These properties are likely to be the ones the visual system normally employs in its initial task of segregating figure from ground.

It turns out that boundaries are salient between elements that differ in simple properties such as color, brightness, and line orientation but not between elements that differ in how their properties are combined or arranged (Figure 1). For example, a region of Ts segregates well from a region of tilted Ts but not from a region of Is made of the same components as the Ts (a horizontal line and a vertical line). By the same token, a mixture of blue Vs and red Os does not segregate from a mixture of red Vs and blue Os. It seems that the early "parsing" of the visual field is mediated by separate properties, not by particular combinations of properties. That is, analysis of properties and parts precedes their synthesis. And if parts or properties are identified before they are conjoined with objects, they must have some independent psychological existence.

This leads to a strong prediction, which is that errors of synthesis should sometimes take place. In other words, subjects should sometimes see illusory conjunctions of parts or properties drawn from different areas of the visual field. In certain conditions such illusions take place frequently. In one experiment my colleagues and I flashed three colored letters, say a blue X, a green T, and a red O, for a brief period (200 milliseconds, or a fifth of a second) and diverted our subjects' attention by asking them to report first a digit shown at each side of the display and only then the colored letters. In about one trial in three, the subjects reported the wrong combinations—perhaps a red X, a green O, or a blue T.
Features and Objects in Visual Processing

**Figure 1** Boundaries that "pop out" of a scene are likely to reveal the simple properties, or features, of the visual world that are seized on by the initial stage of visual processing. For example, a boundary between T's and tilted T's pops out, whereas a boundary between T's and L's does not (a). The implication is that line orientations are important features in early visual processing but that particular arrangements of conjunctions of lines are not. A boundary between O's and V's pops out (b). The implication is that simple shape properties (such as line curvature) are important.

The subjects made these conjunction errors much more often than they reported a color or shape that was not present in the display, which suggests that the errors reflect genuine exchanges of properties rather than simply misperceptions of a single object. Many of these errors appear to be real illusions, so convincing that subjects demand to see the display again to convince themselves that the errors were indeed mistakes.
We have looked for constraints on the occurrence of such illusory conjunctions. For example, we have asked whether objects must be similar for their properties to be exchanged. It seems they do not: Subjects exchanged colors between a small, red outline of a triangle and a large, solid blue circle just as readily as they exchanged colors between two small outline triangles. It is as if the red color of the triangle were represented by an abstract code for red rather than being incorporated into a kind of analogue of the triangle that also encodes the object's size and shape.

We also asked if it would be harder to create illusory conjunctions by detaching a part from a simple unitary shape, such as a triangle, than by moving a loose line. The answer again was no. Our subjects saw illusory dollar signs in a display of $S$s and lines. They also saw the illusory signs in a display of $S$s and triangles in which each triangle incorporated the line the illusion required (Figure 2). In conscious experience the triangle looks like a cohesive whole. Nevertheless, at the preattentive level, its component lines seem to be detected independently.

To be sure, the triangle may have an additional feature, namely the fact that its constituent lines enclose an area, and this property of closure might be detected preattentively. If so, the perception of a triangle might require the detection of its three component lines in the correct orientations and also the detection of closure. We should then find that subjects do not see illusory triangles when they are given only the triangles' separate lines in the proper orientations (Figure 3). They may need a further stimulus, a different closed shape (perhaps a circle), in order to assemble illusory triangles. That is indeed what we found.

Another way to make the early, preattentive level of visual processing the subject of laboratory investigation is to assign visual-search tasks. That is, we ask subjects to find a target item in the midst of other, "distractor" items. The assumption is that if the preattentive processing occurs automatically and across the visual field, a target that

**Figure 2** Illusory dollar signs are an instance of false conjunctions of features. Subjects were asked to look for dollar signs in the midst of $S$s and line segments (a). They often reported seeing the signs when the displays to which they were briefly exposed contained none (b). They had the same experience about as often when the line segment needed to complete a sign was embedded in a triangle (c). The experiment suggests that early visual processing can detect the presence of features independent of location.
Illusory triangles constitute a test of what features must be available to support the perception of triangles. Subjects seldom reported seeing a triangle when they were briefly exposed to displays consisting of the line segments that make up a triangle (a). They saw triangles far more often when the displays also included closed stimuli, that is, shapes that enclose a space, in this case Os (b). Evidently, closure is a feature analyzed in early visual processing. This conclusion was supported by showing displays that lack the diagonal line to make a triangle (c, d). Subjects seldom saw triangles in such displays.

is distinct from its neighbors in its preattentive representation in the brain should “pop out” of the display. The proverbial needle in a haystack is hard to find because it shares properties of length, thickness and orientation with the hay in which it is hidden. A red poppy in a haystack is a much easier target; its unique color and shape are detected automatically.

We find that if a target differs from the distractors in some simple property, such as orientation or color or curvature, the target is detected about equally fast in an array of 30 items and in an array of three items. Such targets pop out of the display, so that the time it takes to find them is independent of the number of distractors. This independence holds true even when subjects are not told what the unique property of the target will be. The subjects take slightly longer overall, but the number of distractors still has little or no effect.

On the other hand, we find that if a target is characterized only by a conjunction of properties (for example, a red O among red Vs and green Os), or if it is defined only by its particular combination of components (for example, an R among Ps and Qs that
together incorporate all the parts of the \( R \), the time taken to find the target or to
decide that the target is not present increases linearly with the number of distractors. It
is as if the subjects who are placed in these circumstances are forced to focus attention
in turn on each item in the display in order to determine how the item's properties or
parts are conjoined. In a positive trial (a trial in which a target is present) the search
ends when the target is found; on the average, therefore, it ends after half of the distrac-
tors have been examined. In a negative trial (in which no target is present) all the dis-
tractors have to be checked. As distractors are added to the displays, the search time in
positive trials therefore increases at half the rate of the search time in negative trials.

The difference between a search for simple features and a search for conjunc-
tions of features could have implications in industrial settings. Quality-control inspec-
tors might, for example, take more time to check manufactured items if the possible
errors in manufacture are characterized by faulty combinations of properties than
they do if the errors always result in a salient change in a single property. Similarly,
each of the symbols representing, say, the destinations for baggage handled at airline
terminals should be characterized by a unique combination of properties.

In a further series of experiments on visual-search tasks, we explored the effect of
exchanging the target and the distractors. That is, we required subjects to find a target
distinguished by the fact that it lacks a feature present in all the distractors. For ex-
ample, we employed displays consisting of Os and Qs, so that the difference between
the target and the distractors is that one is simply a circle whereas the other is a circle
intersected by a line segment (Figure 4). We found a remarkable difference in the
search time depending on whether the target was the Q and had the line or was the O
and lacked the line. When the target had the line, the search time was independent of
the number of distractors. Evidently, the target popped out of the display. When the
target lacked the line, the search time increased linearly with the number of distrac-
tors. Evidently, the items in the display were being subjected to a serial search.

The result goes against one's intuitions. After all, each case involves the same dis-
rimination between the same two stimuli: Os and Qs. The result is consistent, how-
ever, with the idea that a pooled neural signal early in visual processing conveys the
presence but not the absence of a distinctive feature. In other words, early vision
extracts simple properties, and each type of property triggers activity in populations
of specialized detectors. A target with a unique property is detected in the midst of
distractor items simply by a check on whether the relevant detectors are active. Con-
versely, a target lacking a property that is present in the distractors arouses only
slightly less activity than a display consisting exclusively of distractors. We propose,
therefore, that early vision sets up a number of what might be called feature maps.
They are not necessarily to be equated with the specialized visual areas that are
mapped by physiologists, although the correspondence is suggestive.

We have exploited visual-search tasks to test a wide range of candidate features
we thought might pop out of displays and so reveal themselves as primitives: basic
elements in the language of early vision. The candidates fell into a number of cate-
gories: quantitative properties such as length or number; properties of single lines
such as orientation or curvature; properties of line arrangements; topological and
relational properties such as the connectedness of lines, the presence of the free ends
of lines or the ratio of the height to the width of a shape.
Among the quantitative candidates, my colleagues and I found that some targets popped out when their discriminability was great. In particular, the more extreme targets—the longer lines, the darker grays, the pairs of lines (when the distractors were single lines)—were easier to detect. This suggests that the visual system responds positively to “more” in these quantitative properties and that “less” is coded by default. For example, the neural activity signaling line length might increase with increasing length (up to some maximum), so that a longer target is detected against the lower level of background activity produced by short distractors. In contrast, a shorter target, with its concomitant lower rate of firing, is likely to be swamped by the greater activity produced by the longer distractors. Psychophysicists have known for more than a century that the ability to distinguish differences in intensity grows more acute with decreasing background intensity. We suggest that the same phenomenon, which is known as Weber’s law, could account for our findings concerning the quantitative features.

Our tests of two simple properties of lines, orientation and curvature, yielded some surprises. In both cases we found pop-out for one target, a tilted line among vertical distractors and a curved line among straight lines, but not for the converse target, a vertical line among tilted distractors and a straight line among curves. These findings suggest that early vision encodes tilt and curvature but not verticality or straightness. That is, the vertical targets and the straight targets appear to lack a feature the distractors possess, as if they represent null values on their respective dimensions. If our interpretation is correct, it implies that in early vision, tilt and curvature are represented relationally, as deviations from a standard or norm that itself is not positively signaled.

A similar conclusion emerged for the property of closure. We asked subjects to search for complete circles in the midst of circles with gaps and for circles with gaps among complete circles. Again we found a striking asymmetry, this time suggesting that the gap is preattentively detectable but that closure is not—or rather that it becomes preattentively detectable only when the distractors have very large gaps (that is, when they are quite open shapes like semicircles). In other words, closure is preattentively detectable, but only when the distractors do not share it to any significant degree. On the other hand, gaps (or the line ends that gaps create) are found equally easily whatever their size (unless they are too small for a subject, employing peripheral vision, to see).

Finally, we found no evidence that any property of line arrangements is preattentively detectable. We tested intersections, junctions, convergent lines and parallel lines. In every case we found that search time increases with an increasing number of distractors. The targets become salient and obvious only when the subject’s attention is directed to them; they do not emerge automatically when that attention is disseminated throughout the display.

In sum, it seems that only a small number of features are extracted early in visual processing. They include color, size, contrast, tilt, curvature, and line ends. Research by other investigators shows that movement and differences in stereoscopic depth are also extracted automatically in early vision. In general the building blocks of vision appear to be simple properties that characterize local elements, such as points or lines, but not the relations among them. Closure appears to be the most complex property that pops out preattentively. Finally, our findings suggest that several preattentive properties are coded as values of deviation from a null, or reference, value.
**FIGURE 4** Presence or absence of a feature can have remarkably different effects on the time it takes to find a target in the midst of distractors. In one experiment (a) the target was a circle intersected by a vertical line segment or a circle without that feature. The search time for the intersected circle (solid) proved to be largely independent of the number of items in the display, suggesting that the feature popped out. The search time for the plain circle (dashed) increased steeply as distractors were added, suggesting that a serial search of the display was being made. A second experiment (b) required subjects to search for a vertical line (dashed) or a tilted line (solid). The tilted line could be found much faster; evidently only the tilted line popped out of the displays. A third experiment (c) tested an isolated line segment (dashed) or intersecting lines in the form of a plus sign (solid). Evidently neither popped out.
A fourth experiment (d) tested parallel lines (dashed) or converging lines (solid). Again neither popped out. A fifth experiment (e) tested closure with complete circles (dashed) or circles with a gap of a fourth of their circumference (solid). A sixth experiment (f), again testing closure, had complete circles (dashed) or circles with smaller gaps (solid). The size of the gap seemed to make no difference: The incomplete circle popped out. On the other hand, a complete circle became harder to find as the size of the gaps in distractors was reduced. Open dots represent data from trials in which the display included only distractors.
Up to this point I have concentrated on the initial, preattentive stages of vision. I turn now to the later stages. In particular I turn to the evidence that focused attention is required for conjoining the features at a given location in a scene and for establishing structured representations of objects and their relations.

One line of evidence suggesting that conjunctions require attention emerges from experiments in which we asked subjects to identify a target in a display and say where it was positioned. In one type of display only a simple feature distinguished the target from the distractors. For example, the target was a red H in the midst of red Os and blue Xs or an orange X among red Os and blue Xs. In other displays, the target differed only in the way its features were conjoined. For example, it was a blue O or a red X among red Os and blue Xs.

We were particularly interested in the cases in which a subject identified the target correctly but gave it the wrong location. As we expected, the subjects could sometimes identify a simple target, say a target distinguished merely by its color, but get its location wrong. Conjunction targets were different. The correct identification was completely dependent on the correct localization. It does indeed seem that attention must be focused on a location in order to combine the features it contains.

In a natural scene, of course, many conjunctions of features are ruled out by prior knowledge. You seldom come across blue bananas or furry eggs. Preattentive visual processing might be called “bottom up,” in that it happens automatically, without any recourse to such knowledge. Specifically, it happens without recourse to “top down” constraints. One might hypothesize that conjunction illusions in everyday life are prevented when they conflict with top-down expectations. There are many demonstrations that we do use our knowledge of the world to speed up perception and to make it more accurate. For example, Irving Biederman of the State University of New York at Buffalo asked subjects to find a target object such as a bicycle in a photograph of a natural scene or in a jumbled image in which different areas had been randomly interchanged. The subjects did better when the bicycle could be found in a natural context (see the photo).

In order to explore the role of prior knowledge in the conjoining of properties, Deborah Butler and I did a further study of illusory conjunctions. We showed subjects a set of three colored objects flanked on each side by a digit. Then, some 200 milliseconds later, we showed them a pointer, which was accompanied by a random checkerboard in order to wipe out any visual persistence from the initial display. We asked the subjects to attend to the two digits and report them, and then to say which object the pointer had designated. The sequence was too brief to allow the subjects to focus their attention on all three objects.

The crucial aspect of the experiment lay in the labels we gave the objects. We told one group of subjects that the display would consist of “an orange carrot, a blue lake, and a black tire.” Occasional objects (one in four) were shown in the wrong color to ensure that the subjects could not just name the color they would know in advance ought to be associated with a given shape. For another group of subjects the same display was described as “an orange triangle, a blue ellipse, and a black ring.”

The results were significant. The group given arbitrary pairings of colors and shapes reported many illusory conjunctions: 29 percent of their responses represented illusory recombinations of colors and shapes from the display, whereas 13 percent were reports of colors or shapes not present in the display. In contrast, the
Prior knowledge as a guide in visual perception is tested by asking subjects to search for a familiar object in a photograph of an unexceptional scene (top) and in a jumbled photograph of the scene (bottom). Here, the task is simply to find the bicycle. It tends to take longer in the jumbled image. The implication is that knowledge of the world (in this case, expectations about the characteristic locations of bicycles in urban landscapes) speeds up perception and makes it less subject to error. Certain early aspects of the information processing that underlies visual perception nonetheless seem to happen automatically: without the influence of prior knowledge. The illustration was modeled after experiments done by Irving Biederman of the State University of New York at Buffalo.

A group expecting familiar objects saw rather few illusory conjunctions: They wrongly recombined colors and shapes only 5 percent more often than they reported colors and shapes not present in the display.
We occasionally gave a third group of subjects the wrong combinations when they were expecting most objects to be in their natural colors. To our surprise we found no evidence that subjects generated illusory conjunctions to fit their expectations. For example, they were no more likely to see the triangle (the "carrot") as orange when another object in the display was orange than they were when no orange was present. There seem to be two implications: Prior knowledge and expectations do indeed help one to use attention efficiently in conjoining features, but prior knowledge and expectations seem not to induce illusory exchanges of features to make abnormal objects normal again. Thus illusory conjunctions seem to arise at a stage of visual processing that precedes semantic access to knowledge of familiar objects. The conjunctions seem to be generated preattentively from the sensory data, bottom-up, and not to be influenced by top-down constraints.

How are objects perceived once attention has been focused on them and the correct set of properties has been selected from those present in the scene? In particular, how does one generate and maintain an object's perceptual unity even when objects move and change? Imagine a bird perched on a branch, seen from a particular angle and in a particular illumination. Now watch its shape, its size, and its color all change as it preens itself, opens its wings, and flies away. In spite of these major transformations in virtually all its properties, the bird retains its perceptual integrity: It remains the same single object.

Daniel Kahneman of the University of California at Berkeley and I have suggested that object perception is mediated not only by recognition, or matching to a stored label or description, but also by the construction of a temporary representation that is specific to the object's current appearance and is constantly updated as the object changes. We have drawn an analogy to a file in which all the perceptual information about a particular object is entered, just as the police might open a file on a particular crime, in which they collect all the information about the crime as the information accrues. The perceptual continuity of an object would then depend on its current manifestation being allocated to the same file as its earlier appearances. Such allocation is possible if the object remains stationary or if it changes location within constraints that allow the perceptual system to keep track of which file it should belong to.

In order to test this idea we joined with Brian Gibbs in devising a letter-naming task (Figure 5). Two letters were briefly flashed in the centers of two frames. The empty frames then moved to new locations. Next, another letter appeared in one of the two frames. We devised the display so that the temporal and spatial separations between the priming letter and the final letter were always the same; the only thing that differed was the motion of the frames. The subjects’ task was to name the final letter as quickly as possible.

We knew that the prior exposure to a given letter should normally lessen the time it takes to identify the same letter on a subsequent appearance; the effect is known as priming. The question that interested us was whether priming would occur only in particular circumstances. We argued that if the final letter is the same as the priming letter and appears in the same frame as the priming letter, the two should be seen as belonging to the same object; in this case, we could think of the perceptual task as simply re-viewing the original object in its shifted position. If, on the other
Integration of sensory information into what amounts to a file on each perceptual object was tested by the motion of frames. In each trial, two frames appeared, then two letters were briefly flashed in the frames (a). The frames moved to new locations, and a letter appeared in one of the two (b). The subject's task was to name the final letter as quickly as possible. If the final letter matched the initial letter and appeared in the same frame, the naming was faster than if the letter had appeared in the other frame or differed from the initial letter. The implication is that it takes more time to create or update a file on an object than it does simply to perceive the same object a second time.

Hand, a new letter appears in the same frame, the object file should have to be updated, perhaps increasing the time it takes for subjects to become aware of the letter and name it.

Actually the priming was found to be object-specific: Subjects named the final letter some 30 milliseconds faster if the same letter had appeared previously in the same frame. They showed no such benefit if the same letter had appeared previously in the other frame. The result is consistent with the hypothesis that the later stages of visual perception integrate information from the early, feature-sensitive stages in temporary object-specific representations.

The overall scheme I propose for visual processing can be put in the form of a model (Figure 6). The visual system begins by coding a certain number of simple and useful properties in what can be considered a stack of maps. In the brain such maps ordinarily preserve the spatial relations of the visual world itself. Nevertheless, the spatial information they contain may not be directly available to the subsequent stages of visual processing. Instead the presence of each feature may be signaled without a specification of where it is.

In the subsequent stages, focused attention acts. In particular, focused attention is taken to operate by means of a master map of locations, in which the presence of discontinuities in intensity or color is registered without specification of what the discontinuities are. Attention makes use of this master map, simultaneously selecting, by
Hypothetical model of the early stages in visual perception emerges from the author’s experiments. The model proposes that early vision encodes some simple and useful properties of a scene in a number of feature maps, which may preserve the spatial relations of the visual world but do not themselves make spatial information available to subsequent processing stages. Instead, focused attention (employing a master map of locations) selects and integrates the features present at particular locations. At later stages, the integrated information serves to create and update files on perceptual objects. In turn, the file contents are compared with descriptions stored in a recognition network. The network incorporates the attributes, behavior, names, and significance of familiar objects.
means of links to the separate feature maps, all the features that currently are present in a selected location. These are entered into a temporary object representation, or file.

Finally, the model posits that the integrated information about the properties and structural relations in each object file is compared with stored descriptions in a "recognition network." The network specifies the critical attributes of cats, trees, bacon and eggs, one's grandmothers, and all other familiar perceptual objects, allowing access to their names, their likely behavior, and their current significance. I assume that conscious awareness depends on the object files and on the information they contain. It depends, in other words, on representations that collect information about particular objects, both from the analyses of sensory features and from the recognition network, and continually update the information. If a significant discontinuity in space or time occurs, the original file on an object may be canceled: it ceases to be a source of perceptual experience. As for the object, it disappears and is replaced by a new object with its own new temporary file, ready to begin a new perceptual history.

**Questions for Further Thought**

1. What are the two main levels of processing in visual perception, and how do they differ?
2. Describe the kinds of visual-search tasks Treisman has used in her research.
3. What is pop-out?
4. Why did Treisman conclude that a serial search was conducted for a plain circle as distractors were added?
5. What is the role of attention in visual processing?