

## Detecting high-level and low-level properties in visual images and visual percepts

Romke Rouw<sup>a,\*</sup>, Stephen M. Kosslyn<sup>b</sup>, Ronald Hamel<sup>a</sup>

<sup>a</sup>*Department of Psychonomics, University of Amsterdam, Amsterdam, The Netherlands*

<sup>b</sup>*Department of Psychology, Harvard University, Cambridge, MA, USA*

Received 17 April 1996, final version 23 January 1997

---

### Abstract

In this article we provide further evidence that visual mental imagery and visual perception share modality-specific mechanisms, and we find that representing visual information in a mental image (activating stored information to create a picture-like mental representation) preserves relatively low-level visual detail. Subjects either saw or visualized simple pictures, and evaluated them for the presence or absence of six types of non-accidental properties. These properties varied from very 'low-level' ones, such as T junctions, to very 'high-level' ones, such as global symmetry. The question was whether both sorts of information are equally accessible in percepts and mental images. If mental images are equivalent to descriptions of perceptual units and their organization, as some have argued, then subjects should have greater difficulty accessing low-level properties in a mental image compared to the difficulty they experience when the drawing is visible. The results of two experiments were clearcut: Subjects could evaluate high-level properties more easily than low-level ones, but this difference was the same in imagery and perception. These findings suggest that mental images preserve relatively low-level visual features, and are not simply descriptions of a pattern.

---

### 1. Introduction

Inspecting an object or scene in a visual mental image is accompanied by a subjective experience similar to inspecting an object or scene that is physically present. The similarity of the experiences suggests that the information processing underlying visual imagery and visual perception is similar, but this claim has been

\* Corresponding author. Corresponding address: Tilburg University, Department of Psychology, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Fax: (31-13) 466-2370; e-mail: r.rouw@kub.nl

the subject of debate (for reviews, see Farah, 1988; Finke, 1989; Kosslyn, 1994; Shepard and Cooper, 1982). In this article we provide further evidence that imagery and perception share modality-specific mechanisms, and argue that image representations preserve relatively low-level visual details that are present in a percept.

One of the advances of cognitive psychology (e.g., Neisser, 1967) that has proven enduring is the insight that most cognitive abilities are subserved by a collection of processes, and that to understand an ability one must understand the componential structure of the system that accomplishes it (see Kosslyn and Koenig, 1992). For visual mental imagery, we can distinguish between abilities that create a visual mental image, those that transform patterns in images, those that maintain images over time, and those that interpret the pattern in an image (see Kosslyn, 1994). In this article we focus on what sorts of information are accessible to the ‘image inspection processes,’ which interpret the patterns in an image as depicting specific objects, parts, or other visual properties. One might, for example, never have explicitly thought about what shape a German Shepherd dog’s ears are, but when asked one can derive the right answer by forming a mental image of the dog, and then ‘looking’ at the ears. Thus, as in inspecting a perceived scene, an imaged object can be inspected.

Over the past ten years, an important issue has been raised in the study of image inspection: Can one inspect objects in images with the same flexibility that one examines actual objects, or can one only ‘see’ parts and properties that were explicitly included in the image? That is, a visual percept results from looking at a physically present stimulus, which is organized and interpreted during the act of perception; such information is available in a form that can be reorganized, if need be (this is particularly obvious with ambiguous figures). A mental image, in contrast, is generated on the basis of information already in memory, which has already been organized and interpreted. According to Reisberg and Chambers (1991) and Chambers and Reisberg (1992), in visual mental images only information explicitly included can be retrieved. The included information represents the interpretation of the form. These authors argued that since there is no interpretative process, people cannot re-interpret patterns in images.

These claims led to a series of studies (e.g., Anderson and Helstrup, 1993; Finke et al., 1989; Peterson et al., 1992) that provided evidence that ‘new’ (i.e., not previously described and stored) patterns can in fact be interpreted in visual mental images. These researchers showed that if the subjects understand the point of the task and are provided with strategies that can overcome the limited capacities of short-term memory, they can even reorganize ambiguous figures in images (even if the reinterpretation requires re-assigning the specification of front/back or top/bottom). Such findings led Peterson et al. (1992) to infer that perception and imagery share relatively low-level processes, which operate before perceptual interpretation is complete.

The present experiments were designed to investigate the issue of whether mental images preserve relatively unprocessed perceptual information, or represent a few descriptions of the structure of an object or scene, and thus contain only the

information closely related to the initial interpretation. We begin by drawing a distinction between ‘low-level’ and ‘high-level’ visual properties. Low-level visual properties are the basis on which subsequent perceptual organization takes place. High-level properties arise from specific ways of organizing low-level properties. For example, different types of intersections of edges (e.g., T, Y and L junctions) are low-level properties, whereas closure and symmetry of the entire shape are high-level properties. Since these last two properties are not embedded in a subsequent organization, we expect that they are easier to detect in a picture.

Note that the terms ‘high-level’ and ‘low-level’ are used in this article to refer to the expected ease with which a property can be detected in a picture, depending on the level of organization it influences. The terms do not refer specifically to the level of processing of the property (where ‘low’ stands for ‘early’). For example, we use symmetry as a high-level property, because it is a specification of the whole picture—not an isolated part. It has been argued that symmetry is processed preattentively (Baylis and Driver, 1995). This is not in conflict with our conception because what we call a high-level property might or might not be processed preattentively.

In the present study we sought to determine whether both types of properties are as accessible in a mental image as they are in a percept. If Peterson et al. (1992) are correct, images preserve relatively unprocessed perceptual information and even ‘low-level’ properties can be retrieved from a visual mental image. If Reisberg and Chambers (1991) are correct, a mental image is restricted to limited information, based on the interpretation of the whole picture. If so, then we have no reason to expect people to be able to detect the presence of low-level features—which should be buried in the descriptions of the high-level features they produce. And if they can, detecting the low-level features in an image should be disproportionately more difficult than detecting high-level features. Although the logic of this study is straightforward, the actual mechanics of implementing it had to take into account the fact that images are transient and are constrained by the capacity limits of short-term memory. First, we assume that an image decays over time, and, hence, holding a mental image demands constant regeneration. This could conceivably take place part-by-part or by replacing the entire image at once. In either case, the availability of information in an imaged pattern is in flux, and is not as constant as occurs when one is looking at a picture. Second, regenerating an image requires effort. When faced with an overload of information, one must choose what to include and what to exclude, and subjects may be biased to omit certain features that should in fact be included when performing a particular task.

We attempted to minimize the influences of limited short-term memory capacity and image regeneration biases by using simple line-drawings as stimuli and by only querying properties that should be essential in a picture (these can be high-level as well as low-level); such properties would not be ‘optionally omitted’ to cope with capacity limits. Specifically, we choose to examine so-called *non-accidental properties* (see Biederman, 1987; Lowe, 1985, 1987). A non-accidental property has a low likelihood of having arisen by chance. For example, when

perceiving segments that line up as parallel lines, the chance that these segments are parallel due to chance is very small (Lowe, 1985, 1987; Kanade and Kender, 1983). Because non-accidental properties are robust over various changes in orientation and viewpoint, it is not surprising that there is good evidence that they play an important role in human object recognition (Biederman, 1987; Cave and Kosslyn, 1993), and can be considered essential properties of a picture.

We chose six non-accidental properties by three criteria: (1) there had to be a consensus among researchers that the property met the essential criteria for being considered nonaccidental; (2) empirical evidence supported the claim that it is in fact a non-accidental property; and, (3) the property is clearly definable in such a way that it is suitable for a yes/no question about a picture. We chose two high-level properties, which characterized the entire picture, that were used by Biederman (1987) and Lowe (1985): ‘symmetry’ (the whole picture is exactly symmetrical) and ‘parallelism’ (two or more lines are parallel). In addition, we selected two types of low-level properties. The first was ‘continuation,’ which occurs when a line continues after an interruption. The second was ‘cotermination,’ which occurs when several lines terminate at the same point (e.g., see Biederman, 1987 and Lowe, 1985). We divided cotermination into three special cases; these properties were: ‘T-junction,’ ‘Y-junction,’ or ‘Arrow-junction,’ which refer to three lines meeting such as to form a T, Y and an arrow, respectively. Thus, we had four low-level properties in all. The present issue hinges on the perception of low-level properties, and, hence, it makes sense to examine relatively more of them. Thus, we examined the relative ease of detecting the two classes of properties in images and percepts.

## EXPERIMENT 1

In Experiment 1, a visual perception condition was compared to a visual imagery condition. In the perception condition, a picture appeared and remained on the screen. The name of one of the six properties was presented auditorially, and subjects were asked to decide whether the picture did or did not contain this property. In the imagery condition, a picture appeared on the screen and then disappeared. Subjects were instructed to keep their eyes fixated on the screen, and to hold a mental image of the picture. After a short interval the name of a property was presented auditorially. Subjects were asked to inspect their mental image (‘to look with the mind’s eye’) to decide whether it contained the named property.

## 2. Method

### 2.1. Subjects

Sixteen undergraduate students at Harvard University (11 females and 5 males) volunteered to serve as paid subjects. None of these subjects were aware of the purposes or predictions prior to testing, and none participated in Experiment 2.

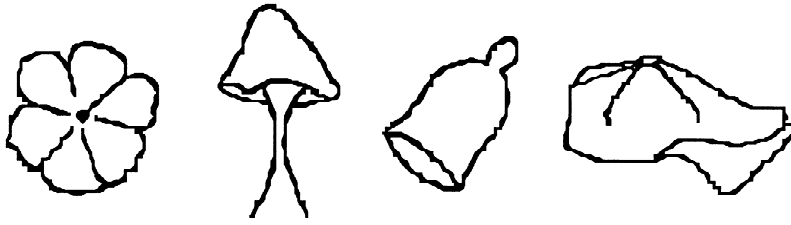


Fig. 1. Some of the line drawings used as stimuli in the experiment.

## 2.2. Materials

Eighteen simple line drawings of common, emotionally neutral, objects were prepared, as illustrated in Fig. 1. The line drawings were created so that 2–6 of the non-accidental properties would be clearly present or absent. The objects were divided into two sets of 9 pictures each (Appendix A). Two sets of items allowed each subject to perform both conditions with different stimuli. In addition, including more pictures reduced the probability that some properties are more difficult to perceive only because they were inadvertently paired with more difficult or more complex pictures.

We also selected a set of properties to be queried. All of these properties were classified as non-accidental properties by several theorists; as noted earlier, we found such a consensus, and at least some empirical evidence, for the following non-accidental properties: parallelism, symmetry, cotermination (lines terminate in a common point) and continuation. We formulated a simple description of each property, which allowed us to create yes/no questions “does this picture contain the property ‘-----’?”. Parallelism was described as: ‘Two or more lines that are separated by an equal distance at every point, which also means they have the same direction at each point’. Symmetry was: ‘The presence or absence of symmetry in the whole picture. This means that one half of the picture is the mirror image of the other half’. Continuation was described as: ‘an interrupted line continuing on the other side of the interruption at the appropriate place, continuing the original trajectory’. We specified three kinds of cotermination, or junctions: T-junction, Y-junction and arrow junction. These junctions have often been used in computer vision systems (e.g., see Winston, 1975). The description of a T-junction was: ‘Two lines in the picture intersect such that they form a T... the lines must be perpendicular’. The description for Y-junction was: ‘A ‘Y’ consists of three line segments. Two of them are each on one side of the third, and extend from this middle line to form a Y’. We emphasized that this meant that a vertical line meets two other lines, both extending in an angle of more than 90° on either side. If a vertical line meets two lines, one on its left and one on its right side, and both are in an angle of less than 90°, an arrow-junction is formed. The description of arrow junction was: ‘An arrow consists of three line segments. Two of them are each on one side of the third, slanting in the same direction as the center line to form an arrow’.

We divided the non-accidental properties into two categories, low-level and high-level. We assume that junctions are used as cues for high-level parts but are not themselves explicit components of a percept; thus, we treated junctions as low-level properties. We also assumed that continuation is a low-level property, but for a different reason: even though seeing ‘good continuation’ is associated with a more holistic interpretation, in order to be aware that a line cuts across perceptual units, one may have to break up higher level perceptual groups (and the same is true of junctions). In contrast, the properties of symmetry and parallelism are closely related to Gestalt laws of grouping, which influence how elements are organized into wholes; we assume that the presence or absence of the properties symmetry and parallelism influence the organization of the whole picture. Thus, we considered these to be high-level properties.

A trial consisted of a combination of a property and a picture. The correct answer could be either ‘yes’ or ‘no’ (the picture does or does not contain this property). The total number of trials was 36; each of the six properties was queried 6 times, three times with ‘yes’ and three times with ‘no’ as the correct response. A Macintosh LC II was used to present the stimuli and record the answers and response times. The pictures subtended a mean visual angle of 7° horizontally and 8.5° vertically, from the subject’s perspective, on the computer’s screen.

### 2.3. Procedure

Subjects began by learning the description of each property, as follows. The name of a property was presented on the computer screen, as well as auditorially (by the computer), followed by a simplified written description of that property. The properties were described with the aid of examples, which illustrated what was and what was not considered an instance of that property. The examples were always presented outside the context of a picture; for instance, examples of junctions were provided by three lines that met in different ways. The subjects were allowed to read the descriptions and study the examples as long as they liked. A second and more detailed description followed, which not only provided the descriptions specified in the materials section but also provided additional examples; moreover, comments on the computer screen also made clear that the line(s) containing the property could be curved, and that they could appear in any size and orientation, and in any location in the picture. For symmetry this meant that a symmetrical figure could contain curved lines, and that it did not matter how the line of symmetry was oriented. The explanation of each property ended with a short definition. After the subject finished the set of descriptive materials for a given property, another was presented, and so on until all 6 were studied. The total training session took approximately 10 min.

After the descriptions of the properties, instructions appeared on the computer screen, followed by practice trials. A practice trial began with a fixation cross, displayed at the center of the screen for 0.75 s. After the cross had disappeared, a picture appeared at the same location. In each subsequent practice trial the same picture was used (a leaf in the imagery condition and a salt shaker in the

perception condition). In a practice trial in the perception condition, after the picture had been on the screen for 1.3 s the name of a property was read aloud by the computer. The presentation of a name always took exactly 1 s. The subjects were asked to indicate as quickly as possible whether the picture contained the named property. The picture remained on the screen until a response was made.

In a practice trial in the imagery condition, after 1.3 s the picture disappeared, and the screen was blank for 2 s. The name of a property was then presented. The subjects were instructed to maintain fixation at what was the picture's location on the screen, and hold a mental image of the picture. They were asked to decide whether the picture contained the property by 'looking with the mind's eye' at the picture in their image. The screen remained blank until the subject responded. In both conditions, a beep sounded after an incorrect answer, and after each practice trial subjects were asked whether they had any questions.

There were six practice trials, with each property being probed once. For both the leaf and the salt shaker stimuli, half the correct responses were 'yes' and half were 'no'. After the six practice trials, the practice picture appeared on the screen, and all properties and their correct answers were repeated. Subjects were encouraged to ask questions.

After all questions had been answered, the test session began. The test trials were identical to the practice trials, with the following exceptions: the pictures were from object set 1 or 2, there was no feedback following the response, and there was no time to ask questions; immediately after the subject responded, the fixation cross of a new trial appeared. A subject always responded with the right hand, pressing either a 'yes' or a 'no' labelled key on the keyboard. After 18 trials the subjects took a short break.

Each subject performed both the perception and the imagery condition, in a counterbalanced order. Each set of objects was used equally often in each condition, as were two trial orders ('forwards' or 'backwards'). The trials were ordered quasi-randomly, with the following constraints: no picture or property would be presented in succession, and all pictures and properties were spread roughly evenly throughout the experiment, not clustering together.

### 3. Results

We performed separate analyses of variance (ANOVAs) for the response times and error rates. All effects and interactions that are not mentioned were not significant,  $p > 0.1$  in all cases.

#### 3.1. Errors

To examine whether subjects were able to extract the low-level properties from an image, we looked at the error rates of these items separately. If subjects were simply guessing, a mean of 50% errors would be expected; however, none of the

subjects made more than 8 (33.4%) errors on the ‘low-level property’ items in the imagery condition.

Further, error rates did not differ between the two conditions, ranging from 1–8 (2.8–22.2%) errors in the perception condition, and from 2–9 (5.6–25%) errors in the imagery condition,  $F < 1$ . Taking the two conditions together, we found that the number of errors differed per property,  $F(5,75) = 4.00$ ,  $p = 0.0028$ . Fewer errors were made for high-level properties than for low-level properties,  $F(1,15) = 28.05$ ,  $p = 0.0001$ , with means, respectively, of 1.69 (of 24 trials is 7%), and 7.88 (of 48 trials is 16.4%). This effect was comparable for both conditions,  $F < 1$  for the interaction of condition and level. Finally, the subjects made more errors on trials with ‘yes’ as the correct response than on trials with ‘no’ as the correct response,  $F(1,15) = 23.31$ ,  $p = 0.0002$ .

### 3.2. Response times

Means and standard deviations of the response times (in ms) for each response category (yes/no) in each task (perception/imagery) were calculated for each property for each subject. Latencies greater than three times the standard deviation of the mean for a given cell for a given person were treated as outliers and excluded prior to taking the mean of that cell; this procedure resulted in our discarding 0.8% of the data. In addition, only times from correct responses were analyzed.

As illustrated in Fig. 2, there were large differences between the mean response times for the different properties,  $F(5,75) = 10.73$ ,  $p < 0.0001$ . The pattern of variation of response times corresponded to the pattern of errors we observed. When response times were collapsed over the two high-level properties (parallelism and symmetry) and the four low-level properties (continuation, T-junction, Y-junction and arrow-junction), we found that subjects took longer to evaluate the low-level properties than the high-level properties (with means of 2.766 s vs. 2.281

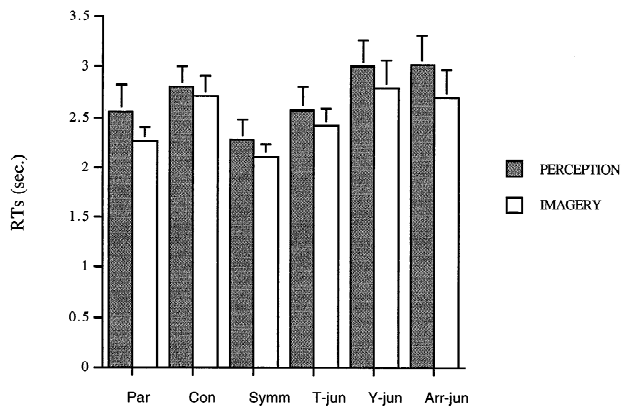


Fig. 2. Mean response times (with standard error) per property in the two conditions of Experiment 1.



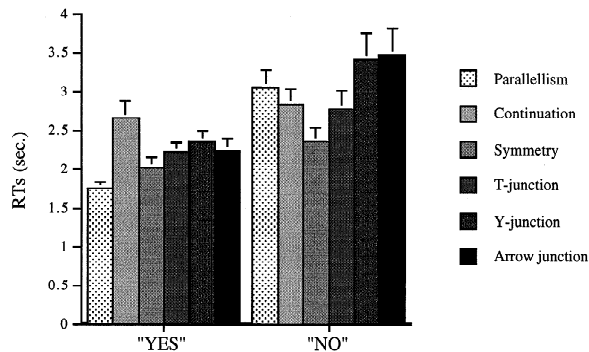


Fig. 3. Mean response times (with standard error) of the six properties per response category in Experiment 1.

s),  $F(1,15) = 27.42$ ,  $p < 0.0001$ . This difference was present in both the imagery and perception conditions; indeed, we found no hint of an interaction between property-level and condition,  $F < 1$ . As evident from Fig. 2, examination of the mean response times of the six properties separately revealed no difference between the two conditions in the relative speed of response per property,  $F < 1$ .

To examine further the similar pattern of response times for the different properties in imagery and perception, we correlated the times for the properties in the two conditions. The order of the properties in both conditions was very highly correlated: the Spearman's correlation coefficient was  $r_s = 0.83$ .

In addition, correct 'yes' responses were faster than correct 'no' responses (with means of 2.211 s for 'yes' and 2.990 s for 'no' answers),  $F(1,15) = 16.23$ ,  $p < 0.0011$ . As illustrated in Fig. 3, the patterns of mean response times for the six properties were different for the two response categories ('yes' and 'no'),  $F(5,75) = 5.900$ ,  $p = 0.0001$ . However, the pattern of mean response times for the properties per response category did not differ between the imagery and perception condition,  $F < 1$ .

#### 4. Discussion

We found that subjects were able to retrieve the low-level properties from a mental image. Further, in both the imagery and the perception condition, the high-level properties were easier to retrieve than the low-level properties. This is not surprising, as the high-level properties are more explicitly present, while low-level properties are buried in the higher level properties. But what is of most interest is that the relative ease of evaluating the different types of properties was the same in both conditions: In particular, consistent with the claims of Peterson et al. (1992), low-level properties were as accessible in images as in percepts. The finding of a very similar pattern of results in the two conditions supports the idea

that images preserve even low-level visual features. Through inspection of the image, these formerly not explicitly known properties can be retrieved.

In this experiment, the delay between the disappearance of the picture and hearing the name of the property was 2 s. After the picture disappeared, the fast decay of information over time can be countered by actively ‘regenerating’ the image. However, when regenerating the image, one must select which visual information to retain. Although 2 s is enough time to require one to regenerate the image of the picture, an alternative explanation of the results of Experiment 1 could be that more information than usual is included in the image. Placing more demands on the processes of imagery, by putting a long delay between seeing the picture and generating the mental image of the picture, will require subjects to select only some of the visual information. Then, if generating an image only activates information explicitly included, a long delay after seeing the picture will unavoidably lead to the inability to retrieve low-level properties from visual mental images.

In addition, it is possible that there is a qualitative difference between images that are formed by retaining perceptual input and those that are formed by activating information in long-term memory. There is now good evidence that the former kinds of images rely on area 17 in the brain (e.g., Le Bihan et al., 1993; Ishai and Sagi, 1995), whereas the evidence is less firm that the second kind rely on this structure (e.g., see Gulyas and Roland, 1994). Thus, we decided to examine whether similar results can be found as in the imagery condition of Experiment 1, if objects are visualized from long-term memory.

## EXPERIMENT 2

Experiment 2 was conducted to discover whether the results of the imagery condition of Experiment 1 were a consequence of the short delay between the disappearance of a picture and extracting a property of that picture. If so, then we would expect different results with a long delay. By asking subjects to visualize previously memorized objects, in this experiment we can also discover whether images that are formed by retaining perceptual input convey properties differently than images formed by activating information in long-term memory. Thus, in the present experiment we again study the relative ease with which the six non-accidental properties can be extracted from a mental image. However, this time subjects generated images of pictures from long-term memory.

We added a control task in this experiment in order to discover whether errors were a consequence of the subject’s not fully understanding the task or of their having difficulty seeing a property in the (image of the) picture. We expected an increase in the number of errors in the imagery condition because representations of the pictures had to be retrieved from long-term memory, and hence, we wanted to ensure that such an increase did not mean that the subjects failed to understand the task. In the control task, the subjects were asked whether properties are present in pictures that were printed on a piece of paper. The subjects were given as much time as they needed to perform this task; hence, subjects should have been very

accurate on this task if they understood the definitions of the properties. Thus, we could use these results to define a criterion for including only those subjects who understood the task in our analysis.

## 5. Method

### 5.1. Subjects

Twenty-one undergraduates (8 males 13 females) at Harvard University volunteered to participate as paid subjects. None of these subjects were aware of the purposes or predictions of the experiment until post-testing debriefing.

### 5.2. Materials

The only difference in materials from Experiment 1 was that each object in the two sets was also named. The names of the objects in set 1 were bell, cap, umbrella, fence, mug, football, comb, nut, pants; the names for set 2 were flower, mushroom, balloon, ironing board, baseball bat, ear, table, cheese, snail shell (see Appendix A for objects). These names were recorded, and during the experiment they were presented aloud by the computer.

### 5.3. Procedure

Each subject began by learning to form mental images of the pictures in one of the sets, and of a drawing of a leaf, which was used in the practice trials. Subjects learned to visualize the pictures as follows. First, the pictures were shown, one at a time, while their names were presented auditorially. Subjects were asked to look carefully at each of the pictures and were told that they later would be asked to form a mental image of these pictures. All pictures were presented a second time, but this time the name of the picture was presented first, and only after the subject had pressed the space bar to indicate that a mental image of the picture was formed did the picture appear on the screen. Subjects were asked to compare their mental images to the pictures, and to correct the images if necessary, and then were to press the space bar so that the next picture would appear. If a subject reported having ‘no or almost no errors in the mental images’ after the third time all pictures were presented, the training session was considered complete. Otherwise, the learning procedure was repeated until the subject claimed to make no errors in a complete sequence. Subjects then learned the definitions of the six properties as in Experiment 1.

The experiment started with instructions and six practice trials. The trials differed from those of Experiment 1 in that subjects did not see the pictures, but rather heard the name of a picture after the fixation cross had disappeared. Subjects were instructed to form a mental image of this picture, and to press the space bar as soon as they had generated the image. Then, as in Experiment 1, the name of

the property was presented auditorially, and subjects indicated whether the imaged picture included the named property. In the practice trials, the picture of the leaf was always used. After these trials, the actual test trials followed. The two sets of drawings were used equally often over subjects. In all other respects, the procedure in Experiment 2 was identical to that in Experiment 1.

Finally, after completing the response time procedure, the subjects were given a paper-and-pencil task. They were given a piece of paper on which the pictures used in the experiment were displayed and again were asked all 36 questions (whether each picture contained each property or not, see Appendix B). Before asking the questions, the experimenter explained to the subject that the aim of this task was to assess accuracy, and that producing a quick response was not relevant at this point.

## 6. Results

The data were analyzed as in Experiment 1. All effects and interactions not mentioned were not significant,  $p > 0.1$  in all cases. Subjects who committed more than 33% errors in the experiment, and/or more than 4 errors in the paper-and-pencil test were excluded ( $n = 5$ ). Sixteen subjects were included in the analysis.

### 6.1. Errors

Given that subjects were asked to choose between a ‘yes’ and a ‘no’ response, a percentage of 50% errors would indicate guessing. We looked at the percentage errors for the ‘low-level property’ items separately, to determine whether subjects were guessing on these items. We found that none of the subjects got half or over half of the items wrong. In fact, of the 16 subjects, 14 had not more than 8 (33.4%) errors on these items, indicating that subjects could indeed retrieve the low-level properties from their mental images.

Collapsing over all trials, subjects made 4–11 errors (11–31%), with a mean of 8.3 (23%). Error rates differed between properties,  $F(5,75) = 3.40$ ,  $p = 0.008$ . Fewer errors were made on the high-level properties than on the low-level ones,  $F(1,15) = 9.832$ ,  $p = 0.007$ .

### 6.2. Response times

Of all response times, 0.35% were excluded as outliers, using the same procedure used in Experiment 1.

As illustrated in Fig. 4, we again found large differences between the mean response times for the different properties,  $F(5,75) = 8.363$ ,  $p = 0.0001$ . Response times were collapsed over the two high-level properties (symmetry and parallelism) and the four low-level properties (continuation, T-junction, Y-junction and arrow-junction). As in Experiment 1, and corresponding to the pattern of errors, subjects required more time to evaluate low-level properties than high-level

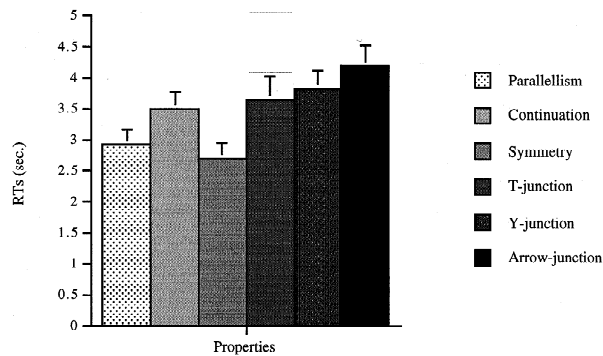


Fig. 4. The mean response times per property in Experiment 2.

properties in the drawings (with mean response times of 3.792 s vs. 2.811 s),  $F(1,15) = 25.82$ ,  $p < 0.0001$ . It is interesting to note that the variation of response times for the different properties was very similar to that of Experiment 1, as can be seen in Fig. 4. And, in fact, Spearman's correlation coefficients indicated that the rank of the properties in this experiment was similar to that of Experiment 1: The correlation with the perception condition of Experiment 1 was  $r_s = 0.94$ , and the correlation with the imagery condition of Experiment 1 was  $r_s = 0.71$ .

### 6.3. Paper and pencil test

Subjects made 0–4 (11%) errors on the paper and pencil test, with a total of 30 (5% of all trials over all subjects), and a mean of 1.88 (5%) errors. Although too few errors were made to draw conclusions about these differences, it is interesting to note that a difference existed between the number of errors on high- and low-level properties: for the two high-level properties, the mean error per property was 2.5 (5 errors in total), whereas for the 4 low-level properties, the mean error per property was 6.25 (25 errors in total).

As noted before, certain items in the experiment were evaluated incorrectly by a majority of the subjects. The paper and pencil test was used to assess whether these errors occurred because of some particularly difficult item or a misinterpreted definition of a property; we assumed that such misidentifications would also lead to the wrong answer when looking at the pictures in the paper and pencil task. If the item is simply harder to inspect under time pressure, because the presence or absence of the property needs a thorough inspection process, we would expect the majority of the subjects to provide the right answer to that item on the paper and pencil test.

In order to determine the number of errors made for individual items, the two object groups were considered separately. Of the 16 subjects, half saw object set 1 and the other half object set 2. Only one item was evaluated incorrectly by more than 4 subjects: 5 subjects failed to give the correct answer ('yes') to the question of whether an arrow junction was present in the drawing of a mushroom. This

arrow-junction, however, was particularly small in size in the picture, so it is possible that it was overlooked. Four subjects did not provide the correct ‘yes’ answer when asked whether the drawing of a flower contained the property of continuation. Only these two items (of the total of 72, since there were 36 per object set) were evaluated incorrectly in the paper and pencil task. Thus, it was clear that subjects understood the task and the definitions of the properties.

## 7. Discussion

Although in this experiment the subjects were forced to use a different method to generate visual mental images, we still found that the subjects could retrieve the low-level properties from their images. This finding supports our hypothesis that visual mental images preserve relatively unprocessed information.

The fact that the same results were found in Experiment 2, even though subjects were required to generate an image from a picture they saw some time ago, supports the findings in Experiment 1. It is important to note here that since subjects saw the pictures only for a relatively short time, and during presentation of the pictures were not aware of the purpose of the study, it is impossible that they memorized the presence or absence of the low-level properties by explicitly paying attention to them. Thus, the fact that they were able to answer the low-level property items in this experiment indicates that they retrieved this information by inspecting their mental images. In addition, the similar pattern of results found in the perception and the imagery conditions in Experiment 1 suggests that the images created in Experiment 2 were qualitatively similar to those retained in Experiment 1.

## GENERAL DISCUSSION

In this article, two important findings are reported. First, we found that people are able to extract low-level properties from mental images. Second, we found a very similar pattern of performance for the different properties in imagery and perception conditions. We did find the expected difference between low-level and high-level properties, with the high-level properties being easier to extract, but this difference was found to the same degree in both the imagery and the perception conditions. What is particularly striking is that subjects were still able to extract low-level properties from a mental image in Experiment 2. Here, many steps (learning the picture, storing it in memory, generating an image) intervened between the perception of the picture and the evaluation of the property in a mental image. These results are consistent with those of Peterson et al. (1992), and contradict the claims of Reisberg and Chambers (1991) and Chambers and Reisberg (1992). A visual mental image is not limited to a representation of a few specifications of the whole object.

Even though our subjects could in fact extract low-level non-accidental

properties from visual images, we did find better performance in both imagery and perception when subjects evaluated the high-level properties than when they evaluated the low-level properties. This was expected if the low-level properties are embedded in higher-order units. If so, then the higher-order units were immediately visible, whereas the lower-level ones required perceptually breaking down these units. However, this remains a topic for further investigation.

The findings reported here are in accordance with other findings of functional equivalence, such as those reported by Podgorny and Shephard (1978); Bagnara et al. (1988), and many others; for reviews, see Finke (1979); Kosslyn (1994); Shepard and Cooper (1982). These findings of functional equivalence support the claim that imagery and perception rely on shared underlying mechanisms. Thus, the claim that imagery is an integral part of the process of perception (e.g., Kosslyn, 1994) is supported by the results. It is worth noting that in order for the process of imagery to be of use in visual perception (e.g., as an aid to top-down hypothesis testing), the two processes must be compatible. Thus, the way information is represented and processed in imagery and in perception should be the same. The parallel findings in imagery and perception, when ordering the properties according to difficulty of extracting them, supports the assumption of similar representations and processes.

In closing, we wish to note that although the functional equivalence we observed does support the view that imagery and perception rely on at least some of the same mechanisms, this theory does not predict that imagery and perception tasks will always show similar results. For both a visual percept and a visual mental image, we can distinguish between the ‘content’ and the ‘format’ of the representation: *what* information is represented, and *how* is it represented. We assume that the content of an image can differ from the content of a percept for a number of reasons: the two representations arise from different sources (what is stored in memory versus what is depicted on the retina); an image is more flexible (perception is dependent of reality, imagery is not); an image takes more effort to maintain (an image needs constant regeneration, which may lead it to become degraded over time); and an image has more severe capacity limitations (only a limited amount of information can be held in an image, e.g., one probably does not portray all the stripes of the zebra). Nevertheless, we have good evidence that images are not, in principle, restricted to represent only a description of an object’s parts and their arrangement; images can also represent relatively low-level visual features, even when they are embedded in higher level ones.

## Acknowledgments

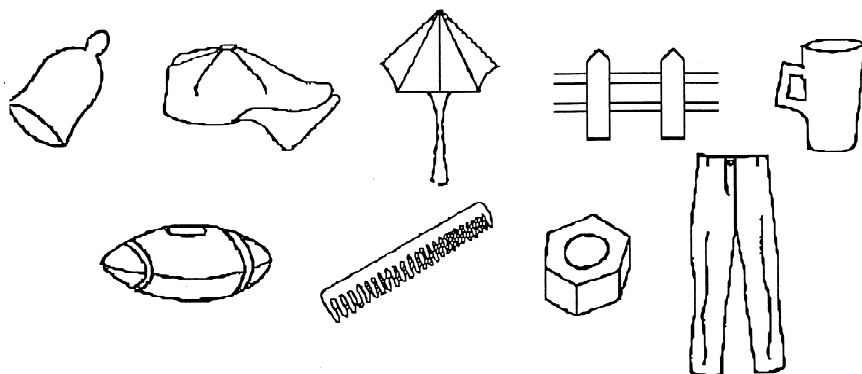
The authors would like to thank Volkert Balk for computer-software assistance. The study was supported by funding provided by the following institutions: Het Algemeen Studiefonds, Het Bekker-LaBastide Fonds, De Schuurman Schimmel van Outeren Stichting, De Vrijvrouwe van Renswoude te Delft, Het Profilerings-

convenant van de Universiteit van Amsterdam, and the US National Institute of Aging, Grant AG12675-01.

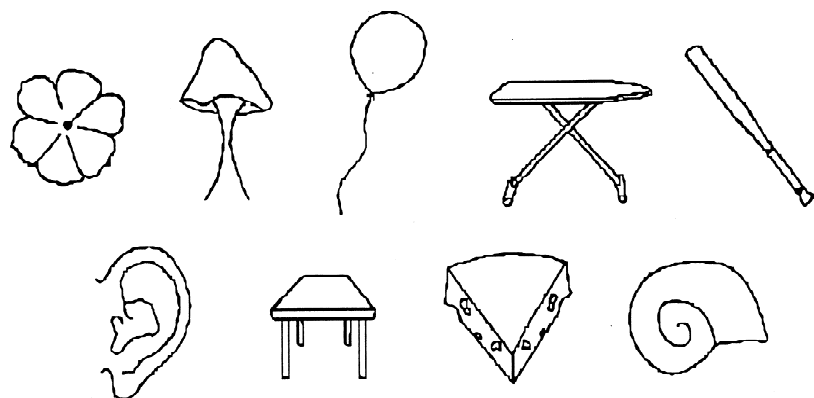
## Appendix A

### The two object groups of nine pictures

#### OBJECT GROUP 1



#### OBJECT GROUP 2





## Appendix B

### The 36 questions (trials) with objectgroup 1, and with objectgroup 2

Objectgroup 1						
	Parallelism	Continuation	Symmetry	T,Y, Arrow-junction		
	P	C	S	T	Y	A
Bell	n	n	y		n	
Cap	n		n			
Umbrella	n	n	y	y		y
Fence	y	y	y	y		n
Mug		y	n	n	n	
Football	y	y	n		y	y
Comb				n	n	n
Nut	y	n		n	y	y
Pants				y	y	n

Objectgroup 2						
	Parallelism	Continuation	Symmetry	T,Y, Arrow-junction		
	P	C	S	T	Y	A
Flower	n	y		n	y	
Mushroom	n	y	n			y
Balloon	n				y	y
Iron.Board	y	y		y		
B.bat	y	n	y	y	n	n
Ear			n	y	n	n
Table	y	n	y		n	
Cheese			y	n		y
Snail		n	n	n	y	n

## References

- Anderson, R.E., and T. Helstrup. (1993). Visual discovery in mind and on paper. *Memory and Cognition*, 21, 283–293.
- Bagnara, S., Simion, F., Tagliabue, M.E. and Umiltà, C. (1988). Comparison processes on visual mental images. *Memory and Cognition*, 16(2), 138–46.

- Baylis, G.C. and Driver, J. (1995). Obligatory edge assignment in vision: The role of figure and part segmentation in symmetry detection. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1323–1342.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115–147.
- Cave, C.B. and Kosslyn, S.M. (1993). The role of parts and spatial relations in object identification. *Perception*, 22, 229–248.
- Chambers, D. and Reisberg, D. (1992). What an image depicts depends on what an image means. *Cognitive Psychology*, 24, 145–174.
- Farah, M.J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review*, 95, 307–317.
- Finke, R.A. (1979). The functional equivalence of mental images and errors of movement. *Cognitive Psychology*, 11, 235–264.
- Finke, R.A. (1989). *Principles of mental imagery*. Cambridge, MA: MIT Press.
- Finke, R.A., Pinker, S. and Farah, M. (1989). Reinterpreting visual patterns in visual imagery. *Cognitive Science*, 13, 51–78.
- Ishai, A. and Sagi, D. (1995). Common mechanisms of visual imagery and perception. *Science*, 268(5218), 1772–1774.
- Kanade, T. and Kender, J.R. (1983). Mapping image properties into shape constraints: Skewed symmetry, affine-transformable patterns and the shape-from-texture paradigm. In J. Beck, B. Hope and A. Rosenfeld, (Eds.), *Human and machine vision*, 1, p. 237–257. New York: Academic Press.
- Kosslyn, S.M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge (MA): MIT Press.
- Kosslyn, S.M. and Koenig, O. (1992). *Wet mind: The new cognitive neuroscience*. New York: The Free Press.
- Le Bihan, D., Turner, R., Zeffiro, T.A., Cunod, C.A., Jezzard, P. and Bonnerot, V. (1993). Activation of human primary visual cortex during visual recall: A magnetic resonance imaging study. *Proceedings of the National Academy of Sciences of the United States of America*, 5(24), 11802–11805.
- Lowe, D.G. (1985). *Perceptual organization and visual recognition*. Boston: Kluwer.
- Lowe, D.G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, 31, 355–395.
- Neisser, U. (1967). *Cognition psychology*. New York: Appleton-Century-Crofts.
- Peterson, M.A., Kihlstrom, J.F., Rose, P.M. and Glisky, M.L. (1992). Mental images can be ambiguous: Reconstruals and reference-frame reversals. *Memory and Cognition*, 20(2), 107–123.
- Podgorny, P. and Shepard, R.N. (1978). Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 21–35.
- Reisberg, D. and Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology*, 45(3), 336–352.
- Roland, P.E. and Gulyas, B. (1994). Visual imagery and visual representation. *Trends in Neurosciences*, 17(7), 281–287.
- Shepard, R.N. and Cooper, L.A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- Winston, P.H. (1975). *The psychology of computer vision*. (Vol. Ed.) (Sec. Ed.). New York: McGraw Hill.