

# Inspecting visual mental images: Can people “see” implicit properties as easily in imagery and perception?

WILLIAM L. THOMPSON

Harvard University, Cambridge, Massachusetts

STEPHEN M. KOSSLYN

Harvard University, Cambridge, Massachusetts  
and Massachusetts General Hospital, Boston, Massachusetts

MICHAEL S. HOFFMAN

University of Pittsburgh, Pittsburgh, Pennsylvania

AND

KATINKA VAN DER KOOLJ

Utrecht University, Utrecht, The Netherlands

---

Can people “see” previously unnoticed properties in objects that they visualize, or are they locked into the organization of the pattern that was encoded during perception? To answer this question, we first asked a group to describe letters of the alphabet and found that some properties (such as the presence of a diagonal line) are often mentioned, whereas others (such as symmetry) are rarely if ever mentioned. Then we showed not only that other participants could correctly detect both kinds of properties in visualized letters, but also that the relative differences in the ease of detecting these two types of properties are highly similar in perception (when the letters are actually visible) and imagery (when the letters are merely visualized). These findings provide support for the view that images can be reinterpreted in ways much like what occurs during perception and speak to the wider issue of the long-standing debate about the format of mental images.

---

To what extent are visual mental images *images*? This question has been debated, sometimes with considerable heat, for centuries (for a review, see Kosslyn, Thompson, & Ganis, 2006). One property of an image in perception is that it can be examined more or less carefully, and additional information can be derived over time. For example, if shown the uppercase letter *A*, people can readily verify that a horizontal line is present. Moreover, they also can verify that there is an enclosed shape and that it is a triangle. They may not have noticed the triangle until asked about it but can easily “look more carefully” at the shape and classify that part appropriately. If objects in visual mental images are like objects we perceive in this regard, people should be able to identify properties in mental images that initially were not noticed.

If mental images allow people to identify properties of objects that were not immediately noticed, such images presumably preserve information that can be organized in multiple ways. The notion that mental images preserve material that is relatively “raw,” that can be organized in numerous ways, has proven controversial (e.g., Peterson, Kihlstrom, Rose, & Glisky, 1992; Reisberg, 1997; Reis-

berg & Chambers, 1991; Rouw, Kosslyn, & Hamel, 1997). One way to address this issue is to compare the ease of accessing two types of properties in perception and imagery: those that are likely, “at first glance,” to be represented explicitly in memory and those that, like the triangle in the *A*, are likely to be only implicit and to be accessible only after more careful scrutiny. That is, we conjecture that some visual features are encoded in memory explicitly by default, whereas others are represented only implicitly and must be inferred or derived from the explicitly encoded information.

The use of the term *implicit* here should not be confused with its use by memory researchers; in our usage, both *explicit* and *implicit* properties are accessible to conscious awareness. However, an explicit visual property is noticed immediately, stored in memory as such, and subsequently can be accessed simply by looking it up in memory; in contrast, an implicit property must be derived by reorganizing what has been encoded explicitly. Even in perception, an internal representation is created as soon as a shape is registered, and this representation makes only some information explicit and immediately accessible (cf. Marr, 1982).

---

S. M. Kosslyn, smkosslyn@wjh.harvard.edu

---

In the experiments reported here, we compared the ease of classifying explicit versus implicit properties in perception and in imagery. If the same representations and processes are at work in both cases, we would expect very similar results in the two conditions. Specifically, although it may be more difficult (as reflected by longer response times [RTs] and increased numbers of errors) to classify implicit properties than explicit ones, this difference should be the same in perception and imagery. We expect differences not only in the ease of accessing specific visual properties, but also in the ease of doing so for different patterns; more visually complex patterns should be more difficult to inspect. But again, if the same representations and processes are at work, we would expect comparable differences in imagery and perception.

The issue of whether mental images can be reexamined in order to extract properties that were previously unnoticed bears directly on the so-called *imagery debate* (e.g., Kosslyn, 1980, 1994; Kosslyn & Pomerantz, 1977; Kosslyn et al., 2006; Pylyshyn, 1973, 1981, 2003). This debate focuses on the format of mental images. The *format* refers to the type of code, which is distinct from the content; the same content (e.g., the information conveyed in this sentence) can be represented in many formats (e.g., in different languages, written, spoken, in Morse code, as sets of bits in a computer memory, etc.). The imagery debate is not about whether people store visual information, or even whether they can recall visual information and use it to perform tasks. These are questions of content. Rather, the imagery debate focuses on the format of the representation, how the information is stored.

Two major theories of the format of mental images have been developed in detail. On the one hand, those who espouse the propositional (also known as the descriptive or symbolic) account of mental imagery (e.g., Anderson & Bower, 1973; Pylyshyn, 2003) argue that mental images are represented symbolically, in a language-like system. In fact, such theorists typically argue that the same symbolic system is used for all internal representation, including in perception and language. On the other hand, those who espouse the depictivist account (e.g., Kosslyn et al., 2006) argue that points in a representational space correspond to points on objects in the world. In this view, every part of the representation corresponds to part of the object, so that the distances among the represented parts mirror the actual distances among the parts themselves. According to this view, a mental image is in fact an *image*; it is not a description. If so, much “raw material” should be present when one visualizes a pattern, and it should be possible to “look again” at an image, extracting shapes or spatial relations from it—even if those shapes or relations had not been considered explicitly in the past. The present experiments put this idea to the test.

## EXPERIMENT 1

We began by asking whether there are grounds for drawing a distinction between explicit and implicit visual properties of letters. To assess the presence of explicit properties, we simply asked participants to look at each of

the uppercase letters of the alphabet, printed in a standard font, and to write descriptions of their appearance. We assumed that because explicit visual properties would be immediately evident, most participants would write them down. Moreover, we assumed that if a property was mentioned by fewer than 5% of the participants, it could be considered *implicit*.

## Method

### Participants

Sixteen participants (8 of them male and 8 female; mean age, 24 years) volunteered to take part in the experiment; all but 1 were right-handed, and none reported deficits in visual or auditory perception. None of the participants were aware of the purposes or predictions of the study prior to testing. The participants were paid \$10, and all gave their informed consent to take part in the experiment. The participants in this, and all the other experiments reported in this article, were tested in accordance with all applicable rules and regulations governing the treatment of human research participants. The study was approved by the Committee on the Use of Human Subjects, Faculty of Arts and Sciences, Harvard University.

### Materials

We prepared two lists of the 26 uppercase versions of the letters of the alphabet, printed in 14-point Arial font. Four blank lines were provided next to each letter, to allow the participants to write their descriptions.

### Procedure

We asked the participants to imagine that they were teaching an elementary English class to a group of adult students who used a different alphabet, such as Chinese or Arabic. In order to teach these students the letters of the alphabet, our participants were to describe each of the 26 uppercase letters of the Roman alphabet. Specifically, we asked the participants to write the description of the appearance of each letter, to help such students recognize them. We asked the participants not to compare the letters to visual objects or scenes, such as *hat* or *mountain valley*, but, instead, to describe the visual properties of the letters.

The participants wrote their descriptions of each letter next to a printed version of that letter in the prepared sheets; they were told that if they needed additional space, they should feel free to use the reverse side of the sheet. We addressed possible effects of practice and fatigue by asking half of the participants to describe the letters of the alphabet in forward order (A–Z) and the other half to describe them in reverse order (Z–A).

The participants also filled out a small set of short questionnaires that assessed characteristics that might influence their imagery of the letters. They responded to the following, in the order listed: (1) a question asking why they had included the visual properties that they mentioned in their letter descriptions, in which possible responses were (a) “I couldn’t think of any other,” (b) “they most easily came to mind,” or (c) “they seemed most relevant”; (2) the “vision” section of a Personal Health History form; and (3) a debriefing form asking about (a) the participants’ general idea of the study, (b) their understanding and following of the instructions, (c) their assessment of how difficult the task was, and (d) the strategies they used to complete the task of writing the descriptions. The participants were given as much time as they wanted to complete the task and typically required between 30 and 45 min to write their descriptions of the letters.

## Results

### Coding

In the experiments reported in this article, we examined a total of eight different properties. The properties were as follows, with the corresponding probe (used in later

experiments) following each: (1) all the lines in the letter were straight, with no curves (the probe for this was “all straight”); (2) there was at least one curved line in the letter (“curved line”); (3) there was at least one diagonal line in the letter (“diagonal line”); (4) the letter contained a fully enclosed space, such as the triangle embedded in the letter “A” (“enclosed space”); (5) the letter contained a semi-circle or an incomplete circle (“semi-circle”); (6) at least one side of the letter was a horizontal or vertical straight line (“straight side”); (7) the letter was symmetrical, either horizontally or vertically (“symmetrical form”); and (8) the letter contained two endpoints that do not intersect with any other lines (“two terminators”).

To be scored as *explicit* in the letter descriptions, the property either had to be mentioned literally or had to be indicated by a synonym. The synonym could be a multiword phrase that described the property or a special case of the property; for example, a “loop” is a certain type of curved line. For each of the noted properties, we coded the property as *explicit* if the participants included the following descriptions: for “curved line,” a “loop,” “circular line,” “oval,” “arc,” or other synonym signifying a curved line; for “all straight,” straight lines being mentioned or implied by noting the horizontal, vertical, or diagonal orientation of a line; for “diagonal line,” “slanted line,” “line going at a 45-degree angle,” “straight line going at an angle,” “a line going from top right to bottom left,” or other synonym; for “semi-circle,” “half-circle,” “incomplete circle,” “circle with a segment missing,” or other synonym; for “enclosed space,” “enclosed area,” “closed,” “surrounded,” “enclave,” or another synonym; for “straight side,” noting that at least one side of the character was composed of a straight line, such as by indicating that “there is a vertical line on the left side”; for “symmetry,” “symmetrical shape,” “mirror image,” or a statement such as “the same on the left and the right” or another synonym; for “two terminators,” “endpoints,” “ending,” “terminating,” or “finishing,” as well as “two,” had to be mentioned for an explicit coding to be assigned.

For each letter that contained a specific visual feature, we computed the percentage of participants who mentioned that property. When a property was mentioned correctly more than 50% of the time, we considered it to be explicit; when a property was present but was mentioned less than 5% of the time, we considered it to be implicit.

Two coders rated the descriptions to determine whether or not a property had been mentioned in a certain description. Reliability of the codings between the two raters ranged from 79% to 100%, depending on the property (see Table 1). The mean reliability was 92%. In cases of disagreement, the coders discussed the disparity, and a consensus was reached. Table 1 shows the percentage of the cases (over the 16 participants) in which each property was mentioned appropriately.

In addition, 63% of the participants indicated that they included in their descriptions the visual properties that came most easily to mind, 6% indicated that they could not think of any other, and the remaining 31% indicated that they included the properties they considered most rel-

**Table 1**  
Percent Mention of Properties, Classification, and Reliability of Classifications in Experiment 1

Property Name	% of Correct Mentions	Explicit or Implicit	Intercoder Reliability (%)
All straight	61	explicit	92
Curved line*	95	explicit	92
Diagonal line*	76	explicit	88
Enclosed space*	2	implicit	100
Semi-circle	60	explicit	85
Straight side	16	undetermined	79
Symmetry*	0	implicit	99
Two terminators	<1	implicit	99

Note—Properties mentioned in 50% or more of the cases were considered to be explicit, whereas those mentioned in 5% or fewer of the cases were considered to be implicit. \*Property was used in Experiment 2.

evant. In short, the descriptions of 69% of the participants apparently reflected the ease with which they could notice the visual properties of the letters. This finding is consistent with our assumption that the descriptions indicate which visual properties are explicitly represented when people see the letters. Perhaps equally important, the fact that many sorts of descriptions were rarely, if ever, mentioned supports our assumption that these properties were not included in explicit representations of the shapes.

## Discussion

We clearly could distinguish between two classes of properties. One class (which included properties such as *diagonal lines*) was explicitly mentioned by at least half of the participants. We conjecture that these properties not only are explicitly mentioned in the descriptions the participants provided, but also are explicitly encoded in the internal representations of the shapes. In contrast, another class (which included properties such as *symmetrical form*) was virtually never mentioned. We conjecture that these properties are not explicitly included in the internal representation.

We should note that the distinction between explicit and implicit properties, as we have defined them here, is different from various other distinctions. For example, it is not the case that implicit properties are simply easier to verbalize than explicit ones, as witnessed by the ease with which we could name members of both categories. Nor is it the case that implicit properties are more complex than explicit ones or vice versa. Similarly, implicit properties need not be smaller or less visible than explicit ones (e.g., symmetry, an implicit property, characterizes the shape as a whole). Moreover, implicit properties need not be local and explicit properties global. However, global properties in general may tend to be stored explicitly because they are, by definition, relatively large and, thus, are likely to be easily noticed. However, as the symmetry example shows, simply being relatively large does not ensure that a property will be stored explicitly. Local properties, in contrast, may often be noticed later and may tend to be stored implicitly. But again, simply being relatively small is not enough to ensure that a property will be stored implicitly; even local features may be stored explicitly if they are distinctive enough (e.g., the tail on a “Q”). Indeed, Kimchi (1992) reviewed the lit-

erature on the global precedence effect and found that many factors, including relative size, overall visual angle, and attention allocation, can affect whether the global properties of an object are, in fact, processed first.

The findings from Experiment 1 provide the bases of the following experiment, which attempts to show that the representations used when one is perceiving a shape are structurally the same as those used when visualizing it.

## EXPERIMENT 2

In perception, an observer can take a "closer look" and glean information not noticed previously. We asked whether the same is true in visual mental imagery and considered two predictions that stem from the two major theories of image format. First, consider a prediction from the theory that visual mental images are stored in a depictive format, much like that used in the early phases of perception (see Kosslyn et al., 2006). According to this theory, it should be comparably easy (on the basis of RTs and error rates [ERs]) to inspect shapes in mental images for implicit properties as it is to inspect shapes during perception for those properties (in both cases, relative to the ease of inspecting shapes for explicit properties). Such a result would be consistent with the view that visual mental images are like perceptual images in that they can be reorganized and inspected with greater care. Second, consider a prediction from the theory that visual mental images are stored in a descriptive format, much like that used in language. According to this theory, it should be more difficult to inspect shapes in mental images for implicit properties than it is to inspect shapes during perception for those properties (relative to explicit properties). This finding would suggest that image representations include only explicit properties and that implicit properties need to be computed and inferred, which is not necessary in perception (where the input provides all the information needed to extract either type of property).

## Method

### Participants

We tested 56 participants (28 of them male, 28 female), almost all of whom were college students (mean age, 21 years). None of these participants had taken part in Experiment 1. Equal numbers of male and female participants were assigned to the two groups (imagery and perception). We tested and included an additional male participant after excluding one of the original male participants because of an excessive (79%) error rate on one of the probes, which suggested that he did not understand the meaning of that probe. All the participants reported having normal hearing and normal or corrected-to-normal vision. All the participants provided informed consent, were compensated at a rate of \$10/h, and were not informed of the purpose or goals of the experiment before they participated.

### Materials

The characters on which participants based their responses were in 130-point Arial font on the computer screen. However, the numerals 1, 2, and 7 (used in practice trials) were corrected because their diagonals were slightly curved; we straightened these curves to eliminate possible confusion regarding whether these characters contained a diagonal. The letters presented in the study phase were saturated black, whereas the letters used in the test phase in the per-

ception condition of the experiment were presented in a medium-dark shade of gray, in an attempt to equate partially for the clarity of mental images. For the imagery condition, we also prepared a 300-msec audiofile of the spoken name of each letter.

We selected the explicit visual properties according to the criteria that they were (1) mentioned more than 50% of the time by the participants who wrote the descriptions in Experiment 1 and (2) were present in 20%–80% of the letters. These criteria led us to choose as explicit properties "curved line" (correctly mentioned 95% of the time, present in 42% of the letters) and "diagonal line" (correctly mentioned 76% of the time, present in 38% of letters). We selected the implicit properties on the basis of the criteria that they were (1) mentioned less than 5% of the time by the participants in Experiment 1 and (2) were present in 20%–80% of the letters. We selected "enclosed space" (correctly mentioned 2% of the time, present in 27% of the letters) and "symmetrical form" (mentioned 0% of the time, present in 73% of the letters). Four audio files were created, one for each probe, each of which was approximately 200 msec in duration; these files contained abbreviated words that designated each of the four auditory probes: "curve" for "curved line," "close" for "enclosed space," "diag" for "diagonal line," and "sym" for "symmetrical form." We reasoned that if we had used the full names of the properties, the participants might have begun processing before the probes were completed; if so, shorter auditory probe durations should result in more accurate measurements of RTs.

A trial consisted of a combination of a letter and a probe that prompted the participants to judge whether the letter had a certain property. Each participant was prompted with the same series of 56 letter/property pairs (e.g., "U . . . enclosed space"—for which the correct response would be *no*, given that the interior space of the letter is not surrounded by lines on all sides) in the same order. All the letters and properties occurred once before any one occurred a second time, all occurred twice before any occurred three times, and so forth. Furthermore, we designed the series of trials so that (1) each of the 26 letters appeared either two or three times; (2) each property was probed exactly 14 times; (3) for each property, there were seven trials on which that property was present and seven trials on which the property was absent; and (4) no more than three consecutive trials could have the same correct response.

We prepared two sets of trials. In the perception condition, a letter was presented at the center of the screen. After 1,500 msec, one of the four auditory probes was presented. The letter remained on the screen until the participant responded either by pressing the "b" key on the keyboard (for *yes*) or the "n" key on the keyboard (for *no*). One second after they responded, they saw the next letter.

The imagery trials had the same format, except that the name of the letter was presented auditorily and the letter itself was not shown. Specifically, the computer presented a spoken letter and then presented one of the four probes (again auditorily) 1,500 msec after the end of the name of the letter. The 1,500-msec delay allowed enough time for the participants to form an image of the character in the imagery condition and to have it ready when the judgment probe was presented; hence, the RT should reflect more accurately the time to make the judgment, and not the time required to generate the image.

Finally, we prepared a block of practice trials for each condition. These trials had the identical format as the test trials but used digits as stimuli.

### Procedure

**Learning phase.** The participants began by memorizing the appearance of the 26 uppercase letters of the Roman alphabet and the numerals 1–9 in 130-point Arial font on a Macintosh G4 with a 17-in. color monitor, 1,024 × 768 pixel resolution. (Numerals, which were not described in Experiment 1, were included in the practice trials, but not in the test trials, and thus did not contribute to the results presented below.) The participants sat approximately 60 cm from the screen. The characters were presented in pseudorandom order until all of them had been presented once, at which point the characters were repeated in the same order. Each character appeared for 3 sec

and was accompanied by its spoken name. After a character had been presented, the participants were asked to visualize it until they felt satisfied that their mental image was accurate. At this point, they pressed the space bar and saw the character again on the computer screen. They now were to compare their mental image with the actual character and to correct their image as appropriate. After a character had been presented, a visual mask appeared in order to eliminate any afterimage of the character. We included the mask to ensure that the participants actually generated the images of each character from memory, rather than relying on a residual iconic image. After the participants felt that they had fully corrected their visual mental image of the character, they pressed the space bar and the next character was presented. The learning phase required about 30 min, after which the participants were allowed to take a short break.

**Defining probes.** Before practicing the experimental task, the participants reviewed the different probes thoroughly. The meanings of the four probe terms were illustrated in figures in which the property was or was not present. Specifically, the meanings of the probe terms were defined with sets of abstract figures that appeared on the screen. The figures (which were mainly based on characters from the Linear B syllabary, or modifications thereof) were roughly as complex as the letters and numbers but did not resemble specific characters. We asked the participants to explain to the investigator why each figure had or did not have a particular property. For each property, they saw a set of three or four unfamiliar symbols (see Figure 1). After seeing each of the probes and the accompanying figures, the participants pressed the space bar to request the correct answers. The participants were instructed to review the correct answers with the investigator if anything was unclear.

Next, we asked the participants to listen to the auditory probes for each property. Each of these probes was approximately 200 msec long and consisted of a syllable or two of the property name, as was noted earlier. Once the participants had listened to each probe twice, we tested their comprehension by asking them to listen to 16 auditory probes (4 for each property, in randomized order) and to say aloud the full name of the property being prompted. If they made an error on the full name of the property, the investigator pointed this out, and they were asked to correct their response.

Finally, after they were familiar with the auditory probes, we retested the participants with new sample shapes, assessing their

knowledge of the property definitions. The participants explained to the investigator why each shape did or did not have the relevant property. The investigator again reminded them of the definitions of each of the four properties.

**Practice trials and test trials.** Half of the participants next took part in the perception condition, and half took part in the imagery condition. Because the same questions and stimuli were used in both the perception and the imagery conditions, once a participant had been exposed to one condition, the correct answers could be rehearsed verbally. Thus, the imagery and perception conditions were performed between participants. Both conditions began with a set of 16 practice trials, and the stimuli were presented using PsyScope software running in Macintosh OS 9. In the perception condition, we asked the participants to fixate their gaze on the center of the screen, view each letter presented to them, listen to the probe, and respond by pressing the appropriate key. The participants were to respond *yes* if the probed property was in fact present in the character and *no* if it was not. When they responded correctly, they saw the next letter. If they responded incorrectly, they heard a short beep, and the screen did not advance until they provided the correct response.

In the imagery condition, we asked the participants to close their eyes, listen to the name of the character, visualize it (as it appeared when they studied it), and listen to the probe before responding *yes* or *no*. Otherwise, the two conditions were identical. In both conditions, the participants were asked to respond as quickly and accurately as possible.

The participants practiced the task using the digit stimuli. The participants received each probe four times, twice when the answer was *yes* and twice when it was *no*. The different number-question-answer combinations were presented in pseudorandom order. The practice phase required about 10 min, and after the practice phase, the participants were given a brief break.

Before beginning the test phase, the investigator reminded the participants that their responses and RTs would be recorded and, so, they should respond as quickly and accurately as possible. In addition, the investigator again reviewed how to perform the task. The test trials were identical to the practice trials, only in this case, the participants were queried about letters instead of numbers for both the explicit and the implicit visual properties. In addition, the computer did not beep when the participants made an error during

### Curved Line

Please indicate which of the shapes below possess the property *curved line*, and, in your own words, explain why.



When you have indicated your answers to the researcher, please press the space bar to continue.

### Diagonal Line

Please indicate which of the shapes below possess the property *diagonal line*, and, in your own words, explain why.



When you have indicated your answers to the researcher, please press the space bar to continue.

### Enclosed Space

Please indicate which of the shapes below possess the property *enclosed space*, and, in your own words, explain why.



When you have indicated your answers to the researcher, please press the space bar to continue.

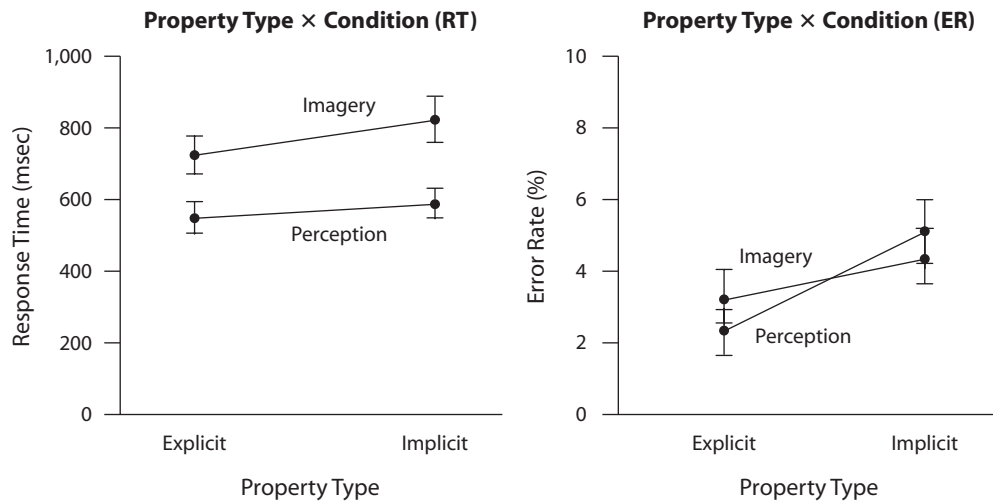
### Symmetrical Form

Please indicate which of the shapes below possess the property *symmetrical form*, and, in your own words, explain why.



When you have indicated your answers to the researcher, please press the space bar to continue.

**Figure 1.** Examples of the figures that were used to train participants to recognize the four properties: curved line, diagonal line, enclosed space, and symmetrical form. The participants were asked to examine the figures and to decide, for each one, whether or not it featured a given property. They were asked to explain their judgment to the investigator.



**Figure 2.** Response times (left) and error rates (right) for property type (explicit/implicit) in each condition (imagery/perception). Participants' pattern of performance in the imagery condition was comparable to that in the perception condition, for both implicit and explicit properties. Error bars represent standard errors of the means.

the actual test trials, and the participants were not required to repeat trials on which they had made an error.

The participants in the perception condition were told that they should view each letter and listen to the name of that letter and that they then would hear the name of one of the property probes and should evaluate it. If the property was in fact included in the letter, they should respond *yes*; if not, they should respond *no*, pressing the appropriate key as quickly and accurately as possible. RTs were recorded from the completion of the property probe. A new trial began 1,000 msec after the response.

The procedure in the imagery condition was identical to that in the perception condition, but with two changes. First, the participants were asked to keep their eyes closed and to form an image of the letter (as studied previously) when they heard its name. Second, when they heard the probe, they were to evaluate the letter in their visual mental image and then respond as quickly and accurately as possible by pressing the appropriate key.

## Results

We performed three sets of analyses on the data, analyzing separately RTs and ERs in each set. First, we used an ANOVA to address the primary question we posed—that is, the relationship between the perception and imagery conditions for the different property types (explicit vs. implicit). If the representations underlying perception and imagery make accessible the same type of information, these two variables should not interact. In addition, in order to investigate possible differences between perception and imagery, we also examined the relative difficulty of the four probed properties and the effects of gender and response type. We were concerned that if we did not find an interaction between the ease of evaluating the two types of properties (implicit vs. explicit) in the two conditions (imagery vs. perception), this would be a type of null effect. Perhaps the data were simply noisy, or our techniques were not sufficiently sensitive. Thus, to demonstrate that we could detect other patterns in the data, we followed up our first analysis

by examining the effects of the visual complexity of the letters on RTs and ERs in the two conditions. Finally, we computed the correlations between the 56 individual items in perception and imagery, separately for RTs and for ERs. If property discriminability plays the same role in perception and imagery, the RTs and ERs for the items in the two conditions should vary in comparable ways.

All analyses of RTs include only data from trials on which the participant provided the correct answer. We also excluded trials with an RT less than 150 msec, reasoning that such a short RT could not reflect processing time for the task and more likely would be due to a reflexive response. Because standard deviations were lower than the mean RTs for a cell, we adopted the criterion of excluding RTs that exceeded 2.5 times the mean for each cell. This criterion represents a compromise between no trimming at all (where long RT trials may or may not reflect task-relevant processing) and trimming based on standard deviations, where more trials likely to reflect actual processing are susceptible to being eliminated. Our procedure led us to treat 2.2% of the data as outliers.

## Effects of Probe Type in Perception and Imagery

**Response times.** A two-way ANOVA provided no evidence that the time to evaluate the probes was different in the perception and imagery conditions [ $F(1,54) = 1.49$ ,  $p > .2$ , for the interaction between condition and probe type]. The means were as follows: perception, explicit, 564 msec; perception, implicit, 596 msec; imagery, explicit, 746 msec; imagery, implicit, 816 msec. These results are illustrated in Figure 2 (left). However, we did find that participants required more time to judge the implicit properties than to judge the explicit ones [ $F(1,54) = 11.28$ ,  $p < .002$ ; explicit  $M = 655$  msec, implicit  $M = 706$  msec]. We also found that the participants in the imagery condition

required more time overall than did the participants in the perception condition [ $F(1,54) = 15.22, p = .0003$ ; imagery  $M = 781$  msec, perception  $M = 580$  msec].

To investigate whether the RT effects observed here may reflect more complex interactions, we performed four-way ANOVAs that included the following variables: condition, property, gender, and response type. We used the same trimming procedure as that previously described, which led us to exclude 1.9% of the data. We again found no hint of an interaction between property and condition ( $F < 1$ ). In addition, we documented that the participants required more time in the imagery condition [ $F(1,52) = 13.65, p = .0005$ ] and also found that the participants required different amounts of time for the different properties [ $F(3,156) = 54.85, p < .0001$ ; with means for curved line = 597 msec, diagonal line = 716 msec, enclosed space = 603 msec, symmetrical form = 844 msec]. Post hoc contrasts revealed that RTs for all pairs of properties, except curved line and enclosed space, were different ( $p < .0001$  in all cases).

We also found that participants required more time in general for *no* responses than for *yes* responses [ $F(1,52) = 17.28, p < .0001$ ;  $M = 715$  vs. 666 msec] and found an interaction between response and property [ $F(3,156) = 12.20, p < .0001$ ]. Post hoc contrasts revealed that the participants required more time to answer *no* than to answer *yes* for symmetrical form ( $p < .0001$ ), whereas there was no such difference for the other three properties. No other main effects or interactions were significant ( $p > .05$  in all cases).

**Error rates.** We performed an analogous ANOVA for ERs. Again, there was no interaction between condition and property type [ $F(1,54) = 1.26, p > .2$ ]. The means, in percentages of errors, were as follows: perception, explicit, 2.4%; perception, implicit, 5%; imagery, explicit: 3.3%; imagery, implicit, 4.2%. These results are illustrated in Figure 2 (right). However, the participants made more errors on trials probing implicit properties than on those probing explicit properties [ $F(1,54) = 5.45, p < .03$ ; implicit  $M = 4.6\%$ , explicit  $M = 2.9\%$ ]. The participants in the two conditions made comparable numbers of errors ( $F < 1$ ). (Note that even here, the participants in the imagery condition were correct for implicit properties at least 95% of the time, which demonstrates that they could, in fact, perform the task well.)

We performed a second ANOVA with the same four variables used in the second analysis of RTs, now considering the ERs. We again did not find an interaction between property and condition ( $F < 1$ ) but did find that the participants made more errors for some of the properties than for others [ $F(3,156) = 9.4, p < .0001$ ]; the mean error rates by property were as follows: curved line = 1.8%, diagonal line = 4.0%, enclosed space = 2.4%, and symmetrical form = 6.8%. Follow-up tests revealed that the participants made more errors on the symmetrical form judgment than on each of the other three ( $p < .01$  in all cases) and more errors on the diagonal line judgment than on the curved line judgment ( $p < .05$ ).

We also found that the participants made more errors when the correct response was *yes* than when it was *no*

[ $F(1,52) = 19.55, p < .0001$ ; 5.2% vs. 2.3%]. Although we did not observe a difference between genders in their overall ERs, there was a three-way interaction between property, gender, and response type [ $F(3,156) = 2.67, p < .05$ ]. Males made relatively more errors on *yes* trials than on *no* trials for the diagonal line property, whereas females made relatively more errors on *yes* trials than on *no* trials for the enclosed space and symmetrical form properties. There were no other significant main effects or interactions.

Crucially, in none of the comparisons did we find an interaction with condition; if a variable affected perception, it also affected imagery—and did so in the same way and to the same extent.

### Complexity Effects

To plumb in greater depth the correspondence between imagery and perception and to investigate further the issue of whether our paradigm simply lacked sensitivity, we considered the effects of visual complexity in both conditions. We began by testing 20 additional participants (10 of them male, 10 female; mean age, 21 years). All of these participants reported normal hearing and normal or corrected-to-normal vision. They were not informed of the purpose of the study before testing and were paid at a rate of \$10/h. None of these people had participated in Experiment 1 or the main study in Experiment 2 (the results of which were just reported).

We showed these participants the stimuli used in this experiment, one at a time, and asked them to study each character, “paying careful attention to the specific *shape* of each one.” After the participants had learned the characters, using a version of the memorization procedure that was used in the main study, we asked them to rate the “complexity of the shape” of each character on a scale from 1 to 7: The participants were instructed to press “1” on the keyboard if they found the shape of the given character to be *very simple* relative to the others, to press “7” if they found it to be *very complex*, and to use intermediate values to designate *intermediate levels of complexity*. To discourage the participants from rating characters for complexity of a different kind (e.g., “Z” might be considered *complex* because it is used relatively infrequently in English words), we included the following sentence in the instructions: “Please remember that you are rating the complexity of the *shape* of each character, not the complexity of the character in relation to any other property.”

The mean complexity ratings for the 26 letters ranged from 1.0 (for “I”) to 5.6 (for “G”), with an overall mean of 3.28 (median: 3.3). Using a median split, the visually *simple* letters were I, O, L, V, T, U, J, C, X, D, H, and P, and the visually *complex* letters were Y, Z, F, N, E, Q, W, A, S, M, B, K, R, and G.

To demonstrate that we had the statistical power to detect effects of interest (if they were present), we tested the hypothesis that it is more difficult to identify properties in more complex letters. We performed a one-way ANOVA on RTs and ERs, with complexity as the independent variable.

**Response times.** Using the rated complexity of the letters, we found a difference between simple and complex

letters [ $F(1,55) = 12.07, p = .001$ ]. The mean RT for simple letters was 657 msec, and the mean RT for complex letters was 695 msec.

**Error rates.** We did not find differences in ERs for simple letters (3.6%) versus complex letters (3.8%) [ $F(1,52) < 1$ ].

Aside from the finding that more time was required to evaluate letters in imagery than in perception (see above), no other main effects or interactions were found when letter complexity was included as a variable ( $p > .1$  in all cases).

This finding that the participants required more time when they evaluated complex letters than when they evaluated simple ones—in addition to all of the other effects and interactions reported above—suggests that we did not lack statistical power to detect findings that were, in fact, present. Nevertheless, condition (imagery vs. perception) never interacted with any other variable.

### Correlational Analysis

**Response times.** We found a very high correlation coefficient when we correlated the mean RT for each item in perception with the corresponding mean RT for that item in imagery ( $r = .87, p < .0001$ ). This is strong evidence that the same factors were responsible for variations in RTs for the items in perception and in imagery.

**Error rates.** We performed the same analysis with the ERs and found a correlation coefficient of  $r = .49, p = .0001$ , between perception and imagery. Again, we have evidence that the same factors that affect processing in perception also affect processing in imagery.

## GENERAL DISCUSSION

The present results provide strong support for the claim that previously unnoticed properties of objects can be reliably detected in visual mental images. Moreover, the results provide solid evidence that the same representations and processes are used in both visual perception and visual mental imagery in this task. In both cases, there is “something to examine more carefully” when one is asked about a property that was not immediately noticed, such as whether a shape is symmetrical along its vertical or horizontal axis. Nobody would be surprised to learn that this is true in perception, where the shape itself remains visible, but some researchers have denied that this is possible in mental imagery. According to these researchers, either an image is accompanied by a description, which locks in a certain organization (cf. Chambers & Reisberg, 1992; Reisberg, 1997), or images are nothing more than descriptions (Pylyshyn, 1973; for a review, see Kosslyn et al., 2006). If either of these possibilities were correct, it should be much more difficult to identify implicit properties, relative to explicit ones, in imagery than in perception; but this was not the case.

We should note that although we used letters as the stimuli in this study, our predictions do not rely specifically on such stimuli. We selected letters as the test stimuli because (1) they are well-known shapes, and thus it is likely that even specific exemplars could be learned ef-

ficiently, and (2) they have clear prototypes, even across most font styles, and thus they lend themselves to the types of descriptions that we requested from the participants in Experiment 1. For a reexamination of the details of an image to be useful, it is essential that the details be recalled vividly and that they be represented accurately in the image. Hence, we used stimuli that we could teach participants quickly and that, in spite of the short training period, were likely to give rise to accurate mental images. We expect the same pattern of results—that is, that previously unnoticed properties can be examined and verified in imagery—for any stimulus that can be visualized with a high degree of accuracy.

Thus, these findings also bear on previous research in which participants have been asked to verify a detail in an image of a more complex object. For example, in some studies, participants have been asked to verify properties of animals, such as the shape of their ears (e.g., Kosslyn, 1975). Many such properties are relatively subtle, and hence it is likely that they are implicit. If so, these previous findings are consistent with the present results, which show that people can inspect objects in images and notice properties that are only implicit in the image.

This ability also allows shapes in mental images to be combined to form novel patterns (with new sets of properties that can be verified using imagery). For example, Finke (1990) asked participants to memorize drawings of complex shapes and then asked them to visualize combinations of the shapes that would create novel objects. The participants were then asked to draw the objects that they had produced using imagery. Some of these objects were very creative, and the participants reported noticing emergent visual properties that they had created using imagery, some of which were totally unexpected.

In order to use imagery in such ways, one must be familiar enough with the stimulus to visualize it in detail. But more than that, because images are representations in working memory, for reinterpretations to take place (or for unnoticed features to be detected), working memory capacity should not be overly stressed, or the details of the image will not be maintained long enough to be inspected accurately. Indeed, Mast and Kosslyn (2002) demonstrated that reducing the load on working memory allowed participants to mentally rotate and reinterpret complex figures.

The fact that previously unnoticed properties are as accessible (relative to explicit properties) in mental imagery as in perception has broad implications for theories of mental imagery and mental representation in general. Pylyshyn (1973, 2002), for example, has argued that mental images are represented as symbols akin to descriptions or propositions. If mental images were represented in a descriptive format, the organization of the representation would probably be based on explicit properties of the stimulus. If the visual imagery system relies on such propositional representations and the visual perceptual system (as is widely accepted) relies in part (particularly during the relatively early phases of processing) on depictive representations, we would expect that implicit properties would be less accessible in imagery than in perception. When information about implicit properties must be retrieved



from memory, as in Experiment 2, we would expect implicit properties to require more time to verify (relative to explicit properties) in imagery than in perception, because a computation would need to derive the properties. This prediction was not borne out by the results.

Depictive representations in general have the characteristic of making emergent visual properties explicit, and thus any visual property, even if not explicitly noticed and stored as a description, can be examined and evaluated in perception. Our results suggest that perceptual and imagery representations in working memory share the same representational format. The evidence from this study supports the view that visual mental imagery, like perception, relies in part on depictive representations. The imagery system *makes* implicit properties explicit and accessible (as when one is examining a picture in front of one's eyes).

These findings relate to an even more fundamental question in psychology, which also has implications in philosophy. Pylyshyn (2002) has claimed that in the absence of convincing evidence for depictive representations, it makes the most sense (i.e., is most parsimonious) to posit a single "language of thought," which relies on symbol systems. We have argued that there is strong evidence for depictive representations in imagery (see Kosslyn, Ganis, & Thompson, 2003; Kosslyn, Thompson, & Ganis, 2002, 2006), and the present results support this view, with the wider implication that there is more than a single language of thought. If the mind were a toolkit, it would have more than a single pair of pliers; it would have a host of different tools, with different ones being appropriate for different tasks.

We note that, from an evolutionary perspective, once the perceptual system is in place, it may be adaptive for memory and reasoning processes to take advantage of these mechanisms. For example, it may be advantageous to be able to reactivate a visual mental image and reexamine it, to search for properties that were not previously noticed. Depictive representations would be especially useful because they allow a large amount of information to be preserved in an efficient format. And the fact that images can then be recombined to produce novel scenes can play an essential role in planning and reasoning more generally.

In short, our results not only provide evidence that visual mental images share a representational system with visual perception, but also support the view that depictive representations are used in both imagery and perception. This finding is consistent with much neuroimaging work (e.g., Kosslyn & Thompson, 2003), which demonstrates that the topographically organized cortex—which literally supports depictive representations—is activated when shapes must be inspected with high resolution, as was the case in the present experiments.

#### AUTHOR NOTE

This research was supported by Grant 2 R01 MH060734-05A1 from the National Institute of Mental Health. Any opinions, findings, conclu-

sions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institute of Mental Health. We thank Graham Blackman for many helpful suggestions and for his work on the Experiment 2 training phase and in testing the participants; in addition, we thank Michael Ermiyas for his help with coding the letter descriptions provided by the participants in Experiment 1. We also thank Tom Henderson for testing participants in Experiment 1. Correspondence concerning this article should be addressed to S. M. Kosslyn, Department of Psychology, Harvard University, 830 William James Hall, 33 Kirkland Street, Cambridge, MA 02138 (e-mail: smkosslyn@wjh.harvard.edu).

#### REFERENCES

- ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory*. Washington, DC: Winston.
- CHAMBERS, D., & REISBERG, D. (1992). What an image depicts depends on what an image means. *Cognitive Psychology*, **24**, 145-174.
- FINKE, R. A. (1990). *Creative imagery: Discoveries and inventions in visualization*. Hillsdale, NJ: Erlbaum.
- KIMCHI, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, **112**, 24-38.
- KOSSLYN, S. M. (1975). Information representation in visual images. *Cognitive Psychology*, **7**, 341-370.
- KOSSLYN, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- KOSSLYN, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
- KOSSLYN, S. M., GANIS, G., & THOMPSON, W. L. (2003). Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Sciences*, **7**, 109-111.
- KOSSLYN, S. M., & POMERANTZ, J. R. (1977). Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, **9**, 52-76.
- KOSSLYN, S. M., & THOMPSON, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological Bulletin*, **129**, 723-746.
- KOSSLYN, S. M., THOMPSON, W. L., & GANIS, G. (2002). Mental imagery doesn't work like that. *Behavioral & Brain Sciences*, **25**, 198-200.
- KOSSLYN, S. M., THOMPSON, W. L., & GANIS, G. (2006). *The case for mental imagery*. New York: Oxford University Press.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- MAST, F. W., & KOSSLYN, S. M. (2002). Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition*, **86**, 57-70.
- PETERSON, M. A., KIHILSTROM, J. F., ROSE, P. M., & GLISKY, M. L. (1992). Mental images can be ambiguous: Reconstructs and reference-frame reversals. *Memory & Cognition*, **20**, 107-123.
- PYLYSHYN, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, **80**, 1-24.
- PYLYSHYN, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, **88**, 16-45.
- PYLYSHYN, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral & Brain Sciences*, **25**, 157-238.
- PYLYSHYN, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.
- REISBERG, D. (1997). Constraints on image-based discovery: A comment on Rouw et al. (1997). *Cognition*, **66**, 95-102.
- REISBERG, D., & CHAMBERS, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology*, **45**, 336-352.
- ROUW, R., KOSSLYN, S. M., & HAMEL, R. (1997). Detecting high-level and low-level properties in visual images and visual percepts. *Cognition*, **63**, 209-226.

(Manuscript received September 14, 2007;  
revision accepted for publication February 6, 2008.)