

Comparison processes on visual mental images

SEBASTIANO BAGNARA
Università di Siena, Siena, Italy

FRANCESCA SIMION and MARIA E. TAGLIABUE
Università di Padova, Padova, Italy

and

CARLO UMILTÀ
Università di Parma, Parma, Italy

Three experiments were conducted to test whether visual mental images and visual perceptual representations possess equivalent structural properties and undergo functionally equivalent comparison processes. In Experiment 1, subjects were required to perform a *same-different* letter classification in which the two letters were shown in succession. The first letter in the pair either was actually presented (perceptual condition) or had to be actively generated (imaginal condition). Both conditions showed that (1) response latencies for *different* decisions decreased as a function of the degree of difference (*segment effect*), and (2) *same* decisions were faster than the fastest *different* ones (*fast same effect*). In Experiment 2, the imaginal condition only was employed, but no imagery instructions were given and very strict time constraints were assigned. In spite of these restrictions, subjects apparently generated and used visual images, as attested by the fact that the results were comparable to those of the previous experiment. In Experiment 3, three experimental manipulations were introduced to prevent the use of visual images. Such manipulations proved effective, as shown by the disappearance of the segment effect. It was concluded that in the visual modality images and percepts are equivalent in structure and are processed in a very similar way.

Since the seminal paper by Shepard and Metzler (1971), much effort has been devoted to the study of visual mental imagery (see reviews in Kosslyn, 1980; Kosslyn, Pinker, Smith, & Shwartz, 1979; Shepard, 1984; Shepard & Cooper, 1982). The field of imagery research has come to be characterized by heated debates (Kolers, 1983) and by various identifiable subfields (Finke, 1985; Pinker, 1984). A core question in the field is still whether cognitive processes on visual images resemble perceptual processes (Finke, 1985; Shepard & Podgorny, 1978), that is, whether mental operations on internally generated visual images are similar to those on visual representations of perceived objects.

The properties possessed by both types of representation and the mental operations that can be performed on them are often thought to be functionally equivalent (Finke, 1980, 1985). It has been shown that visual images possess spatial properties (e.g., extent, Kosslyn,

1980; boundaries, Podgorny & Shepard, 1978), may be constructed from individual parts (Kosslyn, Reiser, Farah, & Fliegel, 1983), and even produce visual illusions (Berbaum & Chung, 1981; Wallace, 1984). Like perceived objects, visual images can be scanned (Finke & Pinker, 1982, 1983; Kosslyn, Ball, & Reiser, 1978; Pinker, Choate, & Finke, 1984) and tested for acuity (Finke & Kurtzman, 1981), and the smaller or concealed ones are examined and recalled less well than larger or unconcealed ones (Keenan, 1983; Kosslyn & Alper, 1977). However, many of the instances of equivalence between images and percepts are not fully convincing because the comparisons were not properly performed. As suggested by Banks (1981, p. 847), in order to test the equivalence hypothesis, "it would seem wiser to compare perception and imagery with a stimulus that has been thoroughly tested in perceptual paradigms." This suggestion has not been followed very often.

One of the most striking instances of equivalence is the availability in images of information as detailed as that present in percepts. Podgorny and Shepard (1978) showed that both imaginal and perceptual representations are formed by units that can be processed independently. It has also been found that the generation of visual images can take place by adding elementary units to the whole configuration (Kosslyn et al., 1983). A true equivalence requires, however, that these units, besides being

The authors wish to thank R. A. Finke and R. W. Proctor for extremely helpful reviews of a previous version of the paper. They are also most grateful to S. Bettella for writing and implementing the computer programs. This work was supported in part by funds from the Consiglio Nazionale delle Ricerche and the Ministero della Pubblica Istruzione. Requests for reprints should be sent to Carlo Umiltà, Istituto di Fisiologia Umana, Università di Parma, via A. Gramsci, 14, 43100 Parma, Italy.

structural components of both types of representation, play the same role in the comparison of images and percepts. This comparison role can be examined in the manner suggested by Banks (1981), because the phenomena that occur in matching visual patterns have been extensively studied.

Elementary components and their use are thought (see, e.g., Bamber, 1969) to be instrumental in determining the well-documented *fast same effect* (FSE) in pattern matching, according to which *same* responses are usually much faster than *different* ones (Farell, 1985; Nickerson, 1978). The FSE has been explained (see, e.g., Taylor, 1976) by proposing a holistic processor that mediates *same* decisions and an analytic processor that mediates *different* decisions. The aim of the present study was to test whether the pattern of results that is taken as supportive of analytic processing (i.e., processing based on component units) could be replicated when visual images were compared with visual percepts. The study used Taylor's (1976) paradigm, in which letters are constructed from component segments. The use of this paradigm does not imply a commitment to the dual-processor model, which has been questioned by many authors (Bagnara, Boles, Simion, & Umiltà, 1983; Farell, 1985; Krueger, 1978, 1983, 1984; Proctor, 1981; Proctor & Rao, 1983). The paradigm was chosen because it seemed particularly well suited to demonstrate the role of the component units in a comparison process.

EXPERIMENT 1

In Taylor's (1976) experiment, subjects were required to make a speeded *same-different* classification of two letters presented one after the other. The letters were formed by component units (segments) and could differ because of the presence or absence of a segment in one, two, or three positions. The results showed that the time needed to classify two different letters decreased with the number of discrepant positions and the time for a correct *same* decision was faster than that for the fastest *different* decision. Taylor concluded that this pattern of results was in accordance with a dual-processor model of comparison, with *same* responses being emitted through a fast holistic processor and *different* responses through a slower analytic processor. Even if this notion is not accepted, there can be little doubt that *different* responses depended on a process that was sensitive to the elementary components, or at least to very fine details, of the two letters.

In order to test whether the above pattern of results could be obtained also in the case of visual images, we designed Experiment 1 so that one condition (the perceptual condition) was identical to that devised by Taylor (1976), whereas the other (the imaginal condition) required the subjects to form a visual image of the first letter.

Method

Subjects. Sixteen subjects from the Università di Padova, 8 of each sex, participated in the two experimental conditions, which were run on successive days. They were paid for their participation.

Apparatus. The subject sat alone in the experimental cubicle and was monitored by one of the experimenters through a unidirectional mirror, thus minimizing the possibility of inadvertently cuing the subject (Intons-Peterson, 1981, 1983). The stimuli were presented on a CRT screen driven by an Apple II computer that controlled the generation and timing of stimuli and responses. The center of the screen was at about eye level, at a distance of 50 cm from the subject.

Stimuli. The stimuli were letters derived from the five basic line segments forming the pattern shown in Figure 1. All contained the vertical segment on the left but varied as to the presence or absence of the other segments. When shown on the screen, a letter measured 14 mm in height and 8 mm in width and each component segment was 2 mm thick. The six stimuli and their relationships in terms of differing segments are shown in Figure 2.

Procedure. All subjects were run in both conditions, but half (4 of each sex) had the perceptual condition and half the imaginal condition on the first day.

Perceptual condition. Trials consisted of the presentation of the first (criterion) letter followed by the second (probe) letter. At the beginning of each trial an acoustic warning signal prompted the subject to fixate a mark in the center of the screen. Upon the offset of the mark (1 sec), the criterion appeared for 500 msec; then, after an interstimulus interval (ISI) of 500 msec, the probe was shown for 500 msec, followed by a 5-sec intertrial interval. The response was emitted by depressing one of two keys on a response panel. The latency between the onset of the probe and the depressing of the key was measured and responses shorter than 150 msec or longer than 1,500 msec were regarded as errors, as were those in which the wrong key was used. Half of the subjects used the right hand for *same* responses and the left for *different* responses, whereas the others had the reverse assignment.

A session consisted of 2 practice and 10 experimental blocks of 36 trials, 18 *same* and 18 *different*. The sequence of presentations was quasi-random, with the following constraints: After each criterion there was .5 probability of either the same or a different probe, and a different probe could differ in one, two, or three positions with equal (.33) probability. Subjects were asked to respond as quickly as possible but were also informed that they would be disqualified if they exceeded an error limit (10%). After 6 blocks there was a 5-min rest period.

Imaginal condition. The imaginal condition replicated the perceptual condition, with one major difference. At the beginning of the session the subject was shown the six letters and was taught

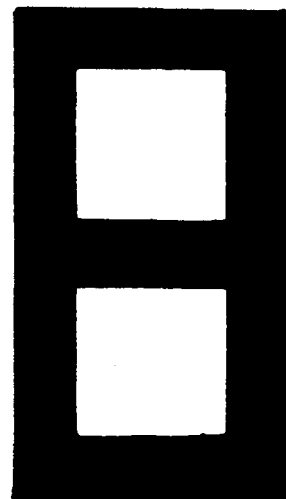


Figure 1. The standard pattern from which all the letter stimuli were constructed (from Taylor, 1976).

	Degree of Difference		
	1	2	3
A	H	E	U
E	F	O	H
O	U	A	F
F	E	H	O
H	A	F	E
U	O	H	A

Figure 2. The letters used in the experiments and the composition of *different* trials (from Taylor, 1976). A probe could differ from the criterion in one, two, or three positions. For example, when A was the criterion, an H probe differed in one position, an E probe in two, and a U probe in three.

to associate a symbol (a comma, an asterisk, a back slash, etc.) with each of them. The learning phase lasted until he/she had reached a criterion of less than four errors in a block of 36 trials. Subjects also learned to form as accurate an image as possible of the proper letter upon onset of the symbol (which occurred in place of the first [criterion] letter in the presentation sequence). The image had to be centered where the fixation mark had been shown. When the imagined criterion letter had reached a satisfactory degree of completeness and vividness, the subject was to depress a foot pedal to start a 500-msec interval, followed by a 500-msec presentation of the probe.

Results

A two-way analysis of variance was carried out on mean correct reaction times (RTs; see Figure 3). The within-subjects factors were condition (perceptual or imaginal)

and type of response (*same* or *different*). The main effect of type of response and the interaction were both significant [$F(1,15) = 65.22$, $MSe = 1,162.97$, $p < .001$, and $F(1,15) = 13.18$, $MSe = 641.38$, $p < .005$, respectively]. *Same* responses were faster than *different* responses (413 vs. 462 msec), irrespective of whether the criterion was imagined (414 vs. 479 msec) or actually presented (412 vs. 444 msec). A set of pairwise comparisons with the Newman-Keuls method showed that the FSE was reliable for both conditions but was larger (65 vs. 32 msec) when the criterion was imagined (all $ps < .01$).

Different responses were submitted to a second analysis of variance with condition and degree of difference (one, two, or three segments) as within-subjects factors. The only significant source of variability was the main effect of degree of difference [$F(2,30) = 49.67$, $MSe = 267.92$, $p < .001$]. Response latencies were faster when the two letters differed by three segments (449 msec) than by two (459 msec) or one (477 msec). RTs for the perceptual condition were faster overall than those for the imaginal condition (444 vs. 479 msec), but this main effect fell just short of significance [$F(1,15) = 3.46$, $MSe = 17,007.29$, $p < .1$]. Two additional one-way analyses showed that the effect of degree of difference was significant for both the perceptual condition [$F(2,30) = 33.54$, $MSe = 181.56$, $p < .001$] and the imaginal condition [$F(2,30) = 24.53$, $MSe = 302.47$, $p < .001$]. Six tests of the contrasted individual means showed the reliability of every pairwise comparison concerning degree of difference (all $ps < .01$): The speed of *different* responses depended on the number of discrepant positions for both the imaginal condition (496, 475, and 467 msec) and the perceptual condition (459, 442, and 432 msec).

In the perceptual condition the errors were 1.1% for *same* trials and 1.6%, 0.4%, and 1.0% for the three types of *different* trials. In the imaginal condition the corresponding figures were 1.6%, 2.2%, 2.3%, and 1.7%. Two analyses of variance, identical to those for RTs, conducted on the number of errors did not show any significant effect.

Discussion

Before discussing the results, it must be considered that in neither condition was the criterion letter physically present at the time of judgments. The two conditions differed in whether the criterion was physically shown shortly before the probe or whether its image was cued. In fact, the perceptual condition could better be construed as a condition based on a short-term visual representation. However, since previous results (Bagnara, Boles, Simion, & Umiltà, 1982) had shown that the features of the comparison process were the same regardless of whether the criterion was physically present or kept in a visual store, it does not seem inappropriate to use the term *perceptual condition* here.

Irrespective of whether the criterion letter was imagined or perceived, there was a clear FSE, and response la-

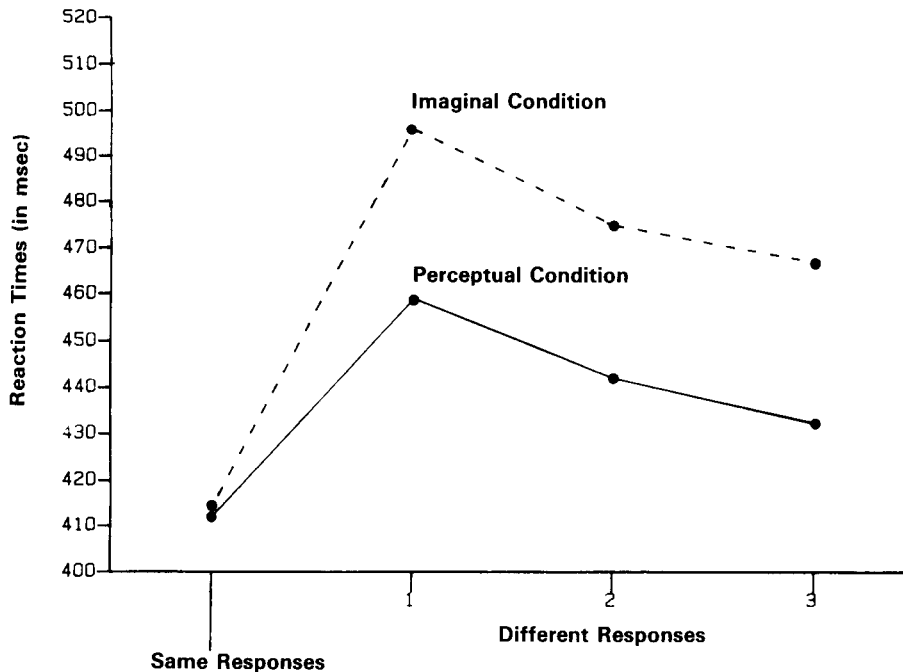


Figure 3. Mean RTs of Experiment 1 plotted as a function of condition (perceptual and imaginal), type of response (*same* or *different*), and degree of difference (for *different* responses only; one, two, or three segments).

tency for *different* decisions depended on the number of discrepant positions (the *segment effect*, or SE). Moreover, *same* responses were faster than the fastest *different* responses. This pattern of results is compatible with a dual-processor model. It seems that the distinction between processes based on the analysis of the component elements and those based on the matching of global configurations, which was proposed by Bamber (1969) for the comparison of real objects, holds true also for internally generated images. Admittedly, however, the above results are not incompatible with other models (see, e.g., Farrell, 1985; Krueger, 1978, 1983, 1984; Proctor, 1981; Proctor & Rao, 1983).

In any event, what matters for the present study is that, whatever the model accepted to explain the FSE, the results clearly showed that the comparison process was very similar for images and percepts, and that the structural details of the images played a role in the comparison process. In other words, it appears that functionally equivalent operations took place when either the perceptual representation or the image of the criterion letter was used in the match. On the other hand, the fact that the FSE was larger and response latency for *different* decisions was slower when the criterion was imagined seems to point to differences between imagery and perception. It could be that it took longer to achieve a fine-grained representation when the criterion letter had to be generated internally, and that this selectively slowed down the responses based on the details of that representation.

It is worth noting that similar results were obtained by Tversky (1969) in a study of name and face matching.

She attempted to manipulate the encoding modality, either pictorial or verbal, of visual stimuli by varying the subjects' expectations about the format of the second of two stimuli to be compared. Her results showed that the match was always in the code of the second stimulus, regardless of whether the stimulus was in the expected format. Thus, subjects apparently generated a visual representation of the face from the stored name when the first stimulus was verbal and the second visual. In contrast to the present experiment, Tversky's subjects were not explicitly instructed to generate images. Therefore, her results probably reflect transformations involving visual codes, rather than consciously generated mental images.¹

Tversky (1969) found what we have termed FSE and SE, but with the former occurring reliably only when the second stimulus was in the expected format. The FSE was of the same magnitude irrespective of whether or not a transformation was required. In our experiment, however, the FSE was larger in the imaginal condition (i.e., the condition corresponding to Tversky's transformation trials) than in the perceptual condition (i.e., the condition corresponding to Tversky's nontransformation trials). This was because *different* responses were much faster when the criterion was actually presented than when it was imagined. This might be attributed to the fact that a better representation of the criterion was available in the former case. It is possible that, when the format of the second stimulus matched subjects' expectations, the visual representation was better in Tversky's than in our experiment. The point of most importance, however, is that Tversky's results are similar to ours in showing both an

FSE and an SE for conditions in which pictorial information was generated, as well as for perceptual conditions.

EXPERIMENT 2

Although the apparent similarities between images and percepts in Experiment 1 suggest functionally equivalent comparison operations, an alternative interpretation is possible. If one admits that the demand characteristics of the task can bring about spurious imagery effects (Banks, 1981; Finke, 1985; Kolars, 1983; Pylyshyn, 1981; respondents to Kosslyn et al., 1979), then the above results become questionable as evidence in favor of the use of mental images in the comparison process. A way to rule out a nonimaginal interpretation of the results of Experiment 1 would be to demonstrate that different results can be obtained when the use of mental images is prevented. This was attempted in Experiments 2 and 3.

Assuming that mental images are generated intentionally and that this process takes time (Kosslyn, 1980; Kosslyn et al., 1983; Shepard, 1975), it would seem that their use could be prevented if subjects were put under time pressure and were not explicitly asked to use imagery. Accordingly, in Experiment 2 we employed a procedure very similar to that of the previous imaginal condition, except for the fact that much stricter time constraints were introduced and subjects were not told to generate visual images.

Method

Subjects. Fourteen new students (7 of each sex) from the Università di Padova subject pool were paid for their participation.

Apparatus. The apparatus was that described before.

Stimuli. The letters and symbols of Experiment 1 were used.

Procedure. The procedure was similar to that of the imaginal condition of Experiment 1, except for the timing of the stimuli. The symbol indicating the criterion letter was shown for 100 msec and was followed by a 100-msec ISI and by a 500-msec presentation of the probe letter. Each session comprised five experimental blocks of 36 trials. During the practice session the subjects did not learn to form images and the experimenter carefully avoided any mention of an imagery strategy. Subjects were simply told that the symbols would "remind" them of the corresponding letters and, of course, were not instructed to depress the pedal to signal that a good image had been formed.

Results

Two one-way within-subjects analyses of variance were performed on correct RTs (see Figure 4), one with type of response and the other with degree of difference as the only factor. As shown by the significance of the type of response main effect, *same* responses were faster than *different* responses (523 vs. 599 msec) [$F(1,13) = 27.28$, $MSe = 1479.11$, $p < .001$]. The main effect of degree of difference was also significant [$F(2,26) = 4.48$, $MSe = 1,410.26$, $p < .025$], and the speed of *different* decisions showed the same trend observed in Experiment 1 (i.e., 619, 602, and 577 msec for one, two, and three discrepant positions, respectively). Pairwise comparisons were performed on *different* responses, and the difference between one segment and three segments was significant ($p < .01$), whereas that between two and three segments fell not far from significance ($p < .1$).

Errors were distributed as follows: 5.0% for *same* trials and 4.3%, 6.4%, and 2.4% for *different* trials when the difference was of one, two, and three segments, respec-

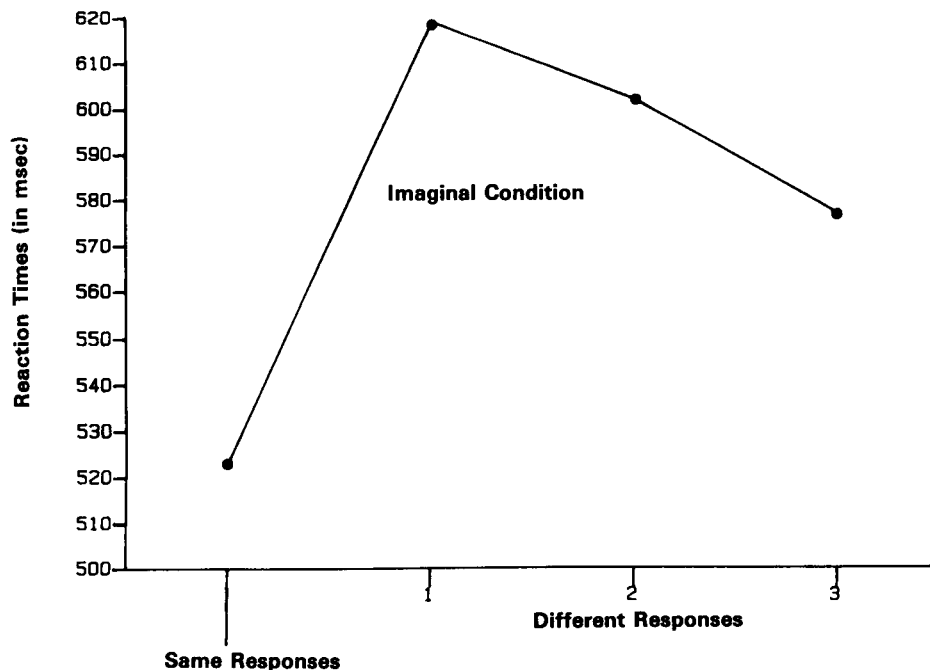


Figure 4. Mean RTs of Experiment 2, plotted as in Figure 3.

tively. Errors were submitted to the two analyses of variance that were performed for RTs, and the main effect of degree of difference was significant [$F(2,26) = 3.50$, $MSe = 1.47$, $p < .05$].

Discussion

Apparently, the results replicated those obtained in the imaginal condition of Experiment 1 and were at odds with our predictions. The FSE was nearly identical to that found in Experiment 1 (76 vs. 65 msec) and the SE showed exactly the same trend.

We had expected that, since no instructions were given to form an image of the criterion letter and the timing of the sequence allowed very little time to adopt an imagery strategy, subjects would not spontaneously generate mental images. In fact, when questioned after completing the experiment, they claimed to have used the name of the criterion letter for comparison with that of the probe and denied that they had formed and used visual images. In spite of that, their performance replicated that of subjects in the imaginal condition of Experiment 1, in which subjects were explicitly instructed to form and use visual images.

Even though it is well known that such introspective reports are questionable, because there is a dissociation between task performance and associated verbalizable knowledge (Berry & Broadbent, 1984), one could legitimately argue that imagery was not involved in either experiment. This crucial point will be dealt with later; for the time being, we will assume that visual images were used in both experiments.

If subjects were able to generate suitable images in the present experiment, then the question arises of how they found the time to do so. The answer might be in the long RTs found in Experiment 2 (i.e., longer than those of the imaginal condition of Experiment 1, 561 vs. 442 msec, overall). It is possible that, although in Experiment 1 the generation process was completed before the subject signaled his/her readiness to see the probe, in Experiment 2 the subject used some extra time to generate the image. In other words, we propose that the subject went on generating the image of the criterion during the ISI and possibly also after the onset of the probe.

Even allowing this extra time, however, it is apparent that images were produced very rapidly, perhaps in a time shorter than the limit of about 500 msec suggested by Shepard (1975). In fact, as observed by Kosslyn et al. (1983), the time needed to form an image can vary and can be very short when its overall shape and component units are simple. In addition, considering the time constraints of the present experiment, the absence of imagery instructions, and the negative subjective reports of imagery, it seems that not only was the use of mental images automatic (see Corballis, 1986), but their generation could also be achieved automatically. Another explanation could be found in the distinction between visual images and visual codes proposed by Posner (1978; see

Note 1): the time needed to access visual codes is much shorter than that needed to generate visual images. Therefore, one could suggest that visual representations were used in both Experiment 1 and Experiment 2, but in the former the representation was a visual image, whereas in the latter it was a visual code.

Finally, these results are interesting also from another point of view: If we assume that visual images were used, they allow us to rule out the possibility that the outcome of the imaginal condition of Experiment 1 was simply due to carryover effects from the perceptual condition. This is because in Experiment 2 the subjects were run only under the imaginal condition. However, as we have already pointed out, the crucial assumption that visual images were used can be questioned exactly because Experiment 2 replicated the results of Experiment 1, even though the experimental manipulations had rendered the use of mental images unlikely. What was needed at this point was a set of experimental manipulations that, by preventing the use of mental images, would yield a different pattern of results.

EXPERIMENT 3

Experiments 1 and 2 clearly showed both the FSE and the SE. Only the SE can be attributed with certainty to the use of a visual representation of the criterion letter, because the FSE is a pervasive effect found in almost every *same-different* paradigm, regardless of the format of the stimulus representation (Farell, 1985; Proctor, 1981). Since the manipulations employed in the present experiment were intended to prevent the use of visual images, their effectiveness should become manifest through the elimination of the SE.

Three conditions were chosen to produce interference with the process of visualization. In the first (the reading condition), subjects were required to read a text after presentation of the symbol and before the probe letter appeared. We reasoned, in accordance with Brooks (1967, 1968, 1970), that reading should suppress visualization, thus rendering the use of visual images unlikely. In the second condition (the masking condition), a pattern was shown in the interval between presentation of the symbol and presentation of the probe letter. Reeves (1980) showed that a visual image can mask a visual pattern, and it would seem likely that the reverse can also happen. If this occurs, the mask should prevent the use of visual images. In the third condition (the delay condition), a long interval elapsed between the symbol and the probe letter. If the maintenance of a visual image requires effort (Cooper & Shepard, 1973; Posner, 1978), subjects not instructed to do so should not employ such a demanding and unnecessary strategy.

Method

Subjects. Twenty-four new subjects (12 of each sex) from the Università di Padova subject pool took part in the experiment. They

were paid for their participation. Each subject participated in only one condition; thus, there were 8 subjects in the reading condition, 8 in the masking condition, and 8 in the delay condition.

Apparatus. A remote eye-movement monitoring system was added to the apparatus employed in Experiment 2 to make sure that the subject kept a good fixation when required to do so.

Stimuli. Besides the usual letters and symbols, in the reading condition a text that occupied the entire screen was presented. The length of the text was such that no subject could read it aloud in less than 20 sec. In the masking condition the standard pattern (see Figure 1) was also used. Four dots delimited an area, centered on the fixation mark, where the probe letter was to be presented.

Procedure. The basic procedure was identical to that of Experiment 2. The symbol indicating the criterion letter was presented for 100 msec and was followed by an ISI that varied according to the experimental condition. The probe letter was then presented for 500 msec. It was the ISI that differentiated the three conditions.

Reading. In the reading condition, the ISI lasted 15 sec, during which time the text was shown and the subject was required to read it aloud at a normal speed but without pausing. Subjects were told that they would later be questioned about the content of the text, which varied from trial to trial. One of the experimenters monitored the speed and accuracy of reading.

Masking. In the masking condition, the standard pattern was shown 100 msec after the disappearance of the symbol. It stayed on for 300 msec and was followed by a 100 msec blank period. Subjects were required not to move or close their eyes during the ISI; if an eye movement was detected, that trial was discarded and replaced.

Delay. The delay condition was identical to the reading condition except for the fact that no text was presented during the 15-sec ISI.

Results

This experiment was intended to show that the SE disappears when visual images cannot be utilized. Given the risks involved in the acceptance of the null hypothesis,

it was thought that the safest way for analyzing the data was through overall analyses of variance that comprised all three experimental conditions.

First, the mean correct RTs were submitted to a two-way analysis of variance in which condition (reading, masking, or delay) was a between-subjects factor, whereas type of response (*same* vs. *different*) was a within-subjects factor. The two main effects were significant [$F(2,21) = 9.73$, $MSe = 11,661.55$, $p < .005$, and $F(1,21) = 58.19$, $MSe = 599.47$, $p < .001$, for condition and type of response, respectively]. Overall, RT was faster for the masking condition (467 msec) than for the other two conditions (611 and 613 msec for reading and delay, respectively). *Same* responses were 53 msec faster than *different* responses (537 vs. 590 msec).

The same analysis on errors showed only a significant effect of condition [$F(2,21) = 4.58$, $MSe = 16.92$, $p < .025$]. Errors were more numerous in the reading condition (8.2%) than in either the masking (4.8%) or delay (3.5%) condition.

Of greater interest for the purpose of this experiment was the two-way analysis of variance carried out on *different* RTs only. In it the between-subjects factor was again condition, whereas the within-subjects factor was degree of difference (one, two, or three segments). The only significant source was condition [$F(2,21) = 10.06$, $MSe = 18,059.22$, $p < .005$], which confirmed that masking (490 msec) was faster than either reading (638 msec) or delay (644 msec). The most important outcome of this analysis, however, was the nonsignificance of either the main effect of degree of difference or the interaction ($F_s = 0.98$ and 0.93 , respectively). As shown in Figure 5 and Table 1, the relationship between degree of differ-

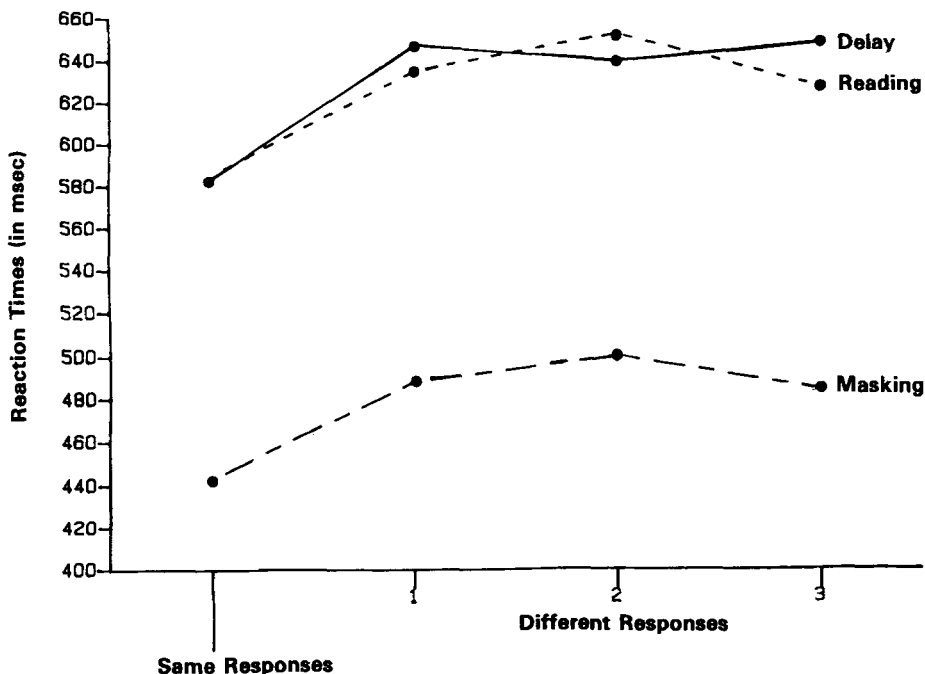


Figure 5. Mean RTs of the three experimental conditions of Experiment 3, plotted as in Figure 3.

Table 1
Mean Reaction Times (RTs, in msec) and Percentages of Errors (%)
for Different Responses as a Function of Experimental Condition
and Degree of Difference (One, Two, or Three Segments)
in Experiment 3

Experimental Condition	Degree of Difference							
	1		2		3		M	
	RT	%	RT	%	RT	%	RT	%
Reading	635	7.5	652	6.7	627	6.0	638	6.7
Masking	488	3.7	500	7.5	484	2.1	490	4.4
Delay	647	2.9	639	3.3	648	2.9	644	3.3
M	590	4.7	597	5.8	586	3.7	591	4.7

ence and speed of response, which was apparent in the previous two experiments, was no longer present.

The same analysis was carried out on errors and did not show any significant source. The F values for the main effect of degree of difference and the interaction were 2.00 and 1.51, respectively. Percentages of errors are shown in Table 1.

Before concluding that the SE was present in Experiment 2 and absent in Experiment 3, a further analysis is needed to compare the effect of degree of difference in the two experiments. This was done by a two-way analysis in which the between-subjects factor was experiment (Experiment 2 and Experiment 3) and the within-subjects factor was degree of difference. The main effect of degree of difference was significant [$F(2,72) = 3.83$, $MSe = 953.75$, $p < .025$]. Of paramount importance was the significance of the interaction [$F(2,72) = 3.53$, $MSe = 953.75$, $p < .05$], which confirmed that the number of segments had an effect on RTs in Experiment 2 only.

Discussion

There can be little doubt that the experimental manipulations were effective in eliminating the SE. We had reasoned that, if the manipulations that were presumed to prevent the use of visual images were effective, then the SE, which was considered to be diagnostic of visualization, should disappear. That was exactly what happened. Therefore, it seems reasonable to maintain that in the imaginal condition of Experiment 1 and in Experiment 2, subjects generated a visual image of the criterion letter and matched it with the physically present probe letter. By contrast, in Experiment 3, either the generation or the use of the visual image of the criterion letter was prevented, and thus no visual representation of it was available at the moment of the comparison.

In the reading and masking conditions, it would seem likely that the visual interference either interrupted the generation of the image or destroyed an already formed one. In the delay condition, the subject could have automatically generated a visual image, as suggested by the outcome of Experiment 2, but then, due to the length of the ISI and the absence of imagery instructions, failed to allocate enough resources to its maintenance, so that the image deteriorated.

One could wonder about the type of internal representation of the criterion letter that was matched to the probe.

A likely candidate is the "name" of the letter, which has been shown to be formed in parallel with its visual representation (Posner, 1978). That would be in agreement with the subjects' verbal reports. However, as already pointed out, such introspections are not reliable. In fact, the subjects were convinced that they had also used a verbal code in Experiment 2, where they had apparently made use of imagery.

The fact that response latency was much faster in the masking condition than in the other two conditions could make one think that a different format of the criterion letter was used in the former. It is difficult to figure out what format, apart from a name or a visual code, could have been produced, and, at any rate, it should not be overlooked that the ISI was 30 times longer in the two conditions that yielded the slowest RTs. A likely possibility, therefore, is that in the reading and delay conditions the subject was unable to maintain a good level of readiness for such a long period.

CONCLUSION

The findings of the present study can be summarized as follows: (1) When the subject is instructed to form a complete visual mental image, the internally generated image possesses structural properties detailed enough to allow a comparison process functionally equivalent to that that takes place when a visual perceptual representation is used. (2) It simply takes longer for the details to become available when images, rather than percepts, are used in the match. (3) Visual images can be generated automatically and very rapidly. (4) The generation of visual images can be disrupted, or their use prevented, if the visual system is engaged in an interfering task or when a visual mask is presented. (5) Visual images decay over time if no effort is made to refresh and maintain them.

These results support the equivalence hypothesis, which maintains that visual mental images and visual perceptual representations are equivalent for both the properties they possess and the operations they undergo. It is true that in the case of images it took longer to achieve the details of the representation than its overall configuration; however, in spite of this difference in the availability of global or local information, the matching process did not substantially differ between images and percepts. Whichever the type of representation, the SE, which indicates visualization, was always present and had more or less the same magnitude, unless, of course, the use of imagery was prevented.

REFERENCES

- BAGNARA, S., BOLES, D. B., SIMION, F., & UMILTÀ, C. (1982). Can an analytic/holistic dichotomy explain hemispheric asymmetries? *Cortex*, *18*, 67-78.
- BAGNARA, S., BOLES, D. B., SIMION, F., & UMILTÀ, C. (1983). Symmetry and similarity effects in the comparison of visual patterns. *Perception & Psychophysics*, *34*, 578-584.
- BAMBER, D. (1969). Reaction times and errors for same-different judg-

- ments of multidimensional stimuli. *Perception & Psychophysics*, *6*, 169-174.
- BANKS, W. P. (1981). Assessing relations between imagery and perception. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 844-847.
- BERBAUM, K., & CHUNG, C. P. (1981). Müller-Lyer illusion induced by imagination. *Journal of Mental Imagery*, *5*, 125-128.
- BERRY, D., & BROADBENT, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, *36A*, 209-231.
- BROOKS, L. R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, *19*, 289-299.
- BROOKS, L. R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, *22*, 349-368.
- BROOKS, L. R. (1970). An extension of the conflict between visualization and reading. *Quarterly Journal of Experimental Psychology*, *22*, 91-96.
- COOPER, L. A., & SHEPARD, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75-176). New York: Academic Press.
- CORBALLIS, M. C. (1986). Is mental rotation controlled or automatic? *Memory & Cognition*, *14*, 124-128.
- FARELL, B. (1985). "Same"- "different" judgments: A review of current controversies in perceptual comparisons. *Psychological Bulletin*, *98*, 419-456.
- FINKE, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review*, *87*, 113-132.
- FINKE, R. A. (1985). Theories relating mental imagery to perception. *Psychological Bulletin*, *98*, 236-259.
- FINKE, R. A., & KURTZMAN, H. S. (1981). Area and contrast effects upon perceptual and imaginal acuity. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 825-832.
- FINKE, R. A., & PINKER, S. (1982). Spontaneous imagery scanning in mental extrapolation. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 142-147.
- FINKE, R. A., & PINKER, S. (1983). Directional scanning of remembered visual patterns. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 126-132.
- INTONS-PETERSON, M. J. (1981). Experimenter naiveté and imaginal judgments. *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 833-843.
- INTONS-PETERSON, M. J. (1983). Imagery paradigms: How vulnerable are they to experimenters' expectations? *Journal of Experimental Psychology: Human Perception & Performance*, *7*, 394-412.
- KEENAN, J. M. (1983). Qualifications and clarifications of images of concealed objects: A reply to Kerr and Neisser. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 222-230.
- KOLERS, P. A. (1983). Perception and representation. *Annual Review of Psychology*, *34*, 129-166.
- KOSSLYN, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- KOSSLYN, S. M., & ALPER, S. N. (1977). On the pictorial properties of visual images: Effects of image size on memory for words. *Canadian Journal of Psychology*, *31*, 32-40.
- KOSSLYN, S. M., BALL, T. M., & REISER, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception & Performance*, *4*, 47-60.
- KOSSLYN, S. M., PINKER, S., SMITH, G. E., & SHWARTZ, S. P. (1979). On the demystification of mental imagery. *Behavioral & Brain Sciences*, *2*, 535-581.
- KOSSLYN, S. M., REISER, B. J., FARAH, M. J., & FLEGEL, S. L. (1983). Generating visual images: Units and relations. *Journal of Experimental Psychology: General*, *112*, 278-303.
- KRUEGER, L. E. (1978). A theory of perceptual matching. *Psychological Review*, *85*, 278-304.
- KRUEGER, L. E. (1983). Probing Proctor's priming principle: The effect of simultaneous and sequential presentation on same-different judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 511-523.
- KRUEGER, L. E. (1984). Self-termination in same-different judgments: Multiletter comparison with simultaneous and sequential presentation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 271-284.
- NICKERSON, R. S. (1978). On the time it takes to tell things apart. In J. Requin (Ed.), *Attention and performance VII* (pp. 77-88). Hillsdale, NJ: Erlbaum.
- PINKER, S. (1984). Visual cognition: An introduction. *Cognition*, *18*, 1-63.
- PINKER, S., CHOATE, P. A., & FINKE, R. A. (1984). Mental extrapolation in patterns constructed from memory. *Memory & Cognition*, *14*, 207-218.
- PODGORNY, P., & SHEPARD, R. N. (1978). Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception & Performance*, *4*, 21-35.
- POSNER, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.
- PROCTOR, R. W. (1981). A unified theory for matching-task phenomena. *Psychological Review*, *88*, 291-326.
- PROCTOR, R. W., & RAO, K. V. (1983). Reinstating the original principles of Proctor's unified theory for matching-task phenomena: An evaluation of Krueger and Shapiro's reformulation. *Psychological Review*, *90*, 21-37.
- PYLISHYN, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*, 16-45.
- REEVES, A. (1980). Visual imagery in backward masking. *Perception & Psychophysics*, *28*, 118-124.
- SHEPARD, R. N. (1975). Form, formation, and transformation of internal representations. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 87-122). Hillsdale, NJ: Erlbaum.
- SHEPARD, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*, 417-447.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- SHEPARD, R. N., & METZLER, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701-703.
- SHEPARD, R. N., & PODGORNY, P. (1978). Cognitive processes that resemble perceptual processes. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 5. Human information processing* (pp. 189-237). Hillsdale, NJ: Erlbaum.
- TAYLOR, D. A. (1976). Holistic and analytic processes in the comparison of letters. *Perception & Psychophysics*, *20*, 187-190.
- TVERSKY, B. (1969). Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics*, *6*, 225-233.
- WALLACE, B. (1984). Apparent equivalence between perception and imagery in the production of various visual illusions. *Memory & Cognition*, *12*, 156-162.

NOTE

1. The importance of distinguishing between a consciously generated visual image and an automatically generated visual code was aptly stressed by Posner (1978). It takes time to generate visual images and this process gives rise to strong subjective reports of imagery. By contrast, visual codes can be contacted very rapidly and do not necessarily give rise to the subjective reports typical of visual images.

(Manuscript received June 4, 1985;
revision accepted for publication October 2, 1987.)