Levels of coding in picture-picture comparison tasks

D. J. BARTRAM

University of Hull, Hull, England HU6 7RX

Two experiments are reported in which subjects had to match pairs of pictures of objects. "Same" pairs could be either identical (Ps), pictures of different views of the same object (Pv), or pictures of different objects having the same name (Pd). With line drawings as stimuli, RTs for Condition Ps were shorter than for Condition Pv, which in turn were shorter than for Condition Pd. Visual similarity had no effect on Pd RTs. However, in Experiment II, where photographs of objects with high-frequency (HF) and low-frequency (LF) names were used, no difference was found between Conditions Ps(HF), Ps(LF) and Condition Pv(HF); and no difference occurred between Conditions Pd(HF), Pd(LF) and Condition Pv(LF), the latter set of conditions being associated with longer RTs than the former. This pattern of results was found with both a .25-sec and a 2-sec ISI. The results are discussed in terms of the levels of coding involved in processing information from picture stimuli. It is concluded that at least two levels are involved in matching photographs of real objects (an object-code level and a nonvisual semantic code level), while a third level may be used in matching tasks involving stylized line drawings (a picture-code level).

Several studies (e.g., Posner, Boies, Eichelman, & Taylor, 1969; Posner & Mitchell, 1967) have shown that a pair of identical letters can be matched more rapidly than a pair of different letters which have the same name. Posner and Mitchell (1967) hypothesized that these two types of matching are based on different underlying stimulus representations. The "identity" match (e.g., AA) is based on a comparison between visual codes, while the "name" match (e.g., Aa) is based on a comparison between verbal codes. Posner and Warren (1972) have suggested an extension of this "dual-code" model to cover studies involving comparisons between picture stimuli. They refer to a study by Frost (1972), who presented line drawings of objects to subjects, some of whom were led to expect a recall task, while others were led to expect a recognition task. After a 15-min delay, both groups were given a set of drawings and told to classify each one according to whether it matched one of the objects seen before. Some of the drawings were identical to those seen during training, others showed the same object from a different viewpoint, and the rest were from different object classes.¹ For the "recognition-set" subjects, RTs were shorter for the "identity match" condition than for the condition involving different views. Posner and Warren (1972) argue that this result supports Posner's hypothesis, implying that comparisons of different views of the same object are verbally mediated. In addition, Klatzky (1972) has shown that "same" judgments for

pairs of identical pictures of objects are faster than for pairs of different objects from the same object class. Thus, if only a "visual" code and a "name" code are available, different objects having the same name and different views of the same object would both have to be compared at the "name" code level. This seems unlikely.

A recent paper by Klatzky and Stoy (1974) suggests that the situation is rather more complex than the simple dual-code model implies. In their first experiment on picture-picture matching, instead of using only two conditions (identity and name match), they introduced a mirror-image match condition. In addition, they looked at the differences between identical, mirror-image, and name match RTs with two sets of line-drawing stimuli. In one set (S) name match pairs of objects were physically similar in appearance, while in the other (D) they were dissimilar. They found virtually no difference between RTs for identity and mirror-image matches, for either set of stimuli. Name match RTs were slower than both of the other conditions. Similarity appeared to have no effect on name match RTs, but both mirror-image and identity matches were faster for the S set of stimuli. From these results, they argued that both mirror and identity match conditions involve the same visual codes, with name matches being mediated either by verbal labels or by some more abstract "visual" code.

Their second experiment (Klatzky & Stoy, 1974) compared the two sets of stimuli in "pure" conditions (all identity matches or all name matches) and in a "mixed" condition (in which both types of match occurred). This experiment showed that a difference between the two sets of stimuli (S vs. D) only occurs in the "mixed" condition. Furthermore, they found that

This research was partly carried out in the Laboratory of Experimental Psychology at the University of Sussex while the author was in receipt of a Science Research Council research studentship. Requests for reprints should be sent to D. Bartram, Department of Psychology, University of Hull, Hull, England HU6 7RX.

this difference applied to both identity and name match conditions. This suggests that the differences between stimulus sets in the mixed conditions, in this and their preceding experiment, are attributable to effects of visual similarity, and not simply a function of differences between the sets with respect to the time taken to perceive or encode the stimuli. The effect of similarity on name match comparisons supports their argument that name matching can involve a visual code of some form. The interaction between ISI and the RT difference between identity and name match conditions suggests that this "abstract" visual code is different from the visual code which mediates identity matching. They conclude by arguing that two types of visual codes are involved in the comparisons they studied: (1) a relatively veridical visual code which mediates both identity and mirror-image comparisons, and (2) a more abstract visual code which may only mediate name match comparisons when the mirror-image match condition is not present. From the data obtained about the effects of ISI, it appears that the original "veridical" code comes to resemble the more "abstract" code over time.

Bartram (1974) presents some data on object naming which suggests a slightly different way of looking at the above results. In this study, subjects named photographs of objects over a number of blocks of practice. For a given object, pictures could either be identical (Ps), different views of the same object (Pv), or different objects having the same name (Pd). Effects of practice were found under all three conditions, with the largest effect occurring for Condition Ps and the smallest for Pd. However, the decrease in naming latencies found with practice on Condition Pd could not be explained simply in terms of "response learning" (i.e., of practice at the name code level): It was found that, if subjects were trained on either Condition Ps or Pv and then transferred to Pd (the set of object names remaining the same), naming latencies increased to near the level found on the initial block of training trials. From this, it was argued that a form of coding exists which mediates the mapping of visual codes onto name codes. It was suggested that this code (which is not itself a verbal code) represents some form of abstract description of the semantic properties of the object class concerned. In Condition Pd, subjects had practice in forming this code every time they named an object. However, in Conditions Ps and Pv, the similarity of the visual codes formed enabled subjects to bypass description formation processes at this nonverbal "semantic" level after the initial block of trials. It is this lack of practice at the semantic level which, it was argued, resulted in the lack of transfer from Conditions Ps and Pv to Pd. The difference between Ps and Pv practice effects could be accounted for in a number of ways. One suggestion was that two interdependent visual codes coexist: a "picture code" (which represents the

stimulus in terms of lines and regions of varying textures and grayness, and which may involve decisions concerning basic figure-ground distinctions) and an "object code" (which is a redescription of the picture code in which regions are redefined as surfaces, inferences are made about hidden features and surfaces, depth relationships conveyed by overlap, foreshortening, and other cues are taken into account to produce a three-dimensional description of the scene portrayed by the stimulus picture). The evidence suggested that the process of recoding the picture code into a code representing that object which the stimulus picture represented was an obligatory one.

As Klatzky and Stoy (1974) rightly point out, a truly isomorphic picture code would result in identity match RTs being no faster than mirror-image match RTs. Either the picture code would have to undergo lateral inversion or mirror images would have to be compared at the object code level (i.e., as if they were two different views of the same object). In both cases, one would expect a difference in RTs to be found. The fact that there is no difference could be accounted for if the assumption about isomorphism is partially relaxed. It could be argued that a mirror-image match is a special case of picture code matching, in that a mirror image retains exactly the same set of spatial relationships between lines, regions, and parts of a picture, the only difference being that all relationships are laterally inverted.

The two experiments reported below attempt to clarify and extend some of the findings reported by Bartram (1974) and Klatzky and Stoy (1974), as well as provide more data on the general question of what coding levels are involved in picture comparisons. Experiment I examines four matching conditions involving pairs of identical pictures (Condition Ps), pairs of different views of the same object (Condition Pv), and pairs of different objects having the same name seen from either the same viewpoint (Condition Pds) or different viewpoints (Condition Pdv). On the basis of the model proposed in Bartram (1974), it was expected that comparison RTs would be fastest for identical pictures (picture code comparisons), longer for different views of the same objects (object code comparisons), and longer still for different objects having the same name (semantic or name code comparisons).

The fact that Klatzky and Stoy (1974) did not obtain an effect of stimulus similarity on name matching in their first experiment, but did in their second, may be accounted for by arguing that the mirror-image match condition in their first experiment somehow interfered with the process of generating abstract visual images. Hence, it might be expected that a situation where both Ps and Pv matches occur may be analogous to Klatzky and Stoy's (1974) experiment involving both identity and mirror-image comparisons in that it will not be possible for subjects to generate abstract visual

codes to mediate name matches (i.e., Conditions Pds and Pdv). If so, then regardless of whether name matches are mediated by verbal or nonverbal semantic codes, there should be no difference in RTs between Conditions Pds and Pdv.

EXPERIMENT I

Method

Subjects. Eight subjects, six undergraduates and two postgraduates, took part and were tested in individual 40-min sessions. Four of the subjects were studying psychology, though none of them knew the purpose of this experiment. All subjects were paid for their participation.

Materials. Simple line drawings of objects were used. For each object class, two objects were chosen, and for each object, two different stimuli were drawn. One showed the object from a viewpoint which clearly showed its shape in perspective (p), and the other was of a characteristic front or side elevation (e). All the perspective drawings were three-quarter views looking down from an elevation of about 30 deg above the horizontal. Thus, a total of 32 pictures were prepared, four from each of the following eight object classes: HOUSE, FACTORY, CHURCH, BRIDGE, TELEVISION, CHAIR, TABLE, WINDOW. The drawings were in black ink on white 15 x 10 cm cards, and occupied a 4 x 4 cm square in the center of the card. Drawings of two additional objects (CUBE and PYRAMID), an elevation and perspective view of each, were used as practice stimuli. A fixation stimulus was drawn consisting of a white 15 x 10 cm card with a black cross at the center surrounded by a 5-cm square. Examples of the stimuli are shown in Figure 1.

Apparatus. A three-field Electronic Developments tachistoscope was used. Field 1 contained the fixation stimulus, Field 2 the first stimulus of each pair, and Field 3 the second stimulus. A Dawe digital timer was started by the onset of the second stimulus and stopped, using a voice key, by the subject's response. Each trial was initiated by the subject pressing a telegraph key.

Experimental design. Four copies of each of the four practice stimuli were used to produce eight pairs of practice stimuli (four "same" and four "different" pairs). Eight copies of each of the 32 test stimuli were produced, giving a total of 128 test pairs. The pairings were carried out to produce two blocks of 64 pairs each, such that within each block there were 32 "same" and 32 "different" pairs. The 32 "same" pairs in each block were equally divided into pairs for each of the following four "same" conditions.

Condition Ps. Same object seen from the same viewpoint: either both elevations, Ps(ee), or both perspective views, Ps(pp).

Condition Pv. Same object seen from different viewpoints: either the first an elevation view and the second a perspective one, Pv(ep), or vice versa, Pv(pe).

Condition Pds. Different objects from the same object class seen from the same type of viewpoint: either both elevations, Pds(ee), or both perspective views, Pds(pp).

Condition Pdv. Different objects from the same object class seen from different types of viewpoint: either the first an elevation and the second a perspective view, Pdv(ep), or vice versa, Pdv(pe). For Conditions Ps and Pds, half were ee pairs and half were pp pairs, while for Conditions Pv and Pdv, half were ep pairs and half were pe pairs. The 32 "different" pairs were divided into four equal sets of ee, pp, ep, pe pairs. Within each block, pairs were randomly ordered, with the constraint that no more than three "same" or three "different" pairs could occur consecutively. The two blocks were each divided into two equal subblocks. Each subject responded to all 128 trials, the order of blocks and of subblocks within blocks being counter-

Figure 1. Examples of stimulus conditions used in Experiment I.

balanced across subjects. To control for within-subblock practice effects, the trials within each subblock were presented to half the subjects in one order and to the rest in the reverse order.

Procedure. The sequence of events on each trial was as follows. Following the instruction "press," the subject pressed the telegraph key, and, after a .5-sec delay, the fixation stimulus was replaced by the first stimulus. This was exposed for .5 sec and appeared centrally located within the area surrounded by the square of the fixation stimulus. The fixation stimulus reappeared during a .5-sec ISI, and was replaced by the second stimulus which appeared in the same position as the first and was exposed for 1.5 sec. Subjects were instructed to say "Yes" if the two stimuli were of objects having the same name and "No" if they had different names. They were told to respond as quickly as possible, but to avoid making errors. Following the eight practice trials, they were told they would receive four blocks each of 32 trials, and were read a list of the eight object-class names. It was made clear that no other objects would occur. They were immediately corrected if they made an error.

Results

Errors. Of the total 1,024 responses, 37 were errors (3.6%), of which 17 were false negative and 20 false positive (see Table 1). Errors were discarded from the results.

Analysis of RT data. Repeated-measures design analyses of variance were carried out: one on both "same" and "different" RTs, and the second just on the "same" RTs. All factors, apart from Subjects, were treated as fixed. As the degrees of freedom for interactions with subjects were small (5), the procedure for pooling within-cell variance estimates and subject interactions, described in Winer (1970, p. 322), was followed. As a result of this preliminary analysis, the

Table 1
Mean RTs (Milliseconds) and Error Rates (Percentage) for "Same" and "Different" Responses (Experiment I)

Type of Viewpoint	Ps		Pv		Pds		Pdv		"Different"	
	RT	error	RT	error	RT	error	RT	error	RT	error
ee	550	1.56			651	4.69			638	2.34
pp	546	3.13			628	3.13			656	5.47
ep			590	3.13			637	4.69	655	6.25
pe			579	1.56			631	4.69	636	1.56
Mean	548	2.35	585	2.35	639	3.91	634	4.69	646	3.91

within-cell sum of squares and all those sums of squares involving interactions with Subjects were pooled to give a single error term. A .25 significance level was used for testing the interactions with Subjects.

"Same" RTs. The results for the "same" conditions are presented in Table 1. The analysis of variance carried out on the "same" response latencies revealed a significant effect of practice (decrease in latencies from the first to the second block of trials) [$F(1,489) = 9.7$, $p < .001$], but this did not interact with any of the other factors ($F < 1$ in all cases). Pairs consisting of pictures of the same object (Conditions Ps and Pv) were responded to faster than pairs of different objects having the same name (Conditions Pds and Pdv) [$F(1,489) = 28$, $p < .001$]. The Type of Object (Ps and Pv vs. Pds and Pdv) by Change in Viewpoint (Ps and Pds vs. Pv and Pdv) interaction just failed to reach significance [$F(1,489) = 2.62$]. However, individual comparisons carried out between the four conditions provide support for the multiple-coding model. Condition Ps latencies were faster than Pv latencies [$F(1,489) = 4$, $p < .01$] and Pdv latencies [$F(1,489) = 6.9$, $p < .01$]. There was no difference between Conditions Pds and Pdv ($F < 1$). Pds(ee) comparisons were 23 msec slower than Pds(pp) comparisons. However, this difference did not reach significance.

"Different" RTs. An analysis of variance revealed that, overall, "same" responses (602 msec) were significantly faster than "different" responses (646 msec) [$F(1,1009) = 21$, $p < .001$]. It can be seen from Table 1 that "different" responses were, on average, as long as or longer than the slowest "same" conditions (Pds and Pdv). Changing the type of viewpoint presented had no effect on "different" responses. The mean for ee and pp-type "different" responses was 647 msec, while for ep and pe "different" responses it was 645 msec. However, "different" responses tended to be faster if the second stimulus was an elevation rather than a perspective view, but this difference was not significant.

The results of Experiment I are consistent with both Klatzky and Stoy's (1974) study and with Bartram's (1974) study on object naming. The lack of difference between Conditions Pds and Pdv implies that these comparisons were not visually mediated. However, it is at present an open question as to whether they

were verbally mediated or mediated by some form of nonverbal semantic code, as suggested in Bartram (1974). The difference between Conditions Ps and Pv is consistent with the argument that the former type of comparison was mediated by a picture code, while the latter involved object coding. Alternative explanations of this difference will be considered after Experiment II has been reported. The finding that "different" response latencies were as long as, or longer than, both Pds and Pdv suggests that, under the present conditions, a mismatch decision cannot be made until the two stimuli have been tested for a match at the highest level of coding.

EXPERIMENT II

Effects of Mode of Representation on Coding Processes

Experiment II was carried out to see whether the differences found between Ps, Pv, and Pd comparisons are limited to experiments using simple stylized line drawings of objects or whether they also occur when photographs of real objects are used. There may well be important differences in the coding mechanisms used to produce internal descriptions of photographs and the highly stylized line drawings typically used in picture matching tasks. Not only do photographs contain more information than stylized line drawings, but also that information is represented in qualitatively different ways. While there is an isomorphic relationship between the way stimulus features and structural relationships are projected onto the retina from photographs, on the one hand, and objects in the real world, on the other hand, stylized line drawings employ a set of artistic conventions to stress important defining features and relationships and to remove "irrelevant" information. Studies comparing the ease with which objects can be identified as a function of mode of representation (e.g., Fraisse & Elkin, 1963; Ryan & Schwartz, 1956) have tended to show a marked advantage for stylized drawings over other forms of representations insofar as identification thresholds are concerned (there being little or no difference between real objects, photographs, and detailed shaded line drawings, and there being a marked disadvantage for simple outline drawings).

Information encoded from photographs may well be treated in the same manner as information encoded from a single glance at a real object. Bartram (1974) found complete transfer from naming identical pictures of an object to naming different views of that same object. This suggests that either the process of recoding from the picture to the object code is obligatory or information from photographs is encoded directly into some form of object representation. Hence, it might be expected that, in a matching task, there will be no difference between comparison times for identical pairs of photographs (Ps) and pairs of photographs consisting of different views of the same object (Pv). In both cases, comparisons would be mediated by some form of object code. Experiment II tested the hypothesis that, contrary to what was found with line drawings (Experiment I), for photographs there will be no difference between Conditions Ps and Pv. In addition, the experiment examines the effects of object-class familiarity and variations in ISI on comparison RTs.

Effects of Familiarity

Previous studies on picture-picture matching have tended to use stylized drawings of common objects as stimuli (Frost, 1972; Klatzky, 1972; Klatzky & Stoy, 1974). Wingfield (1968), however, examined the effect of varying name frequency on matching pairs of drawings of objects (identity match) and matching drawings and names. He found no difference in comparison RTs as a function of frequency. He argues that this supports Oldfield's (1966) hypothesis that the effect of name frequency on object naming RTs arises from differences in the time taken to retrieve common and rare object names from verbal memory, not from differences in "perceptual identification" time. However, whatever the locus of the frequency effect, in a matching task the effect may be absent because of the "priming" effect of the first stimulus in each pair. Given a long enough ISI (and Wingfield used 5 sec), visual or semantic information about the positive object class or the object-class name could be retrieved and, hence, there would be either a much reduced frequency effect or none at all.

There is evidence to suggest that, for object naming, the frequency effect has two components: one related to stimulus variables and affected by practice, and another, more stable, component accounted for in terms of name retrieval differences for common and rare names (Bartram, 1973, 1974). In addition, Seymour (1973) has produced evidence of frequency effects in name-picture and picture-picture matching tasks, using very simple stimuli (square, circle, oblong, rectangle). If the lack of frequency effect in Wingfield's (1968) experiment was due to the priming effect of the first stimulus in each pair, then it is possible that a frequency effect will occur at a short ISI but not at longer ones. The present experiment investigates this possibility by

examining performance at two ISIs: .25 sec and 2 sec. The presence of a frequency effect in Conditions Ps and Pv would strongly suggest that such an effect was a function of differences in the time taken to carry out nonverbal processing on the second stimulus of each pair.

Method

Subjects. Eight psychology undergraduates, five female and three male, were each run in a single 30-min session. None of the subjects knew the purpose of the experiment.

Materials. All the stimuli were black-and-white photographs of objects, prepared as 35-mm slides. Small objects were photographed against plain backgrounds, while large objects (e.g., CAR) were photographed against a natural background which contained no other distracting objects. A set of 20 object-class names was selected from Thorndike and Lorge (1944), 10 with low-frequency names (LF, 15 or less occurrences per million words) and 10 with high-frequency names (HF, 40 or more occurrences per million words). Pictures of objects from 10 of these classes, five LF (TROWEL, TEAPOT, GUITAR, TROMBONE, SAUCEPAN) and five HF (CHICKEN, ARM, FOOT, CUP, PLATE), were used for practice. The remaining five LF (PLIERS, LORRY, CORK-SCREW, SPANNER, SCISSORS) and five HF (CAT, BOOK, CAR, DOG, CHAIR) object classes were used in the main experiment. For each of the 10 test object classes, eight pictures of different viewpoints of one object from that class (for Condition Pv) and eight pictures of different objects from that class (for Condition Pd) were produced. The stimuli for Condition Pv were produced by photographing the same object from eight different viewpoints, with a difference of about 45 deg between each viewpoint (rotation being in the horizontal plane).

Apparatus. Stimuli were back-projected onto a 9 x 6 cm screen from two Kodak Carousel projectors (one for the first and one for the second stimulus in each picture pair). The projected image completely filled the screen and the subject was seated about 150 cm away from it. Each projector was fitted with a solenoid-operated shutter and Polaroid-filter brightness control. The durations of the first stimulus (.5 sec) and the ISI were controlled by timing circuits. Onset of the second stimulus started a Dawe digital timer, and a voice key was used by the subject to terminate the exposure and stop the timer. On every trial, a .5-sec warning tone was relayed to the subject over a pair of headphones immediately preceding the onset of the first stimulus and, for the 2-sec ISI, immediately before the second stimulus.

Experimental design (See Figure 2). For each of the 10 test object classes, the same stimulus always occurred as the first stimulus of a pair, while the second stimulus could be identical (Condition Ps "same"), the same object photographed from a different viewpoint (Condition Pv "same"), a picture of a different object having the same name (Condition Pd "same") or a picture of an object from a different object class (Condition "different"). For each subject, each object occurred 12 times as the first stimulus of a pair. Two pairs were produced for each of the three "same" conditions, and six pairs for the "different" condition. In this way, 120 picture pairs were produced: 60 "same" pairs (10 LF and 10 HF for each of the three "same" conditions) and 60 "different" pairs (30 LF-HF and 30 HF-LF pairs).

The 120 test picture pairs were randomly divided into two equal blocks. Within each block there were five LF and five HF pairs for each of the three "same" conditions and 30 "different" pairs. Two ISIs were used (.25 and 2 sec), one block being assigned to each. The order of ISIs and the assignment of blocks to ISIs was balanced across subjects.

A practice block of 20 picture pairs was constructed from the

Figure 2. Examples of stimulus conditions used in Experiment II.

10 practice object classes. Half of the pairs were from Condition Ps "same" and half from Condition "different." An ISI of 1 sec was used.

Procedure. Subjects were told that they would be shown pairs of pictures and that they were to respond "Yes" if the pictures in a pair were of objects having the same name and "No" if they were of objects having different names. The three "same" conditions were explained to them, and they were told that only Condition Ps "same" and Condition "different" would occur during practice. Before each block, subjects were told the duration of the ISI and were reminded to respond as quickly and accurately as possible. After the practice block, each subject responded to the two test blocks, each of 60 trials. There was an interval of 1 min after every 30 test trials. All errors were immediately corrected.

Table 2
Distribution of Errors in Experiment I

	“Same”						“Different”		
	Ps		Pv		Pd		n	%	
	n	%	n	%	n	%			
LF	1	1.25	3	3.75	5	6.25	HF-LF	11	4.58
HF	4	5.0	6	7.5	5	6.25	LF-HF	14	5.83
Total	5	3.13	9	5.63	10	6.25		25	5.21

Note—n = number of errors per condition.

Results

Errors. Of the 960 responses made, 49 (5.1%) were errors (see Table 2). These were discarded from the data.

Analysis of RT data. As in Experiment I, repeated-measures design analyses of variance were used, and preliminary tests for pooling interactions with Subjects were carried out (using a .25 significance level). For both the analysis of all the RTs and that carried out on the "same" RTs, it was possible to pool all the interactions with Subjects and the within-cell sum of squares to produce a single error term.

"Same" RTs. The results for the "same" conditions and for "different" responses are presented in Figure 3. The analysis of variance showed that the increase in RT from the .25- to the 2-sec ISI was significant [$F(1,461) = 101, p < .001$], but this did not interact with either the effect of Frequency or Condition ($F < 1$ in both cases). The main effect of Conditions was significant [$F(2,461) = 24, p < .001$], as was the main effect of Frequency [$F(1,461) = 8.12, p < .01$] and the interaction of Frequency by Conditions [$F(2,461) = 7.09, p < .01$].² From Figure 3 this interaction can be seen to be due to the fact that, for both ISIs, there is only an effect of Frequency for Condition Pv. Further analysis revealed

Figure 3. Mean RTs and their SEs ($n = 40$) for Conditions Ps, Pv, and Pd, and mean "different" RTs as a function of ISI and frequency (Experiment II).

that the six conditions at each ISI (LF: Ps, Pv, and Pd; HF: Ps, Pv, and Pd) fell into two groups. Fast RTs were obtained for both Ps conditions and for Condition Pv(HF), while RTs were slower for both Pd conditions and Condition Pv(LF). Using Scheffé's test, it was found that the difference between the "fast" and "slow" groups of conditions was significant ($p < .05$) for each ISI, while the differences within each group were not significant (as is clear from Figure 3). The implications of this result for the coding processes involved and the effects of frequency in matching tasks will be discussed later.

"Different" RTs. The second analysis of variance showed no overall difference between "same" and "different" RTs ($F < 1$). This is contrary to what would be expected if "different" judgments took as long as the longest "same" judgments. The mean Pd "same" latency was 692 msec, while "different" latencies averaged 650 msec.

Thus, it may not always be necessary for processing to be carried to the semantic or name level before a "different" judgment can be made. It is possible that there are certain types of difference which are detectable at the picture or object code levels which necessarily imply that the two stimuli are of objects from different object classes. For example, it is not usual for object classes to contain some members which can be characterized as having a generally angular shape while others are rounded. If such a difference between a pair of stimuli (i.e., that one was an angular shape and the other a round one) was detected at a low level, further

processing might be preempted, and a "different" response made.

The objects used in this experiment could be distinguished in three main ways at the object level.³ As well as being classifiable as either rounded (e.g., CAT, DOG) or angular (e.g., CHAIR, SPANNER), they could be classified as either having a skeletal (or open) structure (e.g., CHAIR, PLIERS) or a filled-in (or closed) structure (e.g., CAR, CAT). They could also be divided into those seen against a plain background (e.g., PLIERS, BOOK) and those seen against their natural background (e.g., CAR, LORRY). This last distinction was confounded with object size, as the small objects had been photographed against a plain background, while the large ones had been photographed against their natural background. Each of the 10 object classes was classified according to whether the photographed objects were (a) rounded or angular, (b) open or closed, (c) against a plain or featured background. As each of the five HF object classes could be paired with any of the five LF ones, 25 "different" pairings were possible, of which 19 in fact occurred in the experiment. Each of these 19 pairings was examined to see if any changes occurred within the pair on one or more of the above three dimensions, and, if any changes did occur, to see which dimensions they were on. For 4 of the 19, no changes occurred (i.e., if the first stimulus was open, rounded, and on a plain background, the second was also). Changes occurred on just one of the three dimensions for six of the pairs, on two dimensions for another three pairings, and on all three dimensions for the final

Table 3
 "Different" RTs as a Function of the Dimension on Which
 Members of a Stimulus Pair Differ (Experiment II)

Dimension	Pairs That Differ		Pairs That Do Not Differ		Difference	
	RT	N	RT	N		
Rounded-Angular	639	7	669	12	30	n.s.
Open-Closed	643	13	691	6	48	$p < .01$
Background:						
Featured-Featureless	642	10	677	9	35	$p < .05$
Mean	642		677		35	

Note—N = number of stimulus pairings across which means are calculated.

six pairings. Mean "different" latencies for these four conditions (zero, one, two, and three changes) were 712, 649, 650, and 639 msec, respectively. Latencies for pairs which differed on one or more of the three dimensions were significantly shorter than for those which did not differ on any dimension [$t(17) = 3.25$, $p < .005$]. However, latencies were not significantly affected by the number of changes, only by whether or not at least one change occurred. In Table 3 the effects of change within pairs on a particular dimension are presented. For the curved-angular dimension, the difference between those pairs in which a change did occur and those in which no change on that dimension occurred did not reach significance ($t = 1.56$, $df = 17$). The two remaining differences in Table 2 were significant: A change along the open-closed dimension decreased latencies by 48 msec [$t(17) = 1.67$, $p < .01$], while a change of background reduced latencies by 35 msec [$t(17) = 1.96$, $p < .05$].

While such a post hoc analysis should be treated with caution, it does suggest that one or more changes of the types described above are sufficient for a "different" judgment to be made without having to continue processing of the second stimulus to the "semantic" level. When a "different" judgment cannot be made on the basis of detecting such changes, latencies (712 msec) are slightly longer than Pd "same" RTs (692 msec), as was originally expected.

The analysis of variance carried out on the "same" and "different" responses also revealed an interaction of Frequency by Response [$F(1,945) = 4.8$, $p < .05$]. From Figure 3, it can be seen that a frequency effect occurred for "same" judgments (Condition Pv) but not for "different" judgments. In fact, the effect for "different" judgments is in the wrong direction: If the second stimulus of a mismatch pair is LF, it is responded to slightly faster than if it is HF. (This difference was not significant.)

DISCUSSION

Levels of Coding for Line Drawings of Objects

With simple stylized line drawings as stimuli (Experiment I), increases in "same" RTs were found from Condition Ps to Pv and from Pv to the two Pd con-

ditions. However, there was no difference between Conditions Pds and Pdv. This result is consistent with Klatzky and Stoy's (1974) finding that visual similarity only affects comparisons between different objects having the same name when there are only two matching conditions (in their case, Conditions Ps and Pd). The introduction of Condition Pv in the present experiment appears to have had the same effect on Pd comparisons as their mirror-image matching condition had, in that the presence of a second, more complex, visual matching condition may have prevented the generation of some form of abstract visual code for mediating Pd comparisons.

While Klatzky and Stoy (1974) found no difference in RT between identity matching in their experiment, in the present study a difference was found between Ps and Pv comparisons. This is consistent with the argument that mirror-image matching is a special case of a picture code mediated comparison, while changes in viewpoint (i.e., rotations in depth rather than in the picture plane) result in comparisons being mediated by object codes.

Alternatively, it could be argued that, instead of Condition Pv involving a recoding operation, it may simply involve a rotation operation. That is, orientation-specific object code descriptions are formed of both stimuli and then the description of one is rotated in depth to match the other. Thus, the RT difference between Conditions Ps and Pv would be accounted for by the time taken to carry out this operation. This would imply that the rotation operation is carried out after the second stimulus has been presented and not during the ISI. This would seem to be the optimal strategy to adopt for a number of reasons. First, the operation would only have to be carried out on 25% of all the trials. Second, if the first stimulus was always rotated, then on 50% of the trials it would be necessary to produce an internal representation of a perspective view from an elevation view stimulus. This would involve making inferences about information not present in the stimulus. On the other hand, if the rotation operation is carried out after the second stimulus occurs, the perspective view can always be rotated to an elevation view. This would seem to be a simpler strategy and involve less risk of error.

These two hypotheses, the recoding hypothesis and the rotation hypothesis, are not mutually exclusive. It could be argued that, in those cases where a rotation is not carried out, comparisons are mediated in terms of picture codes (Condition Ps) but, when there is a change in viewpoint, it becomes necessary to rotate the representation of one of the stimuli in depth. To do this, it must be necessary to transform the picture codes into object codes, as, by definition, a picture code cannot be rotated in depth, only in the picture plane. Clearly, further studies are needed to determine whether the difference between the two conditions arises from recoding operations, rotation operations, or both.

Levels of Coding for Photographs of Objects

In Experiment II a rather different pattern of results emerged. For "common" objects there was only a small, statistically nonsignificant, difference between Conditions Ps and Pv, but both were faster than Condition Pd. For "rare" objects Condition Ps produced faster responses than the other two conditions, while Conditions Pv and Pd did not differ significantly. This pattern of results occurred for both the ISIs used.

The presence of a frequency effect for Condition Pv appears to argue against Wingfield's (1968) assertion that frequency does not affect perceptual identification time in matching tasks. However, examination of the data suggests that frequency may not be the crucial variable responsible for the difference between Pv(LF) and Pv(HF) RTs. While there is insufficient data on each of the stimulus objects used to carry out a full statistical analysis, it is interesting to note that the marked rise in RT from Condition Ps to Pv occurs for all the LF objects used except LORRY. For the latter, there is a small rise (12 msec) comparable to the results found for the HF objects. The remaining LF objects are all metallic tools which form a set of, visually, rather confusable stimuli. Hence, it could be argued that, for these objects, the comparison locus for Condition Pv may be transferred to a nonvisual level of coding in order to reduce the risk of error responses.

The fact that the difference in RT between Conditions Ps and Pv(HF) is so small suggests that comparisons in both conditions are carried out at the same level of coding. Given that picture codes can only be used in comparisons between identical stimuli (or rotations of such stimuli in the picture plane, or mirror-image reversals), the comparison locus for Conditions Ps and Pv(HF) in Experiment II is likely to be at the object code level. If this is the case, this implies that the process of producing an object code description of a stimulus from its picture code is obligatory. It will be recalled that the same conclusion was reached in Bartram (1974), where the same stimuli were used in an object naming paradigm.

It seems that the results of Experiment II can be accommodated within the framework of a simple dual-code model. Comparisons in Conditions Ps and Pv (with the exception of the four confusable LF objects in Condition Pv) are mediated by object codes, while for Condition Pd and the visually confusable stimuli in Condition Pv, comparisons are mediated either by name codes or some form of abstract nonvisual code. The results of Experiment I can also be accommodated within this model if it is assumed that the difference between Conditions Ps and Pv reflects the time taken to "rotate" orientation-specific object codes, rather than the time taken to produce an object code from a picture code. However, this explanation still leaves open the question of why there is a difference between Conditions Ps and Pv in Experiment I and not in Experiment II. There seem to be two main factors which may

have induced subjects to adopt different comparison strategies in the two experiments: first, the difference in mode of representation (stylized line drawing as opposed to photograph) and, second, the difference in selection of viewpoints (two fixed types of viewpoint in Experiment I, e and p views, as opposed to a representative sample of different viewpoints in Experiment II). Clearly, these two factors need to be examined independently in order to account for the differences between the results of the two experiments.

Nonvisual Codes

Pd and Pv(LF) comparisons could either be mediated by verbal or by nonverbal semantic codes. Klatzky and Stoy's (1974) study showed that Pd-type comparisons can be affected by visual similarity under certain conditions. This implied that they were mediated by a form of coding which could be used to generate abstract visual images. In order to generate such an image from a name, it must first be necessary to retrieve information about the properties of the object class to which that name refers. Such information would be contained in the stored semantic description of that object class. Bartram (1974) has argued that, in order to name an object, it is necessary to interpret the visual information contained in its object code description in terms of the defining properties of objects belonging to the appropriate object classes. The abstract description produced by this reinterpretation would seem to possess the properties necessary for object-class name retrieval as well as for generating visual images of objects belonging to that object class. Hence, given that the first stimulus of a pair is a picture, it seems more likely that such an image would be generated directly from the semantic description, rather than indirectly from the name.

For Pd comparisons and for Pv(LF) comparisons, it seems that the most likely locus of the comparison process is some form of coding which, while not necessarily verbal, is more abstract than the visual object code. It is unlikely to be, itself, any form of visual code, as no clear difference was found between Condition Pv(LF) and Pd(LF and HF). Considering the fact that Pv pairs of stimuli are visually more similar than Pd pairs, one would expect Pv(LF) RTs to be faster than Pd(LF) RTs if both were mediated by some form of visual code, however abstract. From the above arguments, it seems plausible to identify the code mediating Pd comparisons with the nonverbal abstract semantic code proposed in Bartram (1974). This may well be comparable to the abstract modality-free coding system proposed by Clark and Chase (1972) to account for the nonverbal and nonvisual properties of encoded propositions in proposition-picture matching tasks.

CONCLUSIONS

The results of Experiment I confirm the findings of Klatzky and Stoy (1974) and show that an inter-

mediate level of matching (Condition Pv) may occur. The evidence is consistent with the view that this level of comparison is carried out in terms of a more abstract type of code (i.e., an object code) than that used for Condition Ps (i.e., a picture code). However, the evidence is also consistent with the argument that both types of comparisons are mediated by object codes, the difference in RT being due to the time taken by a rotation operation required in Condition Pv. Further support for the argument that Ps comparisons are mediated by object codes and not by picture codes comes from Experiment II, where it was found that Ps and Pv RTs did not differ significantly.

Biederman (Note 1) suggests a plausible explanation for the difference in results between Experiments I and II. He argues that "a photograph might enable one to infer more features and employ more stimulus features that were orientation independent (e.g., surface texture and grayness) than could be inferred and employed with line drawings If these inferred and stimulus features were processed in parallel, then the difference between Ps and Pv conditions would be reduced relative to the line drawings [but not eliminated]." The data is consistent with this, as it was found that the 37-msec difference between Conditions Ps and Pv in Experiment I was reduced to 14 msec for the HF objects in Experiment II.

One point which clearly emerges from the present study is the danger of regarding different modes of pictorial representation as equivalent. They differ not only with respect to the amount of information, but also with respect to the conventions used to represent that information. These differences may have important consequences for the ways in which such stimuli are treated by coding mechanisms. A systematic study of different modes of representation could well provide some useful information on the nature and flexibility of the coding processes involved in perception.

REFERENCE NOTE

1. Biederman, I. Personal communication, September 8, 1975.

REFERENCES

- BARTRAM, D. J. The effects of familiarity and practice on naming pictures of objects. *Memory & Cognition*, 1973, 1, 101-105.
- BARTRAM, D. J. The role of visual and semantic codes in object naming. *Cognitive Psychology*, 1974, 6, 325-356.
- CLARK, H. H., & CHASE, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, 3, 472-517.
- FRAISSE, P., & ELKIN, E. H. Étude génétique de l'influence des modes de présentation sur le seuil de reconnaissance d'objets familiers. *Année Psychologie*, 1963, 63, 1-12.
- FROST, N. Encoding and retrieval in visual memory tasks. *Journal of Experimental Psychology*, 1972, 95, 317-326.

- KLATZKY, R. L. Visual and verbal coding of laterally presented pictures. *Journal of Experimental Psychology*, 1972, 96, 439-448.
- KLATZKY, R. L., & STOY, A. M. Using visual codes for comparisons of pictures. *Memory & Cognition*, 1974, 2, 727-736.
- OLDFIELD, R. C. Things, words and the brain. *Quarterly Journal of Experimental Psychology*, 1966, 18, 340-353.
- POSNER, M. I., BOIES, S. J., EICHELMAN, W. H., & TAYLOR, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology*, 1969, 79(1, Part 2).
- POSNER, M. I., & MITCHELL, R. F. Chronometric analysis of classification. *Psychological Review*, 1967, 74, 392-409.
- POSNER, M. I., & WARREN, R. E. Traces, concepts and conscious constructions. In A. W. Melton and E. Martin (Eds.), *Coding processes in human memory*. New York: Wiley, 1972.
- RYAN, T. A., & SCHWARTZ, C. B. Speed of perception as a function of mode of presentation. *American Journal of Psychology*, 1956, 69, 60-69.
- SEYMOUR, P. H. K. Rule identity classification of name and shape stimuli. *Acta Psychologica*, 1973, 37, 131-138.
- THORNDIKE, E. L., & LORGE, I. *The teachers word book of 30,000 words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.
- WINER, B. J. *Statistical principles in experimental design*. London: McGraw-Hill, 1970.
- WINGFIELD, A. Effects of frequency on identification and naming of objects. *American Journal of Psychology*, 1968, 81, 226-234.

NOTES

1. An object-class name may be defined as that name which is normally spontaneously used to name a particular object. For example, an Alsatian is normally called "dog" rather than either "Alsatian" or "animal."

2. In Bartram (1974, p. 326) the results of the present experiment were misleadingly reported as follows: "Ps latencies were about 50 msec shorter than Pv latencies, which in turn were about 50 msec shorter than Pd latencies." The original analysis of the data, carried out at Sussex, suggested that there was no interaction of Conditions by Frequency. On recently reexamining the data, it was obvious that this was not the case. The original incorrect analysis was traced to a data tape on which the input parameters were in the incorrect order! Fortunately this mistake does not affect the conclusions drawn in Bartram (1974) with respect to the results of the naming studies (i.e., if 2-D coding occurs in naming photographs of objects, 2-D to 3-D coding is obligatory).

3. These conclusions are based on the results of a sorting task carried out with six subjects. Subjects were given triads of pictures of objects from the experiments and told to pick out the pair which were similar to each other and different from the third item. They then sorted the remaining object classes with respect to this difference. Each subject was presented with 12 randomly selected triads. They were all told to make their decisions on the basis of visual similarities, not in terms of semantic category membership. From the results of these sortings, the three dimensions described above were isolated: These were used consistently by all six subjects along with a small number of other relatively idiosyncratic dimensions.

(Received for publication June 11, 1975;
revision accepted March 3, 1976.)