Psychol Res (1993) 55: 237–247

Psychological Research
Psychologische Forschung

Codes and operations in picture matching

G. Keith Humphrey and Stephen J. Lupker

Department of Psychology, The University of Western Ontario, London, Ontario Canada, N6A 5C2

Received August 3, 1992/Accepted February 8, 1993

Summary. Ellis and Allport (1986; see also Ellis, Allport, Humphreys & Collis, 1989) proposed a model of object perception wherein successively more abstract descriptions are generated as a function of processing time. The aspect of their model that is examined here is the proposal that viewer-centred representations of objects decay rapidly whereas object-centred or semantic-level representations do not. To test the model, a picture-matching task was used in which subjects decided whether successively presented pictures rotated in the frontal plane had the same name. The pictures were either identical pictures, pictures of different objects with the same name, or pictures of objects with different names. The two successive pictures could be in the same orientation or in a different orientation. In Experiment 1, two orientations (0° upright and 120°) and two ISIs were examined (100 ms and 2 s). In Experiment 2, two orientation (0° and 60°) and three ISIs were examined (100 ms, 2 s, and 5 s). In neither experiment was there any evidence that viewpoint-specific representations disappeared at longer ISIs. These results, although consistent with other research on the perception of rotated objects, did not replicate the results of Ellis and Allport (1986) and are inconsistent with their model.

Introduction

Our interactions with objects serve a variety of purposes. For example, sometimes we must manipulate objects, sometimes we must identify them uniquely, and sometimes we must classify them into broad classes. The visual information and representations that guide action, identification, and classification are not necessarily the same. Activities such as reaching for, and grasping, an object rely on a representation that codes the size, location, and orientation of the object with respect to the observer. Such a representation is often called a *viewpoint-dependent* or *viewer-centred representation* in that it captures how an object appears from a particular vantage point. In contrast, identifying an object may involve a more abstract representation in which the information concerning the size, position, and orientation of the object is decoupled from the information that specifies its invariant shape (e. g., discussed in Pinker, 1984). These more abstract representations have been called *viewpoint-independent* or *object-centred representations* in that they characterize the three-dimensional structure of an object regardless of viewpoint. Finally, the classification of distinct objects as members of the same class may require an even more abstract, *semantic* representation, based on more general, structural characteristics.

Considerations of the sort described above have led many to propose models based on the idea that there are several different types or levels of object representation. One of the most influential of these types of theory was developed by Marr (1982; Marr & Nishihara, 1978). According to Marr's theory, visual-object processing proceeds from a viewer-centred to an object-centred representation. In a relatively early stage of visual processing a viewer-centred representation, referred to as the $2^1/_2$-D sketch, is formed. In the final stage of processing, however, the representation of an object consists of a viewpoint-independent structural description of the object. Such a description decomposes an object into generic parts, generalized cylinders according to the model, and the major relations among parts. According to Marr's model, the long-term representation of objects is object centred or viewpoint independent.

A recent model of object representation proposed by R. Ellis and his colleagues (Ellis & Allport, 1986; Ellis, Allport, Humphreys, & Collis, 1989) has some features in common with the proposal of Marr (1982; Marr & Nishihara, 1978). They, like Marr, posit a view-specific representation and an object-centred representation. They propose a quickly decaying view-specific representation called the VIEW code and two other representations that are object centred and long term. One long-term representation, referred to as the OBJECT code, represents a partic-

Correspondence to: either author

ular object independent of viewing position. Another long-term code, the MODEL code, represents concepts, and thus can be used to determine if physically different objects have the same name. This code is obviously based on more abstract features. Ellis and colleagues based their model on results obtained with picture-matching tasks similar to Posner and Mitchell's (1967) letter-matching task. Because the research of Ellis et al. was, to some extent, a follow-up on the earlier research of Bartram (1976), we shall first describe Bartram's experiments.

Bartram (1976) had subjects judge whether sequentially presented pairs of pictured objects would have the same name. The pairs could be the same object from identical viewpoints (hereafter refered to as the *identity* condition), the same object from different viewpoints (hereafter refered to as the *rotated* condition), or two different objects with the same name. The disparate views of the same objects differed by a rotation in depth of about 45°. When line drawings of familiar objects were used with a 500-ms interstimulus interval (ISI), Bartram reported an apparently significant 37-ms difference between the identity and rotated conditions. (In Bartram's experiments the first picture was always presented for 500 ms.) When photographs were used with both 250-ms and 2-s ISIs, there was an even greater difference between these two conditions. Bartram noted, however, that this latter difference seemed to be due to a subset of his stimuli that were all visibly confusable with one another. The other stimuli showed much less of an effect. Thus, the status of this effect with photographs was somewhat unclear. With both types of stimuli, however, there was a clear advantage of the rotated condition over the condition in which the stimuli were different pictures with the same name.

At a theoretical level, Bartram (1976) argued that his experiments could be taken as evidence for the operation of three different levels of representation – a *picture* code which would be, in our terms, viewer-centred; an *object* code which could be considered to be viewpoint-independent; and a *semantic* code. He suggested that the picture code might be used only for line-drawn stimuli and not for photographs. Alternatively, he suggested that there may be only an object code and a semantic code for both types of stimuli and that the differences between the identity and rotated pairs with line drawings may have occurred because of the time taken to align, via mental rotation, the rotated pictures.

Ellis and Allport (1986; see also Ellis et al., 1989) repeated Bartram's (1976) experiments both to evaluate the generalizability of Bartram's results and to examine the temporal-decay rates of different visual codes. They suggested that the stimuli used by Bartram (1976) and others (e.g., Kelter et al., 1984) may have been problematic. The different views of the same objects in these studies were produced by rotation of objects in depth by about 45° (i.e., out of the picture plane). Ellis and Allport point out that such a rotation very often results in different features being visible in the two views. Consequently, any performance differences between identity and rotated conditions could have resulted from the fact that the rotated images actually had somewhat different visual features than the identical, non-rotated images. Ellis and Allport also used objects that were photographed rotated in depth, but made sure that for different views of the same object no major feature was obscured nor was a major axis foreshortened.

Like Bartram (1976), Ellis and Allport (1986) used the matching task in which subjects had to decide if successively presented photographs had the same name. Four trial types were used. *Same* trials included identical views of the same objects (the identity condition), the same object photographed from different angles (the rotated condition), and different objects with the same name. On *different* trials two different objects were presented. The first picture in the pair to be matched was presented for 500 ms and was then followed, after an ISI of 100 ms, 500 ms, or 2 s, by the second picture. At the two short ISIs, *same*-decision latency was fastest for the identity condition, next fastest for the rotated condition, and slowest for the name matches. However, at the 2-s ISI, the identity and rotated conditions did not differ, although both were faster than the same-name condition. Additional results relevant to this issue were that the advantage of the identity condition over the rotated condition also: (a) was eliminated when a pattern mask was shown between the picture pairs even at the 100-ms ISI (Ellis & Allport, 1986) and (b) was significantly reduced when the first and second pictures were different sizes (Ellis et al., 1989).

As was noted earlier, these results led Ellis and his colleagues to propose a three-level representational system for picture matching. One of these levels is called VIEW to emphasize that it is viewpoint specific or viewer centred, although it is non-retinotopic (Ellis et al. 1989). It is this code that accounts for the advantage of the identity condition over the rotated condition. Evidence for its existence and information about its nature come from the three manipulations that cause this advantage to reduce or disappear: (a) the use of a long ISI; (b) the insertion of a mask between the two stimuli; and (c) the use of different-sized first and second stimuli. A second code is called OBJECT, as it is believed to be object centred. An experiment reported in Ellis et al. (1989) indicates that this code requires more time to be generated than VIEW and that it is not disrupted by masking. It is, however, more enduring than VIEW. Finally, they propose another internal code, MODEL, to account for the matching of physically different objects that have the same name. As was discussed above, the model of object representation of Ellis and his colleagues, then, is similar to the proposal of Marr (1982; Marr & Nishihara, 1978) in that view-specific and object-centred representations are posited.

To place the research of Ellis et al. in a larger context, however, it is important to note that there is a continuing debate over whether human object recognition employs a long-term representation that is viewpoint dependent or viewpoint independent. Although Marr proposed an object-centred representation scheme for a variety of reasons, other theorists (e.g. Poggio & Edelman, 1990; Tarr & Pinker, 1989) have proposed viewpoint-dependent representation schemes. Part of the impetus for suggesting viewpoint-dependent schemes comes from research showing that recognition of familiar objects is impaired if objects are seen at an orientation that departs from their usual or canonical orientation (e.g. Jolicoeur, 1985; 1990a; Humphrey & Jolicoeur, 1988; 1993; Palmer, Rosch, & Chase, 1981). Other research shows a deleterious effect on the recognition of novel objects if they are presented in an orientation that deviates from the orientation in which they were viewed originally (Bülthoff & Edelman, 1990; Bülthoff, Edelman, & Sklar, 1991; Humphrey & Khan, 1992; Rock & DiVita, 1987; Rock, DiVita, & Barbeito, 1981; Tarr & Pinker, 1989). Collectively, these results suggest that the viewpoint-dependent representation of objects is very long lasting and certainly does not decay within a couple of seconds.

To explain how objects that deviate from their long-term view-specific representations are recognized, several researchers have argued that normalization operations such as mental rotation are used to align the viewed object with its long-term representation (e.g. Jolicoeur, 1985; Tarr & Pinker, 1989; Ullman, 1989). This suggestion is consistent with the monotonic increase in recognition or naming latency with misorientation of the object in relation to a canonical view (e.g., Jolicoeur, 1985; for review see Jolicoeur, 1990a).

It is apparent that the results and theorizing of Ellis and his colleagues are at odds with a large body of evidence that suggests that long-term representations of objects are viewpoint dependent and that normalizing operations play a significant role in object recognition. In their model the VIEW code is seen to decay quite rapidly and there is no need for normalizing operations. Given these conflicting results and models, a systematic study of the effects of varying ISI in a picture-matching task is warranted. In addition to these concerns, we must point out that there may be potential methodological problems with the research of Ellis and Allport (1986) and Ellis et al. (1989) that could lead to questions about the reliability of their results. One important point is that they apparently did not include a condition in which same-name objects had different orientations. Presumably, the objects used on *different* trials generally had different orientations because of their different shapes. Thus, orientation could be a reasonably reliable cue to the correct response. That is, objects with the same orientation (in the identity and same-name conditions) would always require a "same" response. Those with a different orientation, however, would require a "same" response in the rotated condition and a "different" response in the different condition. The result is that seeing objects in the same orientation would create a bias to respond "same," while seeing objects in different orientations may create a bias to respond "different." In this regard, it is interesting to note that Bartram (1976) did include same-name, rotated pairs and, thus, his data should have been less influenced by these types of bias. As we noted above, Bartram did not find strong support for an identity rotated difference in the condition using photographs of familiar objects – the condition that most closely matches that of Ellis and Allport (1986). This methodological issue, as well as the two others noted below, would appear to raise questions about the reality and/or theoretical interpretation of the difference between the identity and rotated conditions in Ellis and Allport (1986) and Ellis et al. (1989).

Experiment 1

Our main focus is on the difference between the identity condition and the rotated conditions and how this varies as a function of ISI rather than on the other evidence used by Ellis and his colleagues to support the existence of a VIEW code. First of all, with respect to the identity-rotated difference, the question is, is there really a difference between these conditions, as Ellis and his colleagues claim (Ellis & Allport, 1986; Ellis et al., 1989), or are the two conditions essentially the same, as some of Bartram's (1976) results suggest? To control for response biases as much as possible, half of the identical pairs, half of the same-name pairs, and half of the different pairs had the same orientation, while half of each set did not. Further, on the basis of Bartram's results suggesting that this difference may be more reliable with the use of line drawings than with photographs, we also used line drawings. We had, however, a more compelling reason for using line drawings. That is, in comparison to photographs, line drawings are essentially devoid of all texture and grey-level features. Although the role of surface attributes in object recognition is complex (e.g., Biederman & Ju, 1988; Price & Humphreys, 1989), it is possible that surface features may have formed a basis for performing Ellis and Allport's object-matching task. In particular, as we shall argue subsequently, they could have provided a means of matching identical views of the same object faster than any other types of stimuli, at least at short ISIs.

The second issue was whether ISI is really important. One of the cornerstones of the proposal of Ellis and Allport (1986) is that ISI affects OBJECT and VIEW codes differently in that the VIEW code disappears at longer ISIs, thus eliminating any advantage for the identity condition. Bartram (1976), as noted, did not find any type of interaction with ISI. We presented the first picture for 500 ms, as both Bartram (1976) and Ellis and Allport (1986) had done. Two ISIs between the first and second pictures (100 ms and 2 s) were used. These ISIs match two of those used by Ellis and Allport and span the range in which they found that identity matches were faster than matching rotated pictures (100-ms ISI) and they did not differ (2-s ISI).

The final issue concerns the type of rotation used. In all previous studies, rotation in depth was used. Ellis and Allport criticized Bartram's study because the objects were rotated to an extent that many of the features visible in the frontal view were no longer visible in the rotated view. Thus, it may have been the case that the views looked so dissimilar that they could not be matched using the OBJECT code.[1] Ellis and Allport's solution to this was to rotate their objects to such a degree that none of the important features was obscured or a major axis foreshortened in either view. The problem that this creates, however, is that the two views of an object are highly similar (at least for some pictures), making it unclear exactly how subjects are carrying out the matching process in the rotated condition.

[1] However, a characteristic of a truly object-centred representation, like the OBJECT code, is that dissimilar views are all matched to the same representation. In this sense, the presentation of dissimilar views should not matter.

This problem was solved by the use of rotation in the picture plane (120° clockwise in Experiment 1). These types of rotation leave all the features visible, and yet should create a situation in which it would not be possible to match rotated pairs by any other means than a viewpoint-independent code. In any case, the use of frontal-plane rotations will test the generalizability of the model of picture matching developed by Ellis and colleagues, as the model is apparently meant to apply to all orientation changes that do not obscure features or foreshorten axes.

Method

Subjects. Sixty-four subjects from introductory psychology classes at the University of Western Ontario received course credit for participating in the experiment. All were naive with respect to the purpose of the experiment.

Stimuli and apparatus. Line drawings of common objects were used. Examples of these drawings are shown in Figure 1. The line drawings of the objects were photographed as slides on high-contrast copy film with the contours as black lines.

Two Kodak Ektagraphic IIIB slide projectors fitted with Kodak Ektanar C zoom lenses were used to rear-project the slides. An Ilex electronic shutter was mounted in front of each of the lenses and presentation time was controlled by an Apple IIe microcomputer. The slides were rear-projected onto a circular aperture 22 cm in diameter in the centre of a screen positioned 74 cm directly in front of the subject. The longest axis of each object subtended, on average, 8.4° of visual angle (SD = 2.7°). The subject's head was kept from rotating by a chin-rest equipped with a forehead stop and two lateral head stops.

One slide projector projected the first slide, while the other projected the second. The subjects were told that they would be presented with two line drawings of common objects, one after the other. They were told that they were to decide, as quickly and as accurately as possible, whether the two objects depicted had the same name. Matching latency was recorded by the Apple microcomputer using a button-box with two buttons – one for "same", and one for "different", responses. Timing began with the opening of the tachistoscopic shutter on the slide projector projecting the second slide.

Design and procedure. The 12 different trial types used in the experiment are illustrated in Figure 1. Each subject saw 128 pairs of slides during the experiment; half were *same* trials and half were *different* trials. For the 64 *same* trials, half were identical pictures and half were pictures of different objects with the same name. In each of these conditions, half were in the same orientation on both trials and half were in different orientations. For the identical trials in the same orientation, there were 8 trials with both pictures at 0° (upright) orientation[2] and 8 trials with both pictures at 120° (rotated). For the identical trials in different orientations, there were 8 trials in which the first picture was presented at 0°, followed by the second picture at 120°, and 8 trials in which the 120° rotation was first and was followed by 0° pictures. These same trial types were repeated for the same-name pairs (8 trials of each type) and for the different pairs (16 trials of each type).

To create the pairs for the *same* trials, 32 picturable concepts were selected and 2 different pictures of each were found. Each of the 64 pictures was seen by each subject once as the first picture and once as the second picture. There were eight circumstances in which the second stimulus could be presented; two orientations by four conditions. Thus eight groups of four subjects at each ISI were used to allow each second picture to be seen once in each of these eight circumstances. The 64 *different* picture pairs were selected randomly from a separate group of

[2] The 0° orientation refers to what we took to be a canonical view of the picture.

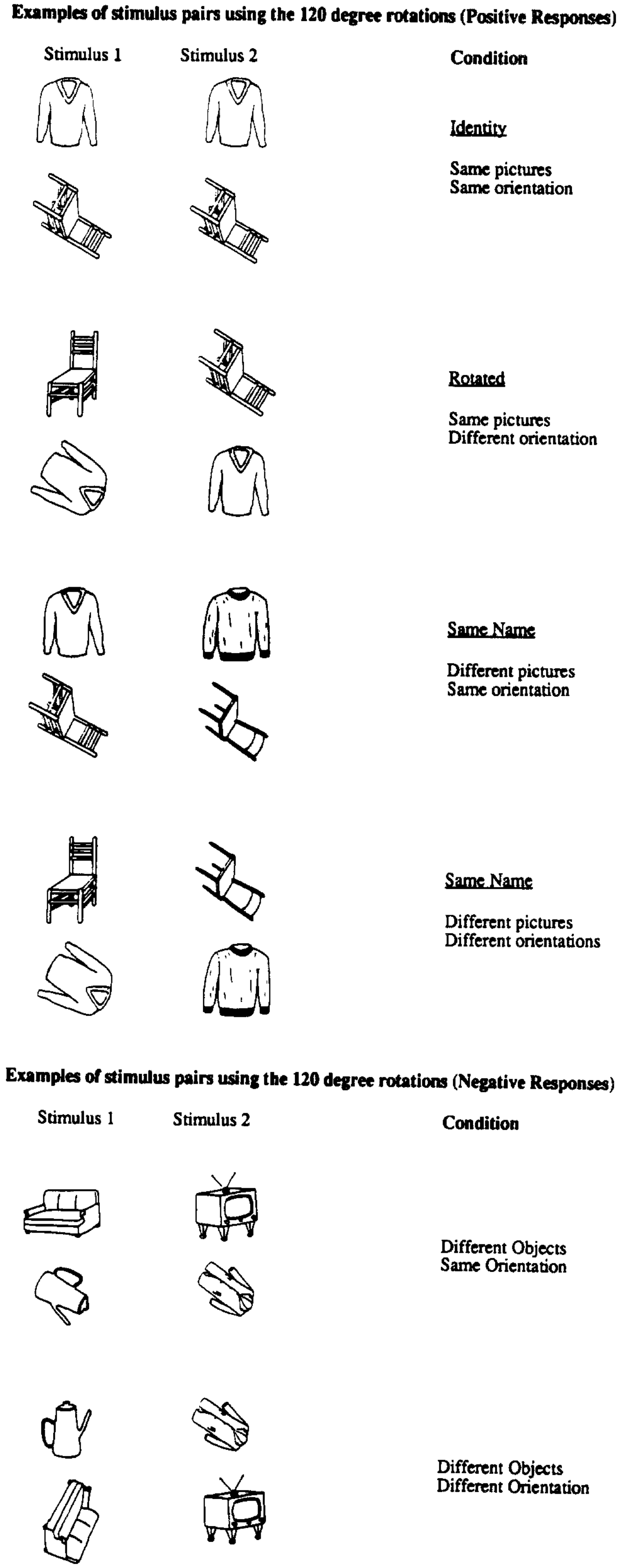

Fig. 1. Examples of the line drawings and an illustration of the trial types used in the experiments

pictures and were the same for all subjects. No attempt was made to control visual similarity.

Each subject was first presented with 12 practice trials in which there was one trial of each of the 12 trial types. The pictures used in the practice

Table 1. The mean reaction time as a function of trial type, orientation and ISI in Experiment 1

	Trial type					
	Identical		Same name		Different name	
	Orientation		Orientation		Orientation	
	Same	Different	Same	Different	Same	Different
ISI						
100 ms	572	618	652	670	690	703
2 s	546	573	587	615	610	630

trials were not used in the experiment proper. After the practice trials, the experimenter answered any questions about the procedure that the subject might have before beginning the 128 experimental trials. For each trial, a subject first heard an orienting tone, which was followed 1 s later by the presentation of the first picture in a pair for 500 ms. After this presentation there was an ISI of 100 ms or 2 s. ISI was a between-subjects factor and 32 subjects were randomly assigned to each of the two ISI conditions. After the ISI, the second picture appeared and remained on until the subject made a response. The inter-trial interval was 2 s.

Results

Same trials. The dependent measure used for all analyses was the mean reaction time for correct trials. The mean reaction time as a function of Trial type (including different trials), Orientation, and ISI is shown in Table 1. An ANOVA was performed on the *same* responses in which the between-subjects factor was ISI (100 ms or 2 s) and the within-subjects factors were Trial type (identical or same name), Orientation (same or different), and Orientation of the first picture in a pair (0° or 120°). The analysis yielded a significant main effect of trial type, $F(1,62) = 67.02$, $MS_e = 5554.6$, $p < .001$, with same-name trials ($M = 631$ ms) being slower than identical ($M = 577$ ms) trials. The main effect of Orientation was also statistically reliable, $F(1,62) = 14.33$, $MS_e = 8052.9$, $p < .001$, with trials in which both pictures were in the same orientation ($M = 589$ ms) being faster than trials with differently oriented pictures ($M = 619$ ms)

The only significant interaction was that between Orientation and Orientation of the first picture, $F(1,62) = 9.29$, $MS_e = 3763.4$, $p < .005$. This interaction is illustrated in Figure 2. Simple effects analyses indicated that for trials with pairs in the same orientation, 0° trials were faster than 120° trials, $F(1,62) = 4.70$, $p < .05$. For trials in which the pairs were in different orientations, trials in which the first picture was oriented at 120° were faster than trials in which the first picture was at 0° orientation, $F(1,62) = 5.22$, $p < .05$. Also, different-orientation trials were slower than same-orientation trials if the orientation of the first picture was 0°, $F(1,62) = 19.59$, $p < .001$; but this was not true if the first picture was 120°, $F(1,62) = 2.47$, *ns*.

The only other interaction even approaching statistical reliability was the interaction between Trial type and ISI, $F(1,62) = 3.53$, $MS_e = 19577.3$, $p < .07$. Although identical trials were faster than same-name trials at both ISIs (both *ps* <.001, according to simple-effects analyses), the difference was somewhat smaller at the 2-s ISI (42 ms) than at the 100-ms ISI (66 ms).

Different trials. As in the analysis of the *same* trials, the dependent measure was the mean reaction time for correct trials. An ANOVA was performed on the different responses in which the between-subjects factor was ISI (100 ms or 2 s) and the within-subjects factors were Orientation (same or different) and the Orientation of the first picture in a pair (0° or 120°).

The main effect of ISI was significant, $F(1,62) = 6.22$, $MS_e = 59904.1$, $p < .02$. Decision latencies were shorter for the 2-s ISI ($M = 620$ ms) than for the 100-ms ISI ($M = 697$ ms). Also significant was the main effect of Orientation, $F(1,62) = 11.20$, $MS_e = 1480.6$, $p < .001$. Trials in which the pictures were in the same orientation were faster ($M = 650$ ms) than trials in which they were differently oriented ($M = 667$ ms). There was also a trend towards a significant interaction between Orientation and Orientation of the first picture, $F(1,62) = 3.48$, $MS_e = 2908.6$, $p < .07$. This interaction took the same form as that found for the *same* trials.

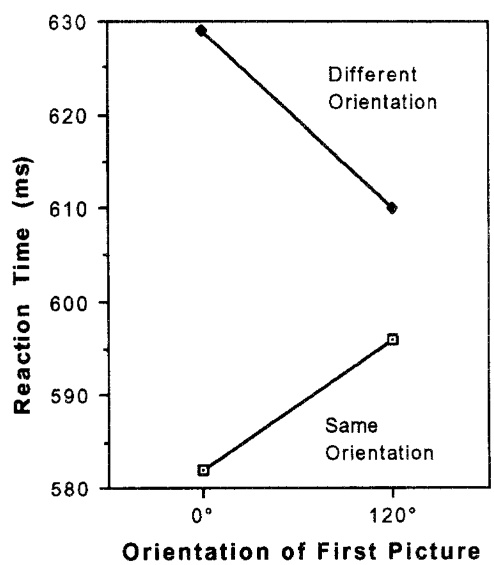

Fig. 2. The mean reaction time as a function of the orientation factor (same or different) and the orientation of the first picture (0° or 120°) in a picture pair in Experiment 1

Errors. The overall error rate was 4%. This included trials on which, for each subject, the latency to respond was more than 3 *SD*s above the mean for that subject. The error rate was 4% for both same and different trials.

Discussion

Some of the results of the *same* trials are consistent with the model of Ellis and colleagues. We found, like others (Bartram, 1976; Ellis & Allport, 1986; Ellis et al., 1989), that matching latency for same-name trials is greater than for identical trials. This result is consistent with the claim that different information is used to match different objects with the same name from that used to match identical objects (i. e., a MODEL code versus less abstract codes).

A second result consistent with the model is that the matching latency for rotated pairs is longer than for same orientation pairs. This finding is consistent with the claim that different codes are used to match identical objects in the same orientation from those used to match identical objects in different orientations. However, there are two aspects of the data that pose problems for this claim and that consequently argue against the model. First, there was no interaction of trial type and orientation. That is, at least statistically, the effect of orientation was the same for both the same-name and the identical trials. According to the Ellis et al. (1989) argument, same-name matches are based on a single code, the MODEL code. Thus, any difference between the same-orientation and the rotated pairs must be due to other factors (e. g., processing operations). Those same factors would presumably also be at work on the identical trials. Thus, given that the size of the orientation effect was statistically the same on the identical trials, the implication would be that the orientation effect on identical trials was (at least partly, if not fully) caused by those factors rather than because different codes were being used in the identity and rotated conditions.

The second problem for the model is the lack of any interaction between Orientation and ISI (or between Trial type, Orientation, and ISI). While the model does not specify how the effect of orientation should vary as a function of ISI for the same-name trials, it is very clear about what should happen on the identical trials. At the longer ISI, the VIEW code should no longer be available and, thus, any effect of orientation on identical trials should have disappeared. Obviously it has not, either statistically or in absolute terms, although it is a bit smaller than at the shorter ISI. Thus, if there is a VIEW code that accounts for some of the difference between these two conditions, its duration must be longer than the longest ISI used here.

There was an unexpected trend towards a Trial type (identical or same-name) by ISI Interaction. As the ISI increased, the difference in matching latency on identical trials and same-name trials decreased. Ellis and Allport (1986) did not report a similar decrease and, in fact, their model would have no reason to predict it. The two codes that are relevant here, the OBJECT and MODEL codes, are both permanent codes. Thus, if this difference is real, it would also require an explanation in terms of a factor such as a decay in the OBJECT code. Such an explanation would, however, also be inconsistent with the characterization of the OBJECT code in Ellis and Allport's model.

One other interesting result was that the orientation factor interacted with the orientation of the first picture of a trial (see Figure 2). If the first picture was oriented at 0° and was followed by another at 0°, matching latency was most rapid. In contrast, if the first picture was at 0° orientation and was followed by a picture at 120°, matching latency was slowest. Intermediate latencies were found if the first picture was at 120°, and for these trials it mattered little whether the second picture was at 120° or at 0°. What is also interesting is that there was no three-way interaction between these factors and Trial type or between these factors and ISI, indicating that the explanation for this interaction must be applicable to both identical and same-name trials at both ISIs. The interaction of the orientation factor with the orientation of the first picture of a trial and the general difficulties that these data present for the Ellis and Allport model leads us to suggest that these results are explained better in terms of a processing account than in terms of the type of representational account offered by Ellis and his colleagues.

We focus first on the interaction between orientation and the orientation of the first picture in a pair. To explain this, it is necessary to propose two processes. We shall refer to these processes as *normalization* and *preparation.* Let us first suppose that the crucial factor in doing the task is accessing the name of the objects to be matched. We also make the assumption that the long-term representation of an object is in a canonical orientation – not an unwarranted assumption, given the many studies showing that naming latency increases as objects depart from their usual, canonical orientation (e. g., Jolicoeur, 1985; Palmer et al., 1981). Specifically, common objects such as those used here are assumed to be stored in a familiar, upright orientation with respect to the viewer. If one is presented with a 120° picture, then some normalization operation(s), including an operation such as mental rotation, will be needed to match the viewed object to the long-term representation before the name is accessed. Accessing the name of a depicted object will be more rapid, however, when presented with a 0° picture, because fewer normalization operations are needed. The first factor, then, is the normalization operation used to process the pictures to gain access to the name of the depicted object (see Ullman, 1989, for a recent discussion of the role of normalization operations in object recognition).

The second factor is preparation by the first picture for the second. The operations used to access the name of the object depicted in the first picture can carry over to the processing of the second picture. If the operations match, as in the 0°–0° and 120°–120° sequences, one is somewhat prepared, and processing will be quite rapid. If the operations do not match, as in the 0°–120° and 120°–0° sequences, processing will be slowed down. The 0°–120° sequence is particularly slow because of the mismatch in operations between the first and second pictures and the need to carry out extensive normalization operations on the second picture. The 120°–0° sequence is more rapid, even though the operations mismatch, because the second 0° picture can access the name of the depicted object rapidly as the long-term representations of objects are stored in such an upright, canonical format.

In order to complete this analysis, we must also assume that this preparation process also causes the name-retrieval process to be faster when it is done through the identical picture twice than through different pictures with the same name. (This advantage may deteriorate with time, as is suggested by the marginal Trial type by ISI Interaction here and the significant Trial type by ISI interaction in Experiment 2). Through this set of assumptions, the Orientation by Orientation of the first picture interaction as well as the overall pattern of the data are explained in terms of operations or processes rather than in terms of representations of specific visual information. This type of explanation, in general, seems much more consistent with the fact that there were similar effects on both identical and same-name trials, as well as with the fact that there was a trend towards a similar interaction with different trials.

Our argument that normalization processes, such as mental rotation, can be primed (i. e., prepared for in advance) has parallels elsewhere in the literature (e. g., Cooper & Shepard, 1973; Jolicoeur, 1990 b; Koriat & Norman, 1988; Robertson, Palmer, & Gomez, 1987). In these studies faster mental rotation was found if the orientation of the second of two patterns presented in close temporal succession was the same as, or similar to, the orientation of the first pattern. However, there is an argument as to whether this occurs because a frame of reference or an image was transformed during the task. Some have argued that the advantage for similarly oriented patterns occurs because a frame of reference that was rotated to align with the first stimulus can remain at that orientation until the second stimulus is presented and thus decrease processing time (Jolicoeur, 1990 b; Robertson et al., 1987). In contrast to the frame-rotation hypothesis, it has been suggested that the faster performance with similarly oriented patterns depends on an image-rotation process in which the second stimulus is brought into alignment with the orientation of the first stimulus (Cooper & Shepard, 1973; Koriat & Norman, 1988). Koriat and Norman (1988) argue that the process is one of image alignment rather than of frame rotation, because they found the advantage only for highly similar forms (but see Jolicoeur, 1990 b; Robertson et al., 1987). They proposed that a low-level template-like representation of the second stimulus is matched to a low-level representation of the first stimulus and that the matching process operates only if the two stimuli are the same form.

While our experiments were not designed to adjudicate between these two hypotheses, our results are not consistent with Koriat and Norman's (1988) model. The effects of orientation congruence in our experiment occurred for both the identical and the same-name objects. Furthermore, there was a trend towards a similar result on the different trials.

Experiment 2

As was mentioned above, one of the cornerstones of Ellis and Allport's (1986) proposal is that at longer ISIs the VIEW code disappears, thus eliminating any advantage for the identity condition. In Experiment 1 we did not find this result, and it is this particular aspect of the data that seems to be most problematic for their model. It is possible, however, that the longest ISI we used was not long enough for the VIEW code to decay. To examine this possibility, in Experiment 2 we added a 5-s ISI to the 100-ms and 2-s conditions. We must also acknowledge that the 120° rotation we used may have been extreme. For example, research by Jolicoeur (1985) shows that 120° rotations are particularly deleterious to object naming. Given these findings, as well as the fact that the rotations used by Ellis and Allport (1986) were rather small in comparison to those used in our Experiment 1, our second experiment used only 60° rotations in the frontal plane. Although pictures oriented at 60° are not named as quickly as upright depictions, they are named substantially faster than pictures oriented at 120° (Jolicoeur, 1985).

Method

Subjects. Ninety-six subjects from introductory psychology classes at the University of Western Ontario received course credit for participating in this experiment. All were naïve with respect to the purpose of the experiment.

Stimuli and apparatus. Except for the fact that the rotated objects were rotated 60° clockwise from the upright, the stimuli and apparatus were the same as those used in Experiment 1.

Design and procedure. As in Experiment 1, ISI was a between-subjects factor and 32 subjects were randomly assigned to each of the 100-ms, 2-s, and 5-s ISI conditions. All other aspects of the design and procedure were identical to those in Experiment 1.

Results

Same trials. The dependent measure used for all analyses was the mean reaction time for correct trials. The mean reaction time as a function of Trial type (including different trials), Orientation, and ISI is shown in Table 2. An ANOVA was performed on the *same* responses in which the between-subjects factor was ISI (100 ms, 2 s, 5 s) and the within-subjects factors were Trial type (identical or same name), Orientation (same or different) and the Orientation of the first picture in a pair (0° or 60°). The analysis yielded a significant main effect of ISI, $F(1,93) = 6.37$, $MS_e = 91683.6$, $p < .003$. Newman-Keuls tests showed that the 5-s ISI produced longer matching latencies (M = 600 ms) than either the 2-s ISI (M = 523 ms) or the 100-ms ISI (M = 514 ms) (both p values <.05). The matching latencies for the 100-ms and 2-s ISIs did not differ. The main effect of Trial type was again statistically reliable, $F(1,93) = 52.24$, $MS_e = 4163.7$, $p < .001$, with identical trials (M = 529 ms) being faster than same-name trials (M = 563 ms). In this experiment the Trial type by ISI interaction was significant, $F(2,93) = 4.31$, $MS_e = 4163.7$, $p < .02$. With increasing ISI the difference between identical and same-name trials decreased, although according to simple-effects analyses, at each ISI identical trials were faster than same-name trials (all p values <.05 or better).

Table 2. The mean reaction time as a function of trial type, orientation and ISI in Experiment 2

	Trial type					
	Identical		Same name		Different name	
	Orientation		Orientation		Orientation	
	Same	Different	Same	Different	Same	Different
ISI						
100 ms	474	504	536	542	575	594
2 s	496	514	538	543	563	577
5 s	583	603	598	618	623	629

The main effect of Orientation was also again statistically reliable, $F(1,93) = 15.78$, $MS_e = 3285.6$, $p < .001$, with trials in which both pictures were in the same orientation ($M = 538$ ms) being faster than trials with differently oriented pictures ($M = 554$ ms). The Orientation of the first picture also produced a reliable main effect, $F(1,93) = 4.97$, $MS_e = 6915.1$, $p < .03$. Trials with a 60° picture presented first were slightly faster ($M = 541$ ms) than trials with an upright picture presented first ($M = 550$ ms). The interaction between Orientation and Orientation of the first picture was not statistically reliable at the conventional level of significance, $F(1,93) = 2.39$, $MS_e = 3191.0$, $p = .13$, in this experiment. Figure 3 shows, however, the means as a functions of orientation (same or different) and which orientation came first. Because this interaction was reliable in Experiment 1 and is theoretically interesting, we performed simple-effects analyses. The analyses indicated that for trials with pairs in the same orientation, 0° trials ($M = 539$ ms) were not different from 60° trials ($M = 536$ ms), $F(1,93) = 0.190$, *ns*. For trials in which the pairs were in different orientations, however, those in which the first picture was oriented at 60° ($M = 546$ ms) were faster than those in which the first picture was at 0° orientation ($M = 562$ ms), $F(1,93) = 7.77$, $p < .01$. Different-orientation trials were also slower than same-orientation trials if the orientation of the first picture was 0°, $F(1,93) = 16.13$, $p < .001$; but evidence for this effect was somewhat weaker if the first picture was oriented at 60°, $F(1,62) = 2.90$, $p < .10$.

Different trials. As in the analysis of the *same* trials, the dependent measure was the mean reaction time for correct trials. An ANOVA was performed on the different responses in which the between-subjects factor was ISI (100 ms, 2 s, or 5 s) and the within-subjects factors were Orientation (same or different) and the Orientation of the first picture in a pair (0° or 60°).

The main effect of Orientation was significant, $F(1,93) = 8.42$, $MS_e = 1920.2$, $p < .005$. Trials in which the pictures were in the same orientation were faster ($M = 587$ ms) than trials in which the pictures were differently oriented ($M = 600$ ms). The main effect of the orientation of the first picture $F(1,93) = 6.46$, $MS_e = 2251.2$, $p < .05$, was also significant. Trials in which the first picture was oriented at 60° were faster ($M = 588$ ms) than trials in which the first picture was oriented at 0° ($M = 600$ ms) as in the *same* trials.

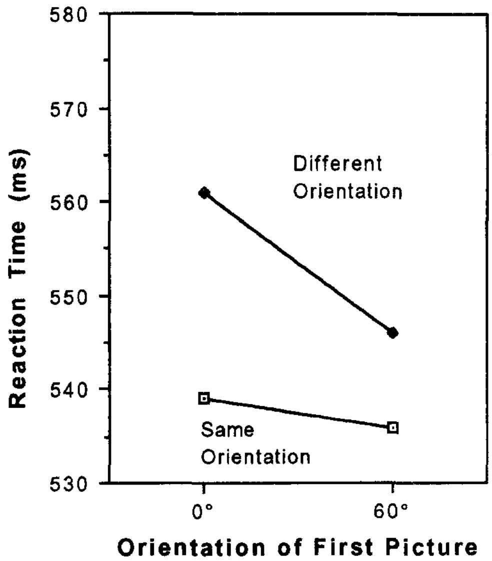

Fig. 3. The mean reaction time as a function of the orientation factor (same or different) and the orientation of the first picture (0° or 60°) in a picture pair in Experiment 2

Errors. The overall error rate was 4.5%. This included trials on which, for each subject, the latency to respond was more than 3 *SD*s above the mean for that subject. The error rate was 5% for *same* trials and 4% for *different* trials.

Discussion

In most important ways the results of Experiment 2 replicated those of Experiment 1. To begin with, as in Experiment 1, there was an overall effect of orientation such that pictures in the same orientation were matched more rapidly than were pictures in different orientations. Again, however, the orientation factor did not interact in any way with ISI nor was there evidence of a three-way interaction between Orientation, ISI, and Trial type. Rather, the difference between the same and different orientation trials in the identical condition seems to be statistically equivalent to that in the same-name condition and it shows very little tendency to disappear, even with a 5-s ISI. Thus, these results also provide very little support for the idea that the difference between the identity and the rotated trials is due to VIEW codes for the same objects in identical orientations that disappear at longer ISIs.

We also found, as in Experiment 1, that matching latency for same-name trials was greater than for identical trials. In Experiment 1 we noted that there was a trend for Trial type to interact with ISI in such a way that at longer ISIs the advantage of identical trials over same-name trials decreased. In the present experiment this interaction was significant and suggested the same pattern. The difference between same-name and identical trials decreased from 50 ms at the 100-ms ISI, to 35 ms at the 2-s ISI, and to 16 ms at the 5-s ISI. If one adopts a representation explanation for these data, the conclusion would have to be that over time there is a gradual loss of the specific visual information that gives the identical trials their advantage. In Ellis and Allport's (1986) terminology, this must be due to a decay of the (supposedly permanent) OBJECT code. This would seem to be a further weakness in the Ellis and Allport model.

In our conceptualization, the interaction between trial type and ISI can be explained in terms of some decay in the preparation for executing the operations used to access the name of a pictured object. We have argued that in the accessing of the name of an object, various operations are applied to the input image. If, on *same* trials, the second picture is identical with the first, these operations will have been primed, leading to a shorter latency in picture matching than when the pictures differ. We suggest that at longer ISIs the priming of these operations may decay, decreasing the difference between identical and same-name trials.

A slight difference between the two experiments was in the Orientation by Orientation of the first stimulus interaction. In Experiment 1, this interaction was significant. In Experiment 2, the interaction was not significant, but had the following form: overall, there was an advantage if the first picture to be matched was oriented at 60° rather than 0°; further, if the two pictures in a trial were in the same orientation, matching was equally fast; but if the picture pairs in a trial differed in orientation, matching was more rapid if a picture at 60° was followed by one at 0° than vice versa. In other words, the 0°–0°, 60°–60° and 60°–0° conditions yielded similar RTs with only the 0°–60° condition being different. Our account of this pattern of results would be quite similar to that offered for the Orientation by Orientation of the first picture interaction that we found in Experiment 1. The only difference between the results of the present experiment and the previous one concerns the trials on which both pictures were in the same orientation. In Experiment 1 we found that the 0°–0° sequence was faster than the 120°–120° sequence, while in Experiment 2 the 0°–0° sequence did not differ from the 60°–60° sequence. Apparently, the preparation afforded by subjects first seeing a 60° rotation is more effective than it is for 120° rotations.

General discussion

Our experiments were motivated by Ellis and Allport's (1986) argument for a separate VIEW code, which represents the particular orientation of an object, but decays within a couple of seconds. Although we too found what could be considered to be evidence for a view-specific representation, in that trials in which the picture pairs were in the same orientation were faster than trials in which the pictures were in different orientations, we did not find a significant decrease in the size of this effect as a function of ISI. Nor did we find that the size of this effect was significantly different for identical objects and same-name objects. According to Ellis and Allport's model, the matching of same-name objects does not involve orientation-specific representations. Our experiments differ from Ellis and Allport's in three ways. First we had complete counterbalancing of trial types and orientations. Second, we used line drawings and they used photographs. Third, our rotations were in the frontal plane, while Ellis and his colleagues used rotation in depth. We shall consider each of these differences in turn.

As was noted earlier, Ellis and colleagues (Ellis & Allport, 1986; Ellis et al., 1989) apparently did not include a condition in which same-name objects had different orientations. Consequently, orientation could be a reasonably reliable cue to the correct response. This could occur because objects with the same orientation (in their identity and same-name conditions) would always require a "same" response. Objects with a different orientation, however, would require a "same" response in the rotated condition and a "different" response in the different condition. The result is that seeing objects in the same orientation could create a bias toward "same" responses, while seeing objects in different orientations may create a bias to respond "different." The result would be an artificially created advantage for the identity pairs over the rotated pairs. Because we also included same-name, rotated pairs, and did obtain a identity-rotated difference in the identical condition, this particular bias explanation of Ellis and Allport's (1986) effect seems unlikely. What is more important, however, is that we can think of no reason why the inclusion of this condition could have caused the major difference between our results and theirs, that is, the maintenance of the identity advantage at longer ISIs in our studies.

With respect to the issue of line drawings versus photographs, as noted, Bartram (1976) argued that there was more evidence of a difference between identity and rotated trials with line drawings than with photographs. Bartram raised the possibility that the lack of difference when photographs were used may have occurred because of the presence of surface features such as texture and grey level, which, he proposed, are essentially orientation independent. That is, the assumption was being made that both identical pictures and rotated pictures of the same object shared these features and, hence, matches made on this basis could be made equally rapidly. If this argument is correct, it suggests that this type of problem might arise whenever photographs are used (as in the research of Ellis and colleagues). The use of such a strategy would not, however, explain the basic discrepancy between the present data (where the identity-rotated difference was observed at all ISIs) and Ellis and Allport's (1986) data, where the difference disappeared at the longer ISI. Following Bartram's logic, the use of this type of a strategy would have the effect of decreasing or eliminating that difference. Thus, in order to explain the difference between the present results and those of Ellis and Allport on this basis, one would have to argue further that this strategy was used by Ellis and Allport's subjects *only* with long ISIs. Such an argument would seem hard to support because, if anything, features such as texture and grey level should be more available at short ISIs.

In fact, the assumption that surface features such as texture and grey level are orientation independent seems wrong. It is more likely that exactly the opposite is true: that is, that identity and rotated pictures actually do differ substantially on features like texture and grey level. If so, one could argue further that these features may have actually served as at least a partial basis for the identity advantage that Ellis and Allport (1986) observed at the short ISI (rather than a VIEW code). Perhaps the particular pattern of surface characteristics decays rapidly, leading to the loss of the view-specific advantage. Thus, at the longer ISI, this information would not be available, leading to a lack of a difference between the identity and the rotated conditions.

If this argument is correct, it would also account for the other two effects that Ellis and Allport used to argue for the existence of a VIEW code. That is, evidence for the VIEW code also comes from the reduction of the identity advantage when a mask is inserted between the first and second stimuli and when the first and second stimuli differ in size (Ellis & Allport, 1986; Ellis et al., 1989). If it is the case that the identity advantage in their studies was mainly due to the use of features such as texture and grey level, it is fairly straightforward to see how these changes would reduce or eliminate the usefulness of these features, hence reducing or eliminating the identity advantage.

If the identity advantage found by Ellis and Allport (1986) was essentially due to the use of texture and grey-level features, it raises an obvious question with respect to the present data. That is, why would there ever be an identity advantage for line drawings that do not have these types of feature? The answer to this question may be found in the third difference between the present studies and Ellis and Allport's: the nature of the rotations used.

Many studies have found that rotations of common objects both in depth (e.g., Humphrey & Jolicoeur, 1988; 1993; Palmer et al., 1981) and in the frontal plane (e.g. Jolicoeur, 1985; reviewed in Jolicoeur, 1990a) away from a canonical orientation produce increases in naming latency. On the basis of this literature, however, it is not clear which type of orientation change is, in general, most disruptive. Recent research by Langdon, Mayhew, and Frisby (1991), using a difference-rating task, however, does suggest that rotations in the picture plane produce a view less similar to a reference view than do equivalent rotations in depth. In their task subjects were required to judge the perceived difference between a reference view of an unfamiliar object and rotated views of the object. Different axes of rotation from the reference view were examined. One depth rotation, called CYLINDER by Langdon et al., was most like that used by Ellis and his colleagues. This type of depth rotation was judged to produce views of the object that were more similar to (i.e., less different from) the standard view than did rotations of the same magnitude in the frontal plane. It is possible, then, that the normalization processes necessary to cope with the 120° rotations, or even with the 60° rotations, used in the present studies were much more demanding than the normalization operations necessary to deal with the depth rotations used by Ellis and his colleagues. Thus, it is reasonable to argue that these processes were the main cause of the identity advantage in the present studies, while playing only a minor role in Ellis and Allport's (1986) experiments.

In conclusion we note that in terms of the general issue of whether long-term object representations are viewpoint independent or viewpoint dependent, our results are most consistent with the latter proposal. Our results suggest that the long-term representation of objects is in a familiar, upright orientation. Like others (e.g., Jolicoeur, 1985; Tarr & Pinker, 1989; Ullman, 1989), we have also proposed that object recognition involves normalization operations such as mental rotation. We also suggest that our data are consistent with the argument that the use of such operations can be primed (see also Cooper & Shepard, 1973; Jolicoeur, 1990b; Koriat & Norman, 1988; Robertson et al., 1987), a factor that obviously must be taken into account in a consideration of data from picture-matching tasks.

Acknowledgements. This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada to G. K. Humphrey (Grant no. A1643) and S. J. Lupker (Grant no. A6333). G. K. Humphrey was also supported by a sabbatical-leave award from the James McKeen Cattell Foundation, which is gratefully acknowledged. This research was presented at the 32nd Annual Meeting of the Psychonomic Society. Thanks to Andrew Herbert, Brent Philips, and Annette St Pierre of the University of Western Ontario for running the experiments. The order of authorship is arbitrary.

References

Bartram, D. J. (1976). Levels of coding in picture–picture comparison tasks. *Memory & Cognition, 4,* 593–602.

Biederman, I. & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20,* 38–64.

Bülthoff, H. H., & Edelman, S. (1990). Recognizing objects from novel viewpoints. Presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida.

Bülthoff, H. H., Edelman, S., & Sklar, E. (1991). Mapping the generalization space in object recognition. Presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida.

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.

Ellis, R., & Allport, D. A. (1986). Multiple levels of representation for visual objects: A behavioural study. In A. G. Cohn and J. R. Thomas (Eds.), *Artificial intelligence and its applications* (pp 245–257). New York: John Wiley & Sons.

Ellis, R., Allport, D. A., Humphreys, G. W., & Collis, J. (1989). Varieties of object constancy. *Quarterly Journal of Experimental Psychology, 41A,* 775–796.

Humphrey, G. K., & Jolicoeur, P. (1988). Visual object identification: Some effects of image foreshortening and monocular depth cues. In Z. W. Pylyshyn (Ed.), *Computational processes in human vision: An interdisciplinary perspective* (pp. 430–444). Norwood, NJ: Ablex.

Humphrey, G. K., & Jolicoeur, P. (1993). An examination of the effects of axis foreshortening, monocular depth cues, and visual field on object identification. *Quarterly Journal of Experimental Psychology A,* 137–159.

Humphrey, G. K., & Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Canadian Journal of Psychology, 46,* 170–190.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13,* 289–303.

Jolicoeur, P. (1990 a). Identification of disoriented objects. A dual-systems theory. *Mind and Language, 5,* 387–410.

Jolicoeur, P. (1990 b). Orientation congruency effects on the identification of disoriented shapes. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 352–364.

Kelter, S., Grötzbach, H., Freiheit, R., Höhle, B., Wutzig, S., & Diesch, E. (1984). Object identification: The mental representation of physical and conceptual attributes. *Memory & Cognition, 12,* 123–133.

Koriat, A., & Norman, J. (1988). Frames and images: Sequential effects in mental rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 93–111.

Langdon, P. M., Mayhew, J. E. W., & Frisby, J. P. (1991). In search of "characteristic view" 3D object representation in human vision using ratings of perceived differences between views. In J. E. W. Mayhew and J. P. Frisby (Eds.), *3D model recognition from stereoscopic cues* (pp. 244–248). Cambridge, MA: MIT Press.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, 200,* 269–294.

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Lawrence Erlbaum.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18,* 1–63.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature, 343,* 263–266.

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review, 74,* 392–409.

Price, C. J., & Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *Quarterly Journal of Experimental Psychology, 41 A,* 797–828.

Robertson, L. C., Palmer, S. E., & Gomez, L. M. (1987). Reference frames in mental rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 368–379.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology, 19,* 280–293.

Rock, I., DiVita, J., & Barbeito, R. (1981). The effect of form perception of change of orientation in the third dimension. *Cognitive Psychology, 19,* 280–293.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233–282.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32,* 193–254.