ELSEVIER

Acta Psychologica 102 (1999) 221–245

acta psychologica

Achieving visual object constancy across plane rotation and depth rotation

Rebecca Lawson [1]

Department of Psychology, University of Liverpool, Bedford Street South, Liverpool, L69 7ZA, UK

Received 16 April 1998; received in revised form 26 October 1998; accepted 27 November 1998

Abstract

Visual object constancy is the ability to recognise an object from its image despite variation in the image when the object is viewed from different angles. I describe research which probes the human visual system's ability to achieve object constancy across plane rotation and depth rotation. I focus on the ecologically important case of recognising familiar objects, although the recognition of novel objects is also discussed. Cognitive neuropsychological studies of patients with specific deficits in achieving object constancy are reviewed, in addition to studies which test neurally intact subjects. In certain cases, the recognition of *invariant features* allows objects to be recognised irrespective of the view depicted, particularly if small, distinctive sets of objects are presented repeatedly. In contrast, in most situations, recognition is sensitive to both the view in-plane and in-depth from which an object is depicted. This result suggests that *multiple, view-specific, stored representations* of familiar objects are accessed in everyday, entry-level visual recognition, or that *transformations* such as mental rotation or interpolation are used to transform between retinal images of objects and view-specific, stored representations. © 1999 Elsevier Science B.V. All rights reserved.

PsycINFO classification: 2323; 2343

Keywords: Visual perception; Perceptual constancy; Recognition

[1] Tel.: +44-151 794 3195; fax: +44-151 794 2945; e-mail: rlawson@liverpool.ac.uk.

PII: S 0 0 0 1 - 6 9 1 8 (9 8) 0 0 0 5 2 - 3

1. Introduction

Visual object constancy is the ability to determine the identity of an object from its variable image. The image varies considerably following plane and depth rotations of the object, due to viewing the object from different positions. Our visual system is said to achieve object constancy because it can usually cope with such variation, by accurately recognising a range of retinal images as depicting the same object. For example, most people would recognise the left four pictures in Fig. 1 as depicting an iron, yet the pictures differ in size, outline shape and in the presence of features and parts.

Although we can achieve object constancy, we are sensitive to image variation and we can accurately report the size and orientation of an object. In addition, speed of object recognition is influenced by the familiarity of a given view, its similarity to views of other objects from which it must be discriminated, and its "goodness" – how well it depicts the object. An aerial view of a house may be recognised accurately, but recognition would usually be slow compared to a street-level view. This is because aerial views of houses are uncommon, similar to aerial views of churches and barns, and "poor" (the view hides the 3D structure of the house and many of its distinctive features). Finally, as it is ecologically important to achieve object constancy efficiently, the visual system has presumably been driven to optimise it, and we may be unaware of the true processing costs involved (see Rock, Schreiber & Ro, 1994).

In addition to achieving object constancy, the visual system must discriminate between stimuli which differ in semantically important ways. Improving the achievement of object constancy will often impair discrimination, so the achievement of these two functions will be in conflict, and the visual system must reach an appropriate compromise between them. If the visual system ignores much image variation, it will be easy to achieve object constancy (because the difference in size and shape between dachshunds and alsations can be ignored, to recognise them both as dogs), but it will then be harder to discriminate between different objects (wolves and alsations, which are visually similar but which have different semantic properties), and vice versa.

The achievement of object constancy can thus only be examined in relation to the difficulty of the discrimination task required of subjects. If we only have to distinguish a red cube from a metal cheese-grater and a yellow pool of paint, then simple surface or texture or shape information will accurately identify all three objects, irrespective of the view presented. The achievement of object constancy will be trivial. In contrast, under everyday viewing conditions, there are usually many objects which could be present in any situation, yet we rarely misidentify objects, even if the object is unlikely in that situation (a frog in our bedroom). Although difficult, discrimination is both rapid and accurate, so achieving object constancy is also hard. The difficulty of discrimination depends not just on the number of objects to be distinguished, but also on their similarity. Animals may be harder to recognise than manmade artefacts because, as a category, animals are more visually similar to each other than are artefacts (Humphreys, Riddoch & Quinlan, 1988).

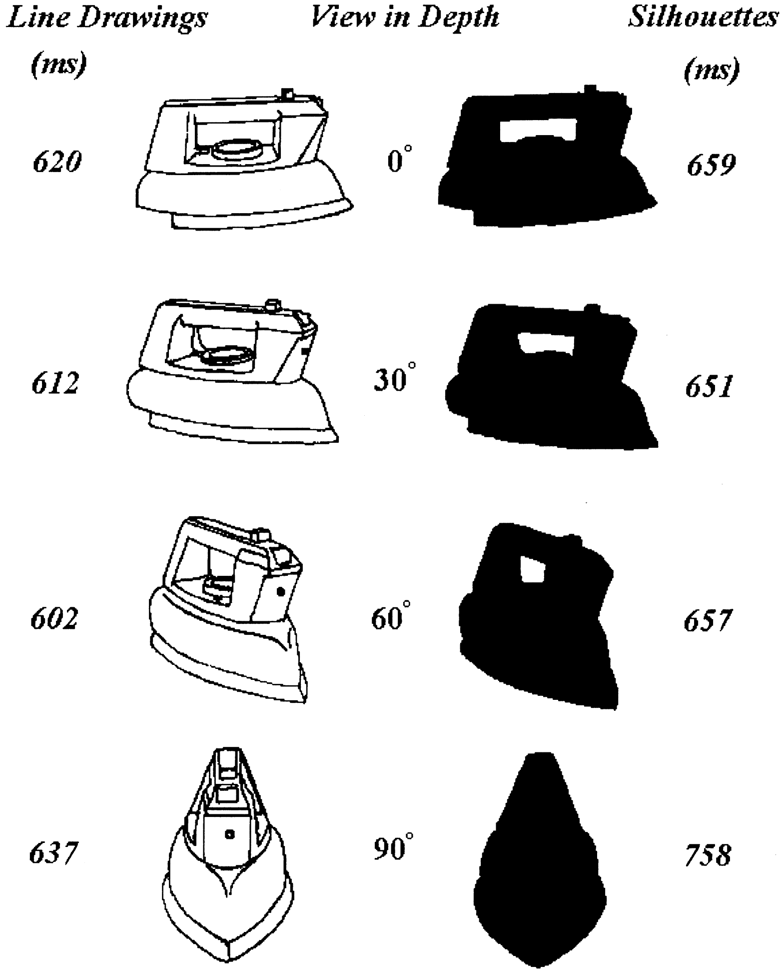

Fig. 1. Eight different views of an iron. On the left are line drawings, on the right are matched silhouettes. The iron rotates in depth through 0° (top), 30°, 60° and 90° (bottom) views. The 0° views fully reveal the main axis of elongation of the object. The 90° views are so foreshortened that the main axis of elongation of the image no longer coincides with the main axis of elongation of the object, but instead is the vertical axis. The 60° view was rated as the most canonical, typical view of the iron. On the left are the RTs to verify the identity of the line drawings in a speeded word-picture verification task; on the right are the analogous RTs for silhouettes, from Experiment 2 of Lawson & Humphreys (1999).

Current theories of human visual object recognition acknowledge the importance of accounting for our ability to achieve object constancy and our ability to discriminate between different classes of objects. However, there is little consensus between these accounts about the representations and processes which are involved (e.g. Biederman, 1987; Bülthoff & Edelman, 1992; Edelman & Weinshall, 1991; Hummel & Stankiewicz, 1997; Jolicoeur, 1990; Lowe, 1987; Marr, 1982; Tarr & Pinker, 1989; Ullman, 1989).

There are three classes of account of the achievement of object constancy. *Invariant features* accounts suggest that invariant features can be used to distinguish most views of one object from most views of all other objects. A feature is only unique to an object in the context of the set of objects from which that object is to be distinguished, i.e. given a particular discrimination task. *Multiple view* accounts suggest that the visual system stores several representations of each object, and that a given view is matched to the nearest view-specific, stored representation. *Transformation* accounts propose that the retinal image can be transformed to reduce differences between the image and a view-specific, stored representation.

Although representation (multiple view) and process (transformation) accounts differ conceptually, they are hard to distinguish empirically. Representations and processes cannot be examined independently of each other. A pattern of performance resulting from a certain representation being stored can be exactly replicated by specifying a particular process. Any well-specified theory of object recognition must describe both the representations stored by the visual system and the processes employed to access those representations. Furthermore, as more interactive models of human information processing become popular (e.g., neural network models), the distinction between representation and process is likely to break down. I have, however, discussed representation and process accounts separately since most theoreticians distinguish between the two. It is worth noting here that there are clear, object-specific effects in recognition. View-specific effects in priming studies are tied to subject's experience with particular objects (e.g., Jolicoeur & Milliken, 1989, and Lawson & Humphreys, 1998a, for plane and depth rotations, respectively). Object-specificity is typically associated with access to different representations rather than the use of different processes.

Invariant features, multiple view and transformation accounts are not incompatible, and it is likely that the visual system employs them all to some degree to achieve object constancy. For example, Jolicoeur (1990) proposed an account of the effects of plane rotation on picture recognition which included all three classes of account. Jolicoeur suggested that plane rotated views of objects are often transformed to match a stored, upright view (transformation account), although plane-disoriented views of objects may also be stored, allowing direct matching of plane rotated views (multiple-views account). In addition, Jolicoeur (1990) suggested that a functionally distinct, feature-based route was also available which enabled pictures to be recognised by matching simple attributes such as colour, texture or shape, many of which are orientation-invariant (invariant features account). This latter route is most useful if distinctive objects are presented, or if a small set of stimuli are presented repeatedly, enabling subjects to learn which features are invariant. Note that subjects do not always use invariant information when it is available, and they may need explicit encouragement or training to take advantage of it (see Takano, 1989).

Similar to Jolicoeur's (1990) account, Tarr and colleagues (e.g., Tarr, 1995; Tarr & Pinker, 1989, 1990) have proposed a multiple-views-plus-transformation account. They suggest that if an object is seen from several, distinct views, representations of these different views will often all be stored. If a view similar to a stored view is presented, it can be matched directly. Otherwise, the image must undergo a time-consuming transformation before it can be matched to a stored representation.

In contrast, only the invariant features and multiple views accounts are included in the theory of object recognition proposed by Biederman and colleagues (Biederman, 1987; Biederman & Gerhardstein, 1993). They hypothesise that the visual system derives a geon structural description from input images. This representation specifies coarsely coded spatial relations (such as left-of, above, below) between view-invariant, volumetric primitives (geons), and is quite insensitive to the view in depth of an object. Nevertheless, multiple geon structural descriptions of a given object may need to be stored if, for example, an important part of the object is occluded in some views, as this would alter the structural description derived from those views. The theory does not posit any image transformation processes. A similar account has been proposed by Hummel & Stankiewicz (1997).

The aim of this paper is to critically assess and integrate the evidence for distinct processes and representations being employed in achieving object constancy across plane and depth rotation, and to determine when invariant features can provide an alternative means of achieving object constancy. To date, there has been insufficient cross-talk between different research areas. For example, there has been considerable neuropsychological empirical research into the achievement of object constancy across depth rotation (Humphreys & Riddoch, 1984; Warrington, 1982; Warrington & Taylor, 1973; see Lawson & Humphreys, 1998b), but there has been little contact between this research and directly related research which has tested non-brain-damaged subjects.

2. The achievement of object constancy across plane rotation

2.1. Is mental rotation involved in mirror-image discrimination or recognition tasks?

Mirror-image discrimination tasks include simultaneous picture matching, where mismatch trials present two mirror-image versions of the same stimulus (Shepard & Metzler, 1971), and left–right direction-of-facing tasks, where subjects must determine the direction a familiar object such as a cow would face, were it upright (Jolicoeur, 1988; Jolicoeur, Corballis & Lawson, 1998; McMullen & Jolicoeur, 1990). These tasks produce approximately linear increases in RTs across increasing plane disorientations. Such tasks are generally assumed to employ mental rotation. Mental rotation is an analogue transformation process that requires more time to rotate an object through a greater angle (see Shepard & Cooper, 1982).

As in mirror-image discrimination tasks, in naming tasks, increasing the plane disorientation of pictures of familiar objects correspondingly increases RTs (Jolicoeur, 1985; Jolicoeur, Corballis & Lawson, 1998; Jolicoeur & Milliken, 1989; McMullen & Jolicoeur, 1990; Murray, 1995). Researchers usually test views rotated 0°, 60°, 120° and 180° away from a canonical, upright view. The increase in RTs across 0° (upright), 60° and 120° views is approximately linear, whilst 180° (inverted) views are often named faster than would be predicted by extrapolating from performance with 0°, 60° and 120° views (see Jolicoeur, 1990). Accounts of the naming of 180° views emphasise the unique qualities of this view (for example, spatial relations are simply left–right and up–down reversed relative to an upright view, and the positions of axes of symmetry and the main axis of elongation are the same as in the upright view, see Hummel & Biederman, 1992; Jolicoeur, 1990; Murray, 1997).

Disregarding the special case of 180° views, it has been widely assumed that mental rotation is used by subjects to recognise plane rotated pictures (e.g. Jolicoeur, 1985; 1990; Murray, 1997; Tarr & Pinker, 1989). This is an appealingly parsimonious proposal, since a single image transformation (mental rotation) would compensate for plane disorientation, regardless of the task (recognition or mirror-image discrimination).

Indirect support for this hypothesis comes from a number of studies which have found similar effects of plane rotation across recognition and mirror-image discrimination tasks. For instance, Jolicoeur (1985, 1988) found similar increases in RTs across plane rotations from 0° to 120°, when subjects named pictures compared to when subjects decided which direction the object faced, were it upright. In addition, Jolicoeur (1988) reported that plane rotation effects on subject's naming RTs correlated with plane rotation effects for the same subjects RTs to decide which way an object faced. Murray (1997) found similar increases in RTs for plane disorientation when subjects named objects compared to when they made mirror/normal judgements to the same objects (after subjects were trained to recognise one mirror-image version of each stimulus as the "normal" version). Finally, Tarr and Pinker have investigated the effects of plane rotation for novel objects using recognition and mirror-image discrimination tasks. Tarr and Pinker (1989, 1990) presented 2D stimuli whilst Tarr (1995) presented analogous 3D stimuli. In all cases, they reported broadly similar increases in RTs with plane disorientation across recognition and mirror-image discrimination tasks. This was the case even when subjects in the recognition task were given explicit instructions that mirror-image versions of stimuli were to be responded to in the same way as "normal" versions, and subjects were given practice at doing this.

All of the above evidence only indicates that plane rotation has broadly similar effects on recognition and mirror-image discrimination tasks. Note, too, that object recognition does not normally require mirror-image discrimination (although there are a few exceptions, such as distinguishing left from right shoes and gloves), and that long-term priming of the visual system appears to be invariant to the mirror-image version of an object presented (Biederman & Cooper, 1991; Lawson & Humphreys, 1996, 1998a; Stankiewicz, Hummel & Cooper, 1998; although the latter paper did report reduced short-term priming with mirror-image compared to identical prime views).

Furthermore, with familiar objects, plane rotation effects on mirror-image discrimination do not reduce with practice (Jolicoeur, 1988), whereas repeated naming of familiar objects does reduce plane rotation effects (Jolicoeur, 1985, 1988; Jolicoeur & Milliken, 1989; Lawson & Jolicoeur, 1999a; McMullen & Jolicoeur, 1992). This is, though, not strong evidence that mental rotation is employed only for mirror-image discrimination and not for recognition. Jolicoeur (1990) suggested that with practice, subjects may learn to recognise objects using features invariant to plane rotation. Such features would allow subjects to recognise plane rotated stimuli directly, without needing to transform images, so plane rotation effects on naming would reduce with practice. In contrast, these features would not distinguish between mirror-images, so image transformation would still have to be used.

More direct tests are necessary to examine the role of mental rotation in object recognition. My colleagues and I have used one such test (Jolicoeur, Corballis & Lawson, 1998). We compared performance on naming and direction of facing tasks, for 0°, 120° and 240° views of familiar objects. As expected, responses were slower to 120° and 240° than to 0° views for both tasks. We examined whether this plane rotation effect interacted with the apparent direction of plane rotation of the object. In the first study, we induced a motion-after-effect in subjects prior to presenting each single, static view of an object, so subjects saw the object appear to rotate either clockwise or anticlockwise. In the second study, each trial consisted of eight brief views of the same object. Each successive view was rotated by a further 2° in the plane, so that on a given trial the object appeared to rotate 14° either clockwise or anticlockwise.

In both studies, for the direction of facing task, subjects were faster when the object appeared to rotate towards, rather than away from, the upright position. The direction of facing task required mirror-image discrimination, and presumably subjects used mental rotation to do the task. Our results suggest that the direction of both a motion-after-effect and the rotation of an object influence the direction that subjects choose to mentally rotate. In contrast, in the naming task, subjects were not influenced by either the direction of the motion-after-effect or the direction of rotation of the object. This suggests that mental rotation is *not* involved in the recognition of plane rotated views of objects.

A second strand of evidence that mental rotation is not employed in recognition tasks comes from an experiment in which we examined the recognition of brief, masked, plane rotated pictures of familiar objects (Lawson & Jolicoeur, 1999a). It has often been claimed that plane rotation effects on picture naming are linear for views rotated between 0° and 120° (e.g. Murray, 1997). However, studies supporting this claim usually only test 0°, 60° and 120° views within this range. We tested the effects of plane rotation more finely, by presenting views rotated successively by 30°, between 0° and 180°. We found a consistently non-monotonic pattern of performance (see Fig. 2). The 30°, 90°, 150° and 180° views were recognised more efficiently than would be predicted, given performance at 0°, 60° and 120° views. This pattern does not support a mental rotation account, which would predict approximately linear (and at least monotonic) effects of plane rotation, as are observed in typical mental rotation tasks which require mirror-image discrimination. Instead, a number of different factors may influence the ease of recognition of plane rotated views. For example, the recognition of 90° and 180° views may benefit from the position of the horizontal and vertical axes of these views being, respectively, perpendicular and coincident with their position in the upright, 0° view, whilst 30° views may benefit from relatively broad-tuned matching to a stored, upright representation of the object.

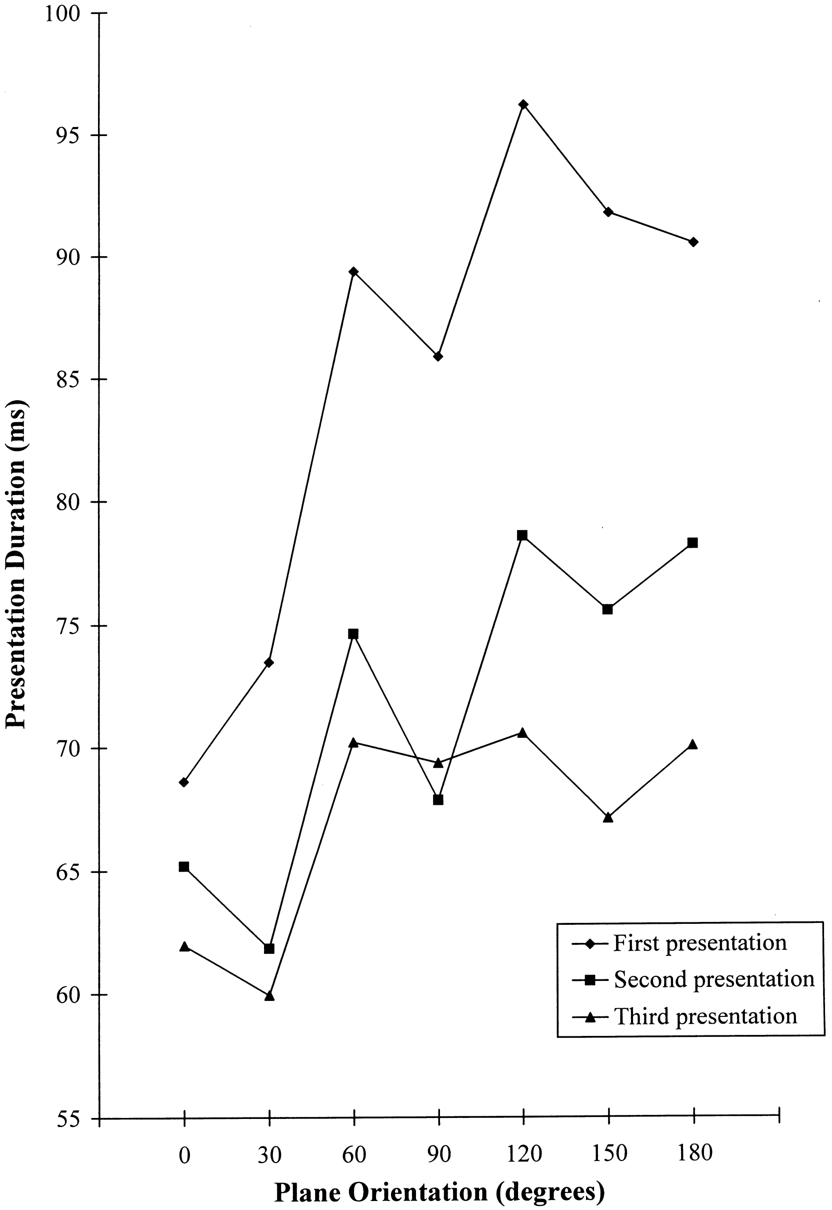

Fig. 2. The mean presentation duration required to correctly verify the identity of line drawings of familiar objects as a function of plane rotation (Lawson & Jolicoeur, 1999a). Stimuli were presented briefly at low contrast and were then masked. On each trial, an object was presented repeatedly at increasing presentation durations until it was correctly recognised. Subjects identified stimuli by making an unspeeded selection from a written list of 126 names of objects. Each object was presented up to three times in separate blocks, each time at a different plane rotation.

Third, neuropsychological evidence suggests that for plane-rotated views, recognition performance can dissociate from performance at mirror-image discrimination (which is typically assumed to involve mental rotation). Farah and Hammond (1988) reported a right-hemisphere lesioned patient, RT, who was very poor at neuropsychological tests of mental rotation, such as the Ratcliff mannikin task. In this task, the patient must specify which hand is depicted as holding a disk, for figures shown upright and inverted, and from front and back views (Ratcliff, 1979). In contrast, RT showed little additional impairment when naming inverted compared to upright views of familiar objects (although he was poor at recognising upright views of familiar objects). When controls were shown stimuli briefly, to equate their performance on upright views to that of RT, there was no difference between the controls and RT in the recognition of inverted views. This suggests that mental rotation is not *necessarily* required to recognise plane disoriented objects. Nevertheless, since RT was worse at naming objects than controls, these results are still consistent with neurally intact subjects using mental rotation when recognising plane disoriented objects. Evidence against this proposal is provided by Turnbull and McCarthy (1996). Their patient RJ could reliably recognise objects, distinguish between two objects differing by only a small perceptual change, and distinguish between upright and inverted objects. In contrast, RJ was unable to discriminate between mirror-image versions of pictures of either novel or familiar objects. Similarly, Turnbull, Laws and McCarthy (1995) reported patient LG, who could recognise an object without being able to reliably determine its plane orientation (see also Solms, Kaplan-Solms, Saling & Miller, 1988).

Neuropsychological single case studies must always be treated with caution. Patients who perform well across a limited battery of simple tests are often found subsequently to be at least mildly impaired when tested under more stringent conditions. In addition, the recognition tasks may simply have been easier than the mirror image discrimination and orientation tasks on which patients showed greater impairments (although inspection of the examples of stimuli presented to RJ by Turnbull & McCarthy (1996) argue against this point). Further evidence is required before strong conclusions can be drawn, but these results suggest that we can recognise a plane-disoriented familiar object without knowing its plane orientation or the direction it faces and without being able to mentally rotate.

Taken together, the above evidence reveals superficial similarities between the effects of plane rotation on the recognition and mirror-image discrimination of familiar objects (Jolicoeur, 1985, 1988; Murray, 1997; Tarr, 1995; Tarr & Pinker, 1989, 1990). However, more direct evidence suggests that the visual system uses different processes to compensate for plane rotation in recognition and mirror-image discrimination tasks, with mental rotation used only in the latter task.

2.2. Double-checking as an account of plane rotation effects on object recognition

An alternative to a mental rotation account of plane rotation effects on recognition is the double-checking account proposed by Corballis (1988) and by De Caro and Reeves (1995). This suggests that for mirror-image discrimination, plane rotated views must undergo time-consuming transformation, but that most plane rotated views can be *recognised* just as efficiently as upright views. Plane rotated views are, though, named slower than upright views because subjects double-check to ensure that they have correctly recognised the object before making their response. Double-checking involves a transformation process which produces the characteristic plane rotation effects. When stimuli are repeated, subjects often realise that double-checking is unnecessary because their initial (orientation-invariant) recognition is usually correct. They therefore double-check less frequently and plane rotation effects reduce.

For the double-checking account to differ from other transformation accounts (such as the mental rotation account assessed in Section 2.1), double-checking must have little effect on accuracy. If double-checking did improve accuracy, then the double-checking account would simply suggest that objects are sometimes recognised by extracting orientation-invariant features (in common with almost any account of plane rotation effects) and that otherwise disoriented stimuli must be transformed before they can be recognised (as in the mental rotation account).

Pierre Jolicoeur and I have examined the double-checking account using a picture-word verification task (Lawson & Jolicoeur, 1998; see also Lawson & Jolicoeur, 1999a). We reasoned that double-checking should only have deleterious effects on performance in speeded tasks, when it would slow responses. In unspeeded tasks, subjects should not be disadvantaged if they double-check, because although double-checking increases RTs, it should not affect accuracy. In our studies, subjects saw brief, low contrast, masked pictures of familiar objects, which they identified by choosing from a set of written alternatives. Subjects were more accurate at recognising upright views than plane disoriented views. Thus even in an unspeeded task, performance was still influenced by plane orientation (see Fig. 2). At least some plane rotation effects cannot then be accounted for by double-checking.

It might be argued that the brief, masked presentation used in our studies would have prevented subjects from double-checking. If this were the case, then again no plane rotation effects would be predicted on the double-checking account, since all views should then have been recognised equally efficiently. Once again, this prediction was not supported by our results.

We could not directly compare plane rotation effects across speeded and unspeeded studies. Therefore, it is possible that double-checking does play *some* role in increasing RTs when identifying plane rotated views. Note, though, that we have found the same pattern of plane rotation effects in unspeeded and speeded tasks, both for initial recognition (Jolicoeur, 1985; Lawson & Jolicoeur, 1999a; for speeded naming and unspeeded verification, respectively) and for the object-specific reduction of plane rotation effects following practice (for speeded naming, see Jolicoeur & Milliken, 1989; for unspeeded verification, see Lawson & Jolicoeur, 1999a, see also Fig. 2). The most parsimonious account of this data is that the same process transforms plane rotated images in both speeded and unspeeded recognition tasks. There is no reason to posit an additional process such as double-checking.

2.3. Are plane disorientation effects eliminated when recognition is at the entry level or when highly distinctive stimuli are presented?

Hamm and McMullen (1998) claimed that plane disorientation effects are found only when subordinate recognition is tested and not when entry level or superordinate recognition is required- so only when alsations and trawlers are specifically identified as alsations and trawlers, and not when they are identified more generally as dogs and boats, or as animals and vehicles. Apparently inconsistent with this claim, strong effects of plane rotation are found in speeded naming tasks (e.g. Jolicoeur, 1985), and here subjects are usually assumed to recognise objects at the entry level. Instead, Hamm and McMullen suggested that plane rotation effects in naming are due to subjects recognising some stimuli at the subordinate level (Hamm and McMullen suggest around 50% of stimuli). In contrast to naming, in verification the experimenter sets the level of recognition of an object, and this is made explicit to the subjects. Using a speeded word-picture verification task, Hamm and McMullen found that plane disoriented views were verified slower than upright views only when a subordinate word was presented (e.g. alsation or trawler) and not when a entry level word was presented (e.g. dog or boat). They concluded that the representations supporting initial, entry-level recognition are invariant to plane rotation.

This claim is important, but there is evidence against it. Murray (1998) failed to replicate Hamm and McMullen's results for entry-level verification. Murray used a speeded word-picture verification task. She found that plane disoriented stimuli were verified slower than upright stimuli for both match and mismatch trials, when a visually similar set of items was tested. When a visually dissimilar set of items was tested, plane disorientation effects were again observed, but only for match trials.

Pierre Jolicoeur and I have found similar results in both speeded and unspeeded word-picture verification studies. In the speeded study, we found plane rotation effects on both match and mismatch trials for entry-level verification (Lawson & Jolicoeur, 1999b). In the unspeeded studies which we described in Section 2.2, there were clear effects of plane disorientation for entry-level verification (Lawson & Jolicoeur, 1999a; see Fig. 2). In addition, we have manipulated the visual similarity of the written distractors. For example, for the picture of a deer, the response alternatives presented either visually similar distractors (deer, goat or donkey) or visually dissimilar distractors (deer, bowl or bow). There were greater effects of plane rotation when similar distractors were presented, and plane rotation effects were found even with the dissimilar distractors (Lawson & Jolicoeur, 1998).

Why have effects of plane rotation been found in some studies of entry-level verification but not in others (see Table 1)? Two factors seem important. First, practise at recognising the stimuli, and second, task difficulty in terms of ease of discrimination between items. With practise, plane rotation effects reduce, both for speeded and unspeeded recognition (e.g. Jolicoeur, 1985; Lawson & Jolicoeur, 1999a, respectively; see also Fig. 2). This is probably because subjects learn to recognise some stimuli using view-invariant features. Such features are easiest to extract if subjects need only discriminate between a small set of visually dissimilar items. In Hamm and McMullen (1998) studies, only six entry-level objects were presented (car, boat, aircraft, bird, dog

Table 1
Word-picture verification studies presenting line drawings of familiar objects

	Task (un/speeded; with visually dis/similar stimuli tested; number of stimuli presented and number of exposures across both match and mismatch trials)	Entry-level effects of plane rotation
Hamm & McMullen (1998)	Speeded – dissimilar – 6[a] items seen 12 times, 6/exemplar	No
Lawson & Jolicoeur (1998)	Unspeeded – dis/similar – 126 items seen two or four times	Yes
Lawson & Jolicoeur (1999a)	Unspeeded – similar – 126 items seen two or three times	Yes
Lawson & Jolicoeur (1999b)	Speeded – dissimilar – 60 items seen eight times	Yes
Murray (1998)	Speeded – dis/similar – 18 items seen two times	Yes

[a] Two different subordinate-level exemplars of each of the six entry-level items were presented to each subject, e.g. for the dog, an alsation and a collie were both seen by a given subject.

and bug[a]), these were visually distinct, and the same stimuli were presented many times during the experiment. No plane rotation effects on entry-level verification were found. In contrast, for the remaining studies listed in Table 1, more stimuli were presented for fewer repetitions. Plane rotation effects on entry-level verification were found, and these effects were greater when more visually similar stimuli were presented.

Nevertheless, plane rotation effects will typically be weaker for verification compared to naming tasks, since discrimination is usually easier. In a typical verification task, relatively coarse or simple orientation-invariant information is often sufficient to decide whether a picture and a word match. This same information would not be precise enough to name the picture. If discrimination is made more difficult in the verification task (for example, by increasing the similarity of stimuli on mismatch trials or by increasing the similarity of response alternatives in an unspeeded task), then it will be harder to achieve object constancy across plane rotation. This is an example of the interaction between ease of discrimination and ease of achieving object constancy which was discussed in the introduction.

2.4. Summary: the achievement of object constancy across plane rotation

For mirror-image discrimination, mental rotation may be used to achieve object constancy across plane rotation. In contrast, for the recognition of familiar objects, neither mental rotation nor double-checking can adequately account for plane rotation effects, and these effects are found for entry-level as well as for subordinate recognition. The process which enables the visual system to compensate for plane disorientation has not yet been satisfactorily specified. *Transformation* processes which are promising alternatives to mental rotation include image alignment (Ullman, 1989) and view interpolation (Bülthoff & Edelman, 1992, 1993; Ullman & Basri, 1991). The identification of *invariant features* (Jolicoeur, 1990) also plays a role in achieving object constancy across plane rotation. Invariant features will be most useful when recognising small sets of distinctive objects, especially if the objects are presented repeatedly (Lawson & Jolicoeur, 1998, 1999a). To date, these alternatives have not been properly assessed in behavioural studies.

3. The achievement of object constancy across depth rotation

Plane rotations are a special case of image transformations about a unique axis along the line of sight of the viewer, and are rarely observed in everyday situations. I will describe rotations about any other axis as a depth rotation. In my own studies I have rotated objects about a vertical axis. This is probably the most common rotation observed in ecological viewing situations, although to date more research has investigated the effects of plane rotation. Indeed, there is often an implicit assumption that the visual system compensates for plane and depth rotations in the same way. This assumption may have arisen from Shepard and Metzler's (1971) finding that the angular separation between stimuli in a mirror-image discrimination task produces similar effects whether the separation is in the plane or in depth.

One reason for the bias towards investigating the effects of plane rather than depth rotations is the relative ease of producing plane rotated stimuli. A further reason is that differences between plane rotated views of the same object are well-defined and are restricted to changes in spatial relations relative to the viewer. All plane rotated views of a given object reveal the same features and parts in the same spatial relations to other object features and parts. The angle of plane rotation between two views of an object provides a psychologically plausible and readily assessed measure of the visual similarity of the two views, unlike the angle of depth rotation between two views (see Lawson & Humphreys, 1998a). For depth rotation, global shape and visibility of features and parts will usually change less following a rotation between 0° and 30° views than following the same 30° rotation but between 60° and 90° views (see Fig. 1, left column). Finally, for the objects typically presented in plane rotation studies, the canonical, upright view (from which the angle of plane rotation is measured) can be defined unequivocally as the usual view in-plane from which the object is seen. For most objects, there is not an equivalent, single, canonical depth-rotated view which can be specified independent of behavioural data (Newell & Findlay, 1997). Instead, preferred, canonical views in depth vary from object to object (Palmer, Rosch & Chase, 1981).

Recent technical improvements have made it easier to produce depth-rotated views of familiar objects. This has been a catalyst for a rapid increase in research into the achievement of object constancy across depth rotation. Researchers have generally manipulated view in depth relative to the most foreshortened view of the object (e.g. Lawson & Humphreys, 1996, 1998a, 1999; Newell & Findlay, 1997; Warrington & James, 1986, 1991). This view can be specified independent of behavioural data, and depth rotation can then be manipulated systematically, though a more psychologically plausible measure of visual changes due to depth-rotation would be preferable.

3.1. Are stored object representations invariant to depth rotation?

It is clear that depth rotation does influence object recognition. Severely foreshortened views and other unusual, depth-rotated views are recognised slower and less accurately than canonical views, so object constancy across depth rotation is not perfect (Humphrey & Jolicoeur, 1993; Lawson & Humphreys, 1996, 1998a, 1999; Newell & Findlay, 1997; Srinivas, 1993, 1995). These depth rotation effects on initial recognition could be due to certain views being intrinsically difficult to recognise (as blurred or small depictions of objects are hard to recognise). Alternatively, the effects might reflect access to stored view-specific object representations. If only canonical views of objects are stored, then unusual views (such as foreshortened views) would be expected to be recognised less efficiently than canonical views.

Priming studies allow us to dissociate the intrinsic difficulty of naming a particular view from effects of access to view-specific representations. In priming studies, subjects first see a prime view of an object, then they see a target view. Any intrinsic difficulty in recognising a given view should be similar for both the prime and the subsequent target views. In contrast, effects of view-specific representations on the recognition of target views should be influenced by the prime view. Views which were disadvantaged on initial recognition (such as foreshortened views) may subsequently be recognised faster than more canonical views if they benefit from view-specific priming. View-specific priming would occur when the prime and target views were similar or identical.

Glyn Humphreys and I (Lawson & Humphreys, 1996, 1998a) have reported such view-specific priming. Subjects were fastest to recognise targets depicting a depth-rotated view similar to the prime view of the object. This was true even when the target depicted a foreshortened view, see Fig. 3. In Lawson and Humphreys (1998a), we tested speeded naming in prime and target blocks. In Experiment 1, in the prime block, canonical, 150° views were named faster than foreshortened 90° views, as expected. In the target block, subjects were again faster to name 150° views than 90° views for trials primed by a 150° view of the object (top two conditions in Fig. 3). In contrast, and most critically, 90° targets were actually named faster than 150° targets for trials primed by a 90° view of the object (lower two conditions in Fig. 3). View-specific effects were not simply determined by the canonicality of a view. Naming was faster when the prime and target were identical relative to when they were dissimilar in view, even for foreshortened views. In Experiment 3 of Lawson and Humphreys (1998a), we reported view-specific priming effects for depth rotations as small as 10°.

We have also reported an analogous series of studies which tested speeded, sequential picture-picture matching (Lawson & Humphreys, 1996). In Experiment 4, 150° primes were matched faster to 150° targets than to 90° targets, as expected (top two conditions in Fig. 3). Most importantly, 90° primes were matched faster to 90° targets than to 150° targets (lower two conditions in Fig. 3). Thus, as in the naming studies, view-specific priming effects strongly influenced performance, suggesting that view-specific, stored representations are involved in the recognition of familiar objects.

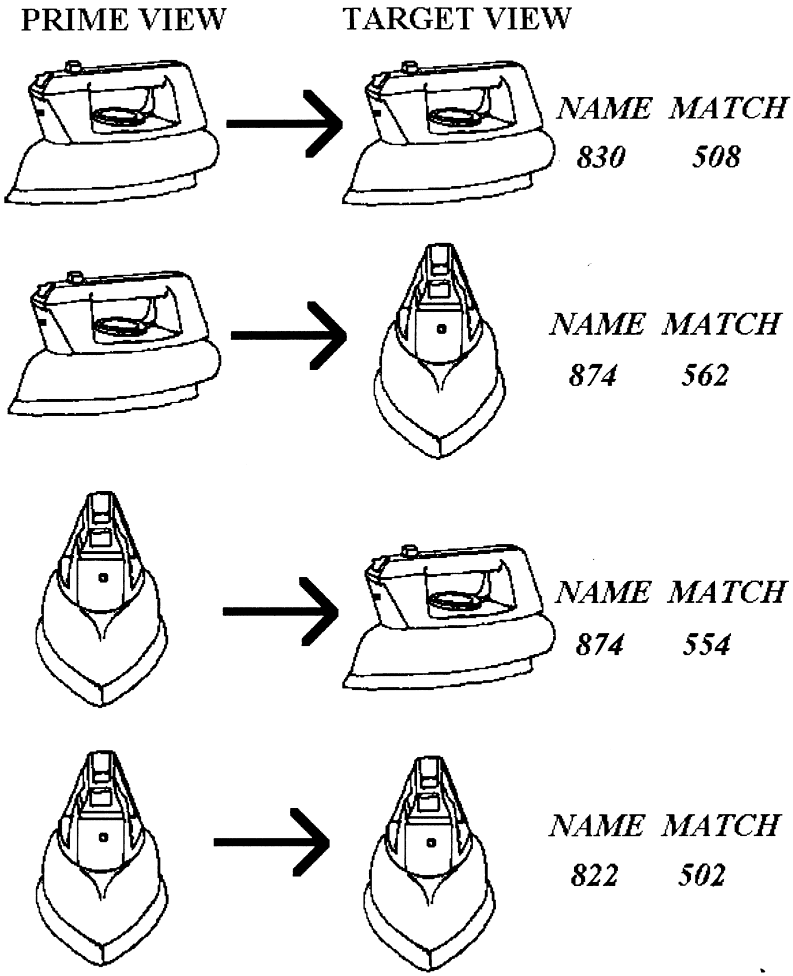

Fig. 3. The critical comparison conditions involving 90° (foreshortened) and 150° (canonical) views from two priming tasks, picture–picture matching and picture naming. The prime was presented immediately before the target in the matching task. The prime was presented in the previous block to the target in the naming task. On each row, the relevant prime and target views are depicted, followed by the RT to respond on prime-target match trials in the matching task (Experiment 4, Lawson & Humphreys, 1996) and the RT to name target views in Experiment 1 of Lawson and Humphreys (1998a).

In studies presenting novel 3D objects in naming and picture–picture matching tasks, Hayward and Tarr (1997) reported view-specific priming effects similar to our results for familiar objects. Finally, similar results (though with broader tuning of priming to depth rotation) were reported by Srinivas (1995) for long-term priming in an object decision task (object / non-object discrimination) for both familiar and novel objects. Again, these studies suggest that view-specific, stored representations mediate human visual object recognition.

The results of Biederman and Gerhardstein (1993) appear to contradict this conclusion. Their Experiment 1 was similar to Experiments 1 and 3 of Lawson and Humphreys (1998a), in being a priming study which required the speeded naming of line drawings of depth-rotated views of familiar objects. Biederman and Gerhardstein (1993) reported that the view of the prime had little effect on target naming, unless prime and target views revealed different parts or different spatial relations between parts. They claimed that in everyday situations, object recognition is largely invariant to depth rotation. They suggested that the strong effects of depth rotation reported in several studies (Edelman & Bülthoff, 1992; Rock & DiVita, 1987; Tarr & Pinker, 1989) were due to testing novel objects that, unlike the objects that humans must typically recognise, "failed to meet at least one of the conditions for viewpoint invariance, either because the stimuli did not decompose into a geon structural description or because the set members did not activate distinctive geon structural descriptions or produced non-stable part structures" (Biederman & Gerhardstein, 1993, p. 1166).

In contrast, the stimuli presented in Lawson and Humphreys (1996, 1998a) were line drawings of familiar objects similar to those presented by Biederman and Gerhardstein (1993). These stimuli could easily be decomposed into sets of distinctive parts and simple spatial relations that would allow the different objects to be distinguished. Nevertheless, we found view-specific priming effects whilst Biederman and Gerhardstein (1993) reported largely view-invariant priming. Note, though, that the trend of Biederman and Gerhardstein's data mirrored our data (Lawson & Humphreys, 1996; 1998a) in revealing greater priming when the prime and target views of an object were similar. Furthermore, Biederman and Gerhardstein tested only 24 objects (our studies tested 36 or 72 objects), so their studies may not have been as sensitive. In addition, some of the depth rotations which Biederman and Gerhardstein tested resulted in near mirror-image versions of stimuli being presented. This is a special case of depth rotation for which long-term priming of the visual system appears to be invariant (see above; Biederman & Cooper, 1991; Lawson & Humphreys, 1996, 1998a; Stankiewicz, Hummel & Cooper, 1998). Finally, for some objects in our studies (Lawson & Humphreys, 1996, 1998a) parts may have been occluded in certain views, resulting in different structural descriptions for those views. Interested readers are referred to Tarr and Bülthoff (1995) for further discussion of these and related issues; see also the response by Biederman and Gerhardstein (1995).

One note of caution should be made in the interpretation of these results from priming studies. In "long-term" priming studies (in which the target is seen several minutes after the prime and following many intervening items, e.g. Biederman & Gerhardstein, 1993; Lawson & Humphreys, 1998a), it is usually assumed that the representations mediating priming are the stored representations used in everyday object recognition. Nevertheless, it is possible that instead information in a short term store or memory cache causes the view-specific priming effects. This is most likely when the prime precedes the target by only a few seconds and there is either no intervening stimulus or only a mask between the prime and the target, as in the picture–picture matching studies of Lawson and Humphreys (1996).

In order to test this alternative interpretation of view-specific priming effects, studies should test very long-term priming, over days, weeks, and even months. These studies should disguise the repetition of prime objects by including many filler items and by changing the task and the experimental situation between the prime and target blocks. Such studies have been conducted with face stimuli (Bruce, Carson, Burton & Kelly, 1998), and have reported view-specific priming effects. If analogous studies with objects revealed view-specific priming, then we could be more confident in claiming that priming effects reflect access to stored object representations.

Additional evidence that view-specific, stored representations are used to recognise familiar objects comes from a series of studies in which my colleagues and I asked subjects to recognise familiar objects from a sequence of 12 depth-rotated views (Lawson, Humphrey & Watson, 1994). In Experiments 1 and 3, each successive view was presented for just 30 ms and was then pattern masked. The percentage correct recognition of objects from four types of sequence was:

(i) 41.9% – random sequences

(ii) 63.0% – coherent sequences (each successive view was rotated by a further 30°- the object appeared to rotate in depth smoothly, in the same direction)

(iii) 41.7% – coherent sequences (as (ii) except each successive view was rotated by a further 60°)

(iv) 68.6% – incoherent but visually similar sequences (the direction of rotation of the object reversed for each successive pair of views, and successive pairs were rotated from each other by a large angle, but within each pair of views, views were rotated by only 30°)

Performance was poor in (iii) although these sequences were coherent. Thus, in this task, the visual system could not use the coherent, structured information from the overall view sequence (although see Stone, 1998). In contrast, increasing the visual similarity of successive views in the sequence improved performance, both for (ii) coherent sequences and for (iv) incoherent sequences. When successive pairs of views are separated by just a 30° depth rotation, both views may access the same view-specific, stored representation, improving the likelihood of recognising that object. Such representations must, though, be narrowly tuned to view in depth- when successive views were rotated by 60° in (iii) sequences, objects were much more difficult to recognise.

If we do store multiple view-specific representations of a given object, a further question is how are these views related together? Recent work has started to investigate whether there are links between different, stored views of an object and whether apparent motion enables different views to prime each other (see Kourtzi & Shiffrar, 1999).

3.2. Cognitive neuropsychological evidence for multiple routes to object constancy

Over the past four decades, cognitive neuropsychologists have conducted many investigations into the achievement of object constancy following neurological damage (Lawson & Humphreys, 1998b). Warrington and colleagues (Warrington &

James, 1988; Warrington & Taylor, 1973; 1978) tested patients at matching a canonical view to a view which obscured important features or which foreshortened the object. Patients with right posterior damage were particularly poor at this test. Layman and Greene (1988) made similar observations and further noted that right-hemisphere lesioned patients found it harder to compensate for depth rotation than for plane rotation.

Warrington and Taylor (1973, 1978) suggested that visual object recognition could be divided into two main stages, first, perceptual categorisation (including the achievement of object constancy), which relied primarily on right hemisphere processing, then semantic categorisation, which mainly involved left hemisphere processing. If perceptual categorisation was impaired, it might only support the minimal processing required for conventional, canonical views of objects, which could then still access the semantic categorisation stage. Supporting this hypothesis, right-lesioned patients are often found to have difficulty in recognising unusual, atypical stimuli although canonical views of objects may still be recognised quite efficiently (e.g. Rudge & Warrington, 1991; Warrington & James, 1988).

Evidence for the existence of multiple routes to object constancy across depth rotation comes from Humphreys and Riddoch's (1984) investigation of five patients, all of whom revealed deficits in the unspeeded recognition of unusual views of familiar objects. H.J.A., a patient with bilateral occipital lesions, revealed a specific deficit in matching and naming views where the main distinguishing feature of the object was obscured. In contrast, four right-lesioned patients were particularly impaired in matching and naming foreshortened views of objects. Control subjects rated the feature-obscured views as having lower feature saliency but higher figural goodness and familiarity relative to the foreshortened views.

Humphreys and Riddoch (1984) proposed that there are at least two independent routes to object constancy across depth rotation. The first employs a distinctive local feature analysis; the second depends on a global shape analysis, is axis-based and may include the encoding of depth cues. HJA had a deficit in the global shape route and was reliant on identifying local distinguishing features. In contrast, the right-lesioned patients had a deficit in the local features route, and so were reliant on global shape analyses. Their difficulty in recognising foreshortened views was due to the reduced salience of the main axis of elongation of the object in these views, which often led them to impose an incorrect 2D rather than a 3D structure on such images. Their performance improved when the foreshortened stimuli were depicted against graph paper, presumably because this made global orientation information more accessible, by adding linear perspective cues. Similar results have been reported by Humphrey and Jolicoeur (1993) for neurologically intact subjects.

3.3. Do internal details aid in the achievement of object constancy?

Marr (1982) proposed that all views of an object access the same, view-invariant, stored structural description. These structural descriptions are described in relation to an object-centred co-ordinate system based around the object's main axis. Before an input image can be matched to a stored structural description, the image must be described with respect to the main axis of the object. Marr (1982) suggested that for certain views, the main axis was difficult to derive from the image, making these views hard to recognise. For example, in foreshortened views, the longest 2D axis of the image often does not coincide with the main axis of elongation of the object (see Fig. 1).

The evidence presented in Section 3.1 for view-specific, stored representations contradicts Marr's (1982) theory, which proposes that stored representations are view-invariant. Glyn Humphreys and myself (Lawson & Humphreys, 1999) have examined a further prediction related to Marr's theoretical claims. If all effects of depth rotation are due to difficulties in assigning the main axis of elongation to a given image, and if the ease of assigning the main axis is matched across two sets of stimuli, then depth rotation effects should be equal across those stimulus sets. We compared the efficiency of recognition of matched line drawings and silhouettes in a speeded word-picture verification task (see Fig. 1, left compared to right views of an iron). The outline global shape and aspect ratio of each matched line drawing and silhouette was identical, but the silhouettes lacked internal detail. We tested only clearly elongated objects (e.g. pencil, stapler, hammer) for which the principal axis of the object is most likely to be the main axis of elongation.

View effects were not equal across line drawings and silhouettes (see Fig. 1). First, across 0°, 30° and 60° views, silhouettes were verified only 45 ms slower than line drawings. Thus recognition was reasonably efficient even when internal details were absent, since the global shape of the object was informative. Second, when global shape was uninformative, as for most 90° views, then internal detail could still support quite efficient recognition, since 90° line drawings were verified only 26 ms slower than 0°, 30° and 60° line drawings. Third, foreshortened, 90° silhouettes, which lacked both global shape and internal detail information, were verified much slower than all other stimuli- 102 ms slower than 0°, 30° and 60° silhouettes. The same pattern of results was found for errors. This result indicates that either internal details aid recognition *directly,* or that internal details aid the extraction of the principal axis or secondary axes of description of the stimulus, such that the occluding contour is not the only source of information in locating axes.

In two further studies which have compared the recognition of shaded pictures of familiar objects and matched silhouettes, little or no disadvantage for silhouettes was found, except for foreshortened views (Hayward, 1998; Newell & Findlay, 1997). Together these three studies indicate that, as Marr (1982) had predicted, under normal conditions, internal detail and shading of the object are not necessary for fast and accurate recognition. There are, though, clearly circumstances in which objects are difficult or impossible to recognise from a silhouette. These include objects which differ only on surface information (boxes of cereal and boxes of washing powder), objects which differ only on "concave" information (a bowl with and without cereal in it) and, as noted above, objects for which most views are identifiable, but certain unusual views are not (such as foreshortened views of mugs and jugs).

Warrington and James (1986) tested the recognition of 3D silhouettes of familiar objects which fell into the latter category of being unidentifiable only when depicted from certain views. The silhouettes were initially presented from an unusual view (either a view of the base of the object or an upright, foreshortened view). The object was then gradually rotated (vertically or horizontally respectively) towards a canonical view, until the object was recognised. The view at which the object was first recognised was termed the minimal view. The minimal view varied across vertical and horizontal directions of rotations for a given object. It also varied across different objects for a given axis of rotation. However, the minimal view was consistent across subjects for a given object rotating about a particular axis. This provides further evidence for the conclusion drawn from Section 3 that the canonical view in depth varies across objects (Newell & Findlay, 1997; Palmer, Rosch & Chase, 1981).

Warrington and James (1986) suggested that accurate recognition required distinguishing features of the object to be visible. Their results suggest that subjects use the same distinguishing features given a particular object rotating about a given axis. Warrington and James (1986) compared right-lesioned patients to neurally intact control subjects on the silhouette recognition task. The minimal rotations required by the patients were greater than those needed by the controls, but patients and controls revealed the same pattern of performance (objects rotating about a given axis which were relatively easy for the patients to recognise were also easy for the controls to recognise). This suggests that patients and controls were using the same information to recognise objects (possibly the same distinguishing features).

A subsequent study by Warrington and James (1991) presented single, 2D silhouettes at the view at which 75% of control subjects could recognise the object. In an object decision task in which subjects had to discriminate between these silhouettes of familiar objects and silhouetted non-objects, right-lesioned patients performed worse than left-lesioned patients, whose performance was no worse than that of controls. For all three groups of subjects, performance on this object decision task correlated to their performance on the unusual views test described above, suggesting that the two tests measure the same process of achieving object constancy across depth rotation. Results from these two tasks indicate that right but not left-hemisphere lesioned patients have specific difficulties in achieving object constancy for depth-rotated silhouettes, supporting the conclusions of Section 3.2.

3.4. Summary: the achievement of object constancy across depth rotation

The visual system is not perfectly efficient at achieving object constancy across depth rotation. Some unusual views (such as foreshortened views) are recognised slower and less accurately than more canonical views. View-specific effects are not dependent solely on the typicality, familiarity or quality of views, since foreshortened views can be recognised more efficiently than canonical views if they are primed by a similar depth-rotated view (see Fig. 3). This suggests that view-specific object representations are stored by the visual system. The intact visual system uses information from different sources and at different spatial scales (global shape and internal detail). For most views of an object, either local or global information is, by itself, sufficient to recognise the object. Performance of the visual system only starts to break down if stimulus information is impoverished (as for most foreshortened silhouettes, where both local and global information is either absent or misleading) or if the system is itself damaged, preventing it from using certain types of information (as for HJA and certain right-lesioned patients).

4. What are the effects of combining depth rotation and plane rotation?

Sections 2 and 3 reviewed studies investigating the individual effects of plane and depth rotation on the recognition of familiar objects. Following from this, one obvious question is whether common visual transformations compensate for both plane disorientation and foreshortening on object recognition. Plane disorientation and foreshortening both increase naming RTs and errors. In addition, the effects of both manipulations reduce with practise at recognising a fixed set of stimuli (for plane rotation, see Fig. 2; see also Jolicoeur, 1985; Lawson & Jolicoeur, 1999a; for depth rotation, see Lawson & Humphreys, 1998a). Nevertheless, these similarities may be superficial. More direct comparisons of the effects of plane and depth rotation are required before strong conclusions can be drawn. My colleagues and I have started to conduct such research by comparing the achievement of object constancy across combined transformations of plane and depth rotation for familiar objects, both on initial recognition and following training (Lawson, Humphreys & Jolicoeur, 1999; Lawson & Jolicoeur, 1999b).

5. What are the routes to object constancy?

There are similarities between the effects of the different transformations (plane rotation, depth rotation, removing internal detail), tasks (speeded and unspeeded; mirror-image discrimination, naming, word-picture verification, sequence recognition and picture-matching) and subject groups (neurally intact and brain-damaged) used in the studies reviewed here. However, I have noted important differences in the achievement of object constancy, first, given different image information tested with the same task (e.g., global shape versus local, distinguishing features, Humphreys & Riddoch, 1984; drawings with and without internal detail, Hayward, 1998; Lawson & Humphreys, 1999; Newell & Findlay, 1997), and second, given different tasks presenting the same stimuli (e.g., comparing mirror-image discrimination to recognition for plane-rotated views of familiar objects, Farah & Hammond, 1988; Jolicoeur Corballis & Lawson, 1998; Lawson & Jolicoeur, 1998, 1999a; Turnbull & McCarthy, 1996).

The extraction of view-invariant features is not likely to be an effective means of achieving object constancy for the initial, unconstrained, entry-level recognition of objects in everyday situations. Nevertheless, if a set of readily distinguishable stimuli are presented many times for recognition, subjects may learn to extract such features efficiently, if such features are available. In addition, subjects may be forced to rely on using view-invariant features under extreme conditions (for instance, when pictures of objects are presented briefly, at low contrast, and are then masked, Lawson & Jolicoeur, 1998), although performance is then likely to be inaccurate. Finally, view-invariant features may suffice to classify an object at the superordinate level (an animal) and to indicate its likely entry-level identity (probably a dog, maybe a cat).

Under most circumstances, it appears that view-specific transformation processes or view-specific, stored representations are required to achieve object constancy. The route to achieving object constancy will depend on the stimuli presented, the context in which the stimuli must be recognised and the task required. Unfortunately, most current accounts are under-specified, making it difficult to devise rigorous empirical tests of their predictions under these diverse experimental conditions. In order for our understanding to increase, much more detailed theoretical hypotheses are now required about the nature of the processes and representations involved in the achievement of object constancy.

References

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, *94*, 115–147.

Biederman, I., & Cooper, E.E. (1991). Evidence for complete translational and reflectional invariance in visual object priming. *Perception*, *20*, 585–593.

Biederman, I., & Gerhardstein, P.C. (1993). Recognising depth-rotated objects: evidence for 3-D viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.

Biederman, I., & Gerhardstein, P.C. (1995). Viewpoint-dependent mechanisms in visual object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1506–1514.

Bruce, V., Carson, D., Burton, A.M., & Kelly, S. (1998). Prime time advertisements: Repetition priming from faces seen on subject recruitment posters. *Memory and Cognition*, *26*, 502–515.

Bülthoff, H.H., & Edelman, S. (1992). Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science, U.S.A*, *89*, 60–64.

Bülthoff, H. H. & Edelman, S. (1993). Evaluating object recognition theories by computer graphics psychophysics. In T. A. Poggio & D. A. Glaser (Eds.), *Exploring brain functions: Models in neuroscience* (139–164). New York: Wiley.

Corballis, M.C. (1988). Recognition of disoriented shapes. *Psychological Review*, *95*, 115–123.

De Caro, S. A. & Reeves, A. (1995, November). *Object recognition without mental rotation.* Paper presented at the Annual Meeting of the Psychonomic Society, Los Angeles.

Edelman, S., & Bülthoff, H.H. (1992). Orientation dependence in the recognition of familiar and novel views of 3D objects. *Vision Research*, *32*, 2385–2400.

Edelman, S., & Weinshall, D. (1991). A self-organising multiple-view representation of 3-D objects. *Biological Cybernetics*, *64*, 209–219.

Farah, M.J., & Hammond, K.M. (1988). Mental rotation and orientation-invariant object recognition: Dissociable processes. *Cognition*, *29*, 29–46.

Hamm, J.P., & McMullen, P.A. (1998). Effects of orientation on the identification of rotated objects depend on the level of identity. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 413–426.

Hayward, W.G. (1998). Effects of outline shape in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 427–440.

Hayward, W.G., & Tarr, M.J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 1511–1521.

Hummel, J.E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480–517.

Hummel, J. E. & Stankiewicz, B. J. (1997). An architecture for rapid, hierarchical structural description. In T. Inui, & J. L. McClelland, *Attention and Performance XVI*, (pp. 93–121). Cambridge MA: MIT Press.
Humphrey, G.K., & Jolicoeur, P. (1993). Visual object identification: Some effects of image foreshortening, monocular depth cues, and visual field on object identification. *Quarterly Journal of Experimental Psychology A*, *46*, 137–159.
Humphreys, G.W., & Riddoch, M.J. (1984). Routes to object constancy: Implications from neurological impairments of object constancy. *Quarterly Journal of Experimental Psychology A*, *37*, 493–495.
Humphreys, G.W., Riddoch, M.J., & Quinlan, P.T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*, 67–103.
Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory and Cognition*, *13*, 289–303.
Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology*, *42*, 461–478.
Jolicoeur, P. (1990). Identification of disoriented objects: a dual systems theory. *Mind and Language*, *5*, 387–410.
Jolicoeur, P., Corballis, M.C., & Lawson, R. (1998). The influence of perceived rotary motion on the recognition of rotated objects. *Psychonomic Bulletin and Review*, *5*, 140–146.
Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: effects of context of prior presentation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 200–210.
Kourtzi, Z. & Shiffrar, M. (1999). The visual representation of rotating, three dimensional objects. *Acta Psychologica,* 265–292.
Lawson, R., & Humphreys, G.W. (1996). View-specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 395–416.
Lawson, R., & Humphreys, G.W. (1998a). View-specific effects of depth rotation and foreshortening on the initial recognition and priming of familiar objects. *Perception and Psychophysics*, *60*, 1052–1066.
Lawson, R. & Humphreys, G. W. (1998b). The neuropsychology of visual object constancy. In V. Walsh & J. Kulikowski, *Perceptual Constancies: Why things look as they do* (pp. 210–228). Cambridge: Cambridge University Press.
Lawson, R. & Humphreys, G. W. (1999). The effects of view in depth on the recognition of line drawings and silhouettes of familiar objects: Normality and pathology. *Visual Cognition, 6*(2), 165–195.
Lawson, R., Humphreys, G. W., & Jolicoeur, P. (1999). The combined effects of plane rotation and foreshortening on picture naming- one manipulation or two? *In Review*.
Lawson, R., Humphreys, G.W., & Watson, D.G. (1994). Object recognition under sequential viewing conditions: Evidence for viewpoint-specific recognition procedures. *Perception*, *23*, 595–614.
Lawson, R., & Jolicoeur, P. (1998). The effects of plane rotation effects on the recognition of brief masked pictures of familiar objects. *Memory and Cognition*, *26*, 791–803.
Lawson, R. & Jolicoeur, P. (1999a). The effect of prior experience on recognition thresholds for plane rotated pictures of familiar objects. *Memory and Cognition,* in press.
Lawson, R. & Jolicoeur, P. (1999b). Within-rotation and between-rotation transfer of training effects for plane and depth transformations of familiar objects. *In preparation*.
Layman, S., & Greene, E. (1988). The effect of stroke on object recognition. *Brain and Cognition*, *7*, 87–114.
Lowe, D. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence*, *31*, 355–395.
Marr, D. (1982). *Vision.* Freeman: San Francisco.
McMullen, P.A., & Jolicoeur, P. (1990). The spatial frame of reference in object naming and discrimination of left-right reflections. *Memory and Cognition*, *18*, 99–115.
McMullen, P.A., & Jolicoeur, P. (1992). The reference frame and effects of orientation on finding the top of rotated objects. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 807–820.
Murray, J.E. (1995). Imagining and naming rotated natural objects. *Psychonomic Bulletin & Review*, *2*, 239–243.

Murray, J.E. (1997). Flipping and spinning: spatial transformation procedures in the identification of rotated natural objects. *Memory and Cognition*, *25*, 96–105.

Murray, J.E. (1998). Is entry-level recognition viewpoint invariant or viewpoint dependent?. *Psychonomic Bulletin & Review*, *5*, 300–304.

Newell, F.N., & Findlay, J.M. (1997). The effect of depth rotation on object identification. *Perception*, *26*, 1231–1257.

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

Ratcliff, G. (1979). Spatial thought, mental rotation, and the right cerebral hemisphere. *Neuropsychologia*, *17*, 49–54.

Rock, I., Schreiber, C., & Ro, T. (1994). The dependence of two-dimensional shape perception on orientation. *Perception*, *23*, 1409–1426.

Rock, I., & DiVita, J. (1987). A case of viewer-centred perception. *Cognitive Psychology*, *21*, 280–293.

Rudge, P., & Warrington, E.K. (1991). Selective impairment of memory and visual perception in splenial tumours. *Brain*, *114*, 349–360.

Shepard, R. N. & Cooper, L. A. (1982). *Mental images and their transformations.* Cambridge, MA: MIT Press.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701–703.

Solms, M., Kaplan-Solms, K., Saling, M., & Miller, P. (1988). Inverted vision after frontal lobe disease. *Cortex*, *24*, 499–509.

Srinivas, K. (1993). Perceptual specificity in nonverbal priming. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 582–602.

Srinivas, K. (1995). Representation of rotated objects in explicit and implicit memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 1019–1036.

Stankiewicz, B.J., Hummel, J.E., & Cooper, E.E. (1998). The role of attention in priming for left-right reflections of object images: Evidence for a dual representation of object shape. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 732–744.

Stone, J.V. (1998). Object recognition using spatiotemporal signatures. *Vision Research*, *38*, 947–951.

Takano, Y. (1989). Perception of rotated forms: A theory of information types. *Cognitive Psychology*, *21*, 1–59.

Tarr, M.J. (1995). Rotating objects to recognition them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychological Bulletin & Review*, *2*, 55–82.

Tarr, M.J., & Bülthoff, H.H. (1995). Is human object recognition better described by geon-structural descriptions or by multiple-views. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1494–1505.

Tarr, M.J., & Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychology*, *21*, 233–283.

Tarr, M.J., & Pinker, S. (1990). When does human object recognition use a viewer-centred reference frame?. *Psychological Science*, *1*, 253–256.

Turnbull, O.H., Laws, K.R., & McCarthy, R.A. (1995). Object recognition without knowledge of object orientation. *Cortex*, *31*, 387–395.

Turnbull, O.H., & McCarthy, R.A. (1996). Failure to discriminate between mirror-image objects: A case of viewpoint-independent object recognition? *Neurocase*, 2, 63–72.

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

Ullman, S., & Basri, R. (1991). Recognition by linear combination of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*, 992–1006.

Warrington, E.K. (1982). Neuropsychological studies of object recognition. *Philosophical Transactions of the Royal Society of London, B*, *298*, 15–33.

Warrington, E.K., & James, M. (1986). Visual object recognition in patients with right-hemisphere lesions: Axes or features?. *Perception*, *15*, 355–366.

Warrington, E.K., & James, M. (1988). Visual apperceptive agnosia: a clinico-anatomical study of three cases. *Cortex*, *24*, 13–32.

Warrington, E.K., & James, M. (1991). A new test of object decision: 2D silhouettes featuring a minimal view. *Cortex*, *27*, 377–383.

Warrington, E.K., & Taylor, A.M. (1973). The contribution of the right parietal lobe to object recognition. *Cortex*, *9*, 152–164.

Warrington, E.K., & Taylor, A.M. (1978). Two categorical stages of object recognition. *Perception*, *7*, 695–705.