

Identifying objects in conventional and contorted poses: contributions of hemisphere-specific mechanisms

Bruno Laeng^{a,b,*}, Jinesh Shah^a, Stephen Kosslyn^a

^a*Harvard University, Cambridge, MA, USA*

^b*Department of Psychology, University of Tromsø, Tromsø N-9037, Norway*

Received 7 July 1998; accepted 26 November 1998

Abstract

Three experiments were designed to test the hypothesis that different mechanisms are used to encode objects seen in unfamiliar contortions than are used to encode objects seen in conventional poses. When a familiar non-rigid form (e.g. an animal) is seen in a contorted pose, we hypothesize that object identification may be achieved by (1) encoding the object's parts separately, (2) encoding the spatial relations among the parts, and (3) matching these encodings to a stored structural description. However, once this form has become familiar, its global shape can be directly matched to information stored in memory. Based on the idea that 'categorical' spatial relations are encoded better by the left cerebral hemisphere and are used in structural descriptions, we predicted a left-hemisphere advantage when one first encodes contorted poses; in contrast, based on the idea that overall shapes are encoded better by the right hemisphere, we predicted a right-hemisphere advantage for encoding the same shapes after they are familiar. Three experiments confirmed these predictions, which supports the hypotheses that different visual mechanisms operate in the recognition of familiar and unfamiliar views of known non-rigid objects. Moreover, correlational analyses between visual-field differences in several perceptual tasks (matching whole pictures to names, body parts to the whole body, and judging categorical spatial relations) revealed that the degree and lateralization of categorical spatial encoding predicts the left hemisphere's initial advantage in the identification of contorted shapes. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Object identification; Spatial cognition; Categorical spatial relations; Cerebral lateralization

* Corresponding author. Fax: +47-77-645291; E-mail: bruno@psyk.uit.no

1. Introduction

Most plants and animals, and many objects, have joints that can bend or surfaces that can be stretched, creased and folded, either by internal or external forces. Hence, such objects project many different images onto our retinas, which somehow must be recognized as arising from the same object. Consider, for example, the varieties of shapes that a cat can take during a very common sequence of feline actions: jumping on a chair, sitting down and grooming its muzzle, curling up asleep. Although rigid objects can project different images when seen from different viewpoints (e.g. consider walking around buildings or mountains), their parts and surfaces always remain at fixed locations relative to one another. In contrast, parts and surfaces of flexible objects change their relative locations when these objects contort. Computationally different problems must be solved by a system that identifies flexible objects than by one that identifies only rigid objects, and different mechanisms may be used in the two cases.

One of the differences between flexible and rigid objects is that the geometric structure of a rigid object can be completely described by a single set of spatial coordinates (e.g. in a Cartesian framework) among its parts. In contrast, the only spatial relations among parts of a flexible object that must remain invariant during its non-rigid transformations are abstract, describing qualitative properties of the connections (such as that one part is above another or on the side of another; cf. Richards et al., 1988). Kosslyn (1987) calls such abstract spatial representations ‘categorical’ because they define an equivalence class, independent of any precise position.

The fact that the structure of flexible objects remains invariant at an abstract level has been taken to imply that we recognize such objects, at least in part, by using a ‘structural description’ that specifies parts and their connections (e.g. cf. Marr, 1982). Consider a case in which a non-rigid object, such as the human body, is seen for the first time in a new contortion. If our visual memories are stored in a viewer-centered format, as much recent data suggests (for reviews, see Kosslyn, 1994; Ullman, 1996), such input would fail to match the stored representations of the object. Only a few parts of the object (e.g. a hand or foot or part of the head) might match stored representations. Although the contortion alters the global shape, the individual parts are left intact (although some parts may have become occluded or may be seen from a novel perspective). Moreover, the spatial relations among parts may also be evident. Hence, the parts and their spatial relations can provide clues as to the identity of the whole. Such information is particularly useful in conjunction with an abstract description of how parts are arranged, which in turn can guide active top-down search for distinctive characteristics (see Kosslyn et al., 1990).

On the basis of these observations, then, we suggest that people will use information based on the encoding and recognition of individual parts of objects when faced with a flexible object seen in an unfamiliar pose; in contrast, if the object and its pose are familiar, then the entire shape can be matched as a single representation. This notion is consistent with theories that posit view-centered representations and matching processes (e.g. Lowe, 1985; Kosslyn, 1994; Ullman, 1996), and is incon-

sistent with theories that posit that parts are *always* utilized during the course of object identification (e.g. Biederman, 1985).

Information about the global shape and the parts of an object may be extracted automatically and in parallel (e.g. a simple feature such as convexity, on which the description of an object's parts may depend, cf. Hoffman and Singh, 1997, can be preattentively salient; Driver and Baylis, 1995; Wolfe and Bennett, 1997). However, both sorts of information may not be used to the same extent in all circumstances. Whether information about global shape or parts will be used in object identification may depend on the viewing conditions and stored visual knowledge. Typically, one will need to encode not simply the parts themselves, but also the spatial relations among them, and this entire description would be compared to stored structural descriptions (i.e. a descriptions of sets of parts and their spatial relations) in order to identify the object. This process is computationally expensive (cf. Ullman, 1996; Saiki and Hummel, 1998), and we surmise that, in visually familiar and unambiguous circumstances, global shape extraction and immediate matching to stored perceptual templates would be a more efficient strategy than one based on the recognition and description of parts. If so, then a global shape matching process should be faster than a parts-based process (cf. Cave and Kosslyn, 1989).

We also hypothesize that both parts and categorical spatial relations among them are encoded better in the left cerebral hemisphere (LH). Robertson and her colleagues (e.g. Delis et al., 1986; Robertson and Delis, 1986, 1991) have presented much evidence that left hemisphere damage selectively disrupts a patient's ability to encode constituent parts of objects, and normal subjects can encode local elements of a hierarchical display (Navon figures; see Navon, 1977) better when the stimuli are presented in the right visual field (RVF, and hence encoded initially by the LH). In addition, there is evidence that categorical spatial relations representations are encoded better in the LH (e.g. Kosslyn, 1987; Hellige and Michimata, 1989; Kosslyn et al., 1989, 1992; Laeng, 1994; Laeng and Peters, 1995; Laeng et al., 1998).

The LH's specialization for categorical spatial relations is in contrast to the well-known specialization of the right hemisphere (RH) for spatial tasks, which seems to be restricted to the use of 'coordinate' spatial relations representations (see Kosslyn, 1987, 1994; Laeng, 1994; Kosslyn and Koenig, 1995). Such representations specify the precise metric distances among objects or parts, and are particularly useful for motor control. There is also evidence that global forms are encoded better when presented in the left visual field (LVF, and hence encoded initially by the RH; e.g. Delis et al., 1986; Robertson and Delis, 1986; Robertson and Lamb, 1991) and that visual representations of specific shapes are matched more effectively in the RH (e.g. Marsolek et al., 1992, 1994; Marsolek, 1995). Thus, we expect that if people are very familiar with a shape (so that it is very likely to be stored as such), they can identify it more effectively in the RH.

2. Experiment 1

Animals are natural objects that often require us to identify a shape from a variety

of unpredictable transformations of its form. Therefore, we used drawings of animals as the stimuli to be identified. We hypothesized that the LH's superior ability to process structural descriptions of flexible shapes would lead it to identify novel views of non-rigid objects better than the RH; in contrast, the RH's superior ability to store the visual appearances of specific shapes would lead it to identify familiar views of objects better than the LH.

In short, we asked subjects to view pictures and hear words that either were or were not appropriate names for the pictured objects. Their job was to decide whether the word correctly named the picture as quickly as possible. The pictures either depicted the object in a contorted pose or a conventional pose, and were presented equally often in the LVF and RVF (and hence were encoded initially by the RH or LH, respectively). We are led to make two predictions: First, drawings of animals will be recognized better, in terms of both response time and accuracy, in the RVF than in the LVF, if the pose of the animal is unfamiliar and clearly 'contorted' (i.e. the animal bends its body or body parts, thus causing unusual juxtapositions and occlusions of relevant body parts). Second, we do not expect such a laterality effect for familiar poses of objects. Rather, as unfamiliar views of objects become familiar, we expect the initial RVF advantage to disappear and be replaced by a LVF advantage (because the specific pictures become stored as particular exemplars of global shapes in the RH). Because one may learn to recognize a specific pose after only a single presentation, we were particularly interested in comparing visual field differences in performance between the very first presentation of a drawing and all the subsequent presentations of the same drawing.

2.1. Method

2.1.1. Subjects

Twenty female and 20 male right-handed students at Harvard University volunteered to serve as paid subjects. Handedness was assessed with the Edinburgh questionnaire (Oldfield, 1971). Each subject was assigned to one of four experimental groups and each of these groups had an identical number of males and females.

2.1.2. Stimuli and apparatus

Sixty-four black-and-white line drawings of animals were selected from zoology books (e.g. Harter, 1979; Leclerc, 1993). All drawings were realistic depictions (either engravings or fine ink pen drawings) in which the correct proportions, natural texture, and shading of the animals' bodies were clearly represented. Two drawings were selected for each species, according to the following criteria: (1) one drawing depicted a specific view of the animal that was judged by the authors to be 'conventional' (i.e. the body appeared in a natural standing position and, typically, as seen from either a side or 3/4 'canonical perspective', cf. Palmer et al., 1981); (2) the other drawing depicted the animal in a pose judged by the authors to be contorted or relatively unconventional and unfamiliar; the animal's whole body and/or body parts appeared bent or flexed (although they were in a position that would be natural

to the animal); however, even in these drawings the perspective could be canonical to the main axis of skew symmetry. Fig. 1 provides examples of the two types of stimuli.

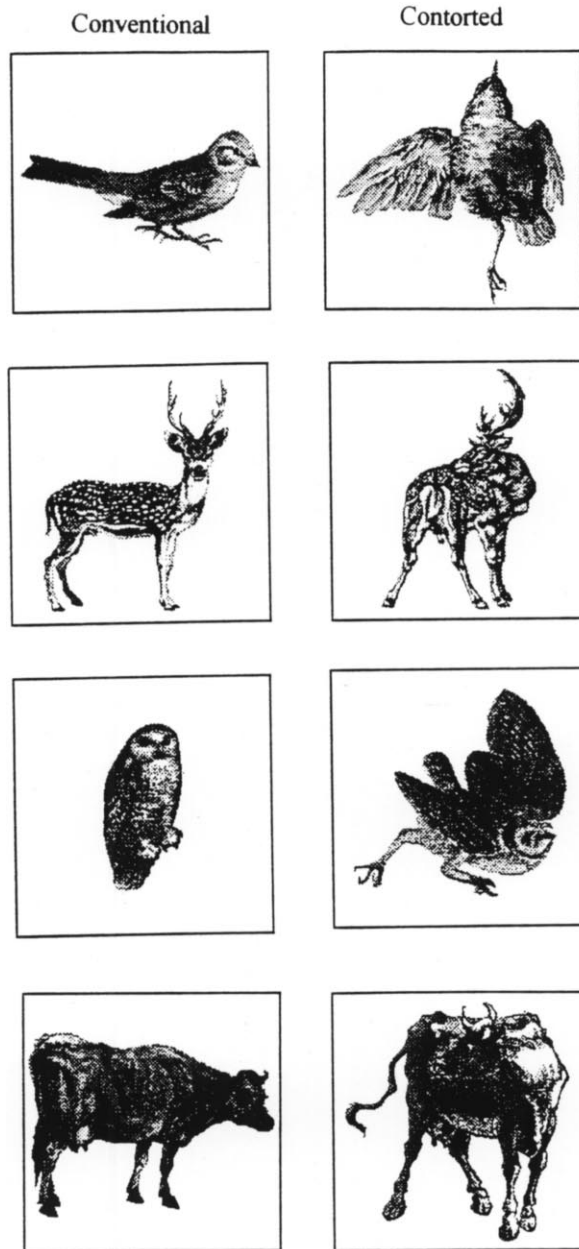


Fig. 1. Examples of the stimuli used in all experiments. The left column shows examples of Conventional poses of animals, the right column shows the same animal in a Contorted pose.

Thirty-two pairs of drawings of animals were obtained and divided into 4 stimulus sets (A–D). Set A included an alligator, anteater, baboon, bat, cat, cow, deer, and dog; set B included a duck, elephant, fish, giraffe, horse, jaguar, lion and pheasant; set C included a lizard, ostrich, otter, owl, panther, parrot, peacock, and rabbit; and, set D included a rat, rooster, seal, shark, sheep, sparrow, toucan, and turtle. These animals' names were also recorded digitally, while spoken slowly and clearly by a male, for auditory presentation. On foil trials, names of dissimilar animals used in the task were paired with an inappropriate picture (e.g. a picture of a 'lion' and the name 'horse'); each name appeared equally often as a distractor and a target.

Drawings were digitized by a Microtek Scanmaker 600ZS scanner and edited with commercial software (e.g. Adobe Photoshop). The background was edited out along the contour of the animal's shape, leaving only the animal's figure. The original orientation of the figure in the drawing was not changed. However, the size of the drawings was normalized by fitting each drawing into the frame of a 7×7 cm empty box; the boundary of the drawing was no more than 5 mm from any side of the frame. The box with an animal inside could appear in one of two different positions on the central horizontal meridian of the Macintosh II screen, either 2.7 cm to the left or 2.7 cm to the right of central fixation. The head of the animal was always toward the screen center (i.e. rightward for LVF trials and leftward for RVF ones) in order to facilitate recognition (details of the head are typically smaller than those of other parts of the body and would therefore benefit from the higher resolution of retinal locations near to the fovea). Two versions of each of the lateralized stimuli were prepared, one in each visual field (which allowed appropriate counterbalancing, as described below). All sequences of trials and storage of key presses were controlled by MacLab software (Costin, 1988).

2.1.3. Design

Subjects were divided into four groups and stimuli into four sets so that individuals in each group saw all four sets randomly intermixed, but one view (contorted or conventional) of a specific animal in a same set would always be presented *in its very first presentation* in one visual field. For example, group 1 saw, initially, (1) contorted poses of animals in set A in the LVF, (2) contorted poses of animals in set B in the RVF, (3) conventional poses of animals in set C in the LVF and (4) conventional poses of animals in set D in the RVF. Each group saw for the first time one version of an animal in one set in a different visual field than any other group of subjects, counterbalancing this factor across the groups.

For the following seven presentations of a same drawing, three were in the same visual field as the first presentation and four in the opposite visual field. Thus, there were a total of 256 trials, with an equal number of trials in which each animal appeared in each visual field and pose. The order of both visual field and pose in the trials following the first presentation was pseudo-random (random but with no more than three consecutive trials in the same visual field or with the same response, and no consecutive presentations of the same animal's drawing or name or those of taxonomically closely related animals).

2.1.4. Procedure

Subjects sat upright and placed their heads on a chin-rest located 45 cm away from the computer screen. There were no practice trials; subjects were told that they would see drawings and hear names of animals, and their job was to determine whether the name was appropriate for that animal. They were told that each trial had the following sequence of events: (1) A blank screen; the subjects would initiate a trial by pressing the space bar of the computer's keyboard; (2) a small fixation cross then appeared at the center of screen, which remained visible for 450 ms; the subjects were asked to gaze directly at the cross and maintain fixation until the end of the trial; (3) a blank screen appeared for 50 ms; (4) a lateralized visual stimulus (i.e. an animal in a box) appeared for 100 ms; (5) a blank screen appeared, and after 50 ms the recorded name of an animal was played through the computer's speaker; (6) the subjects were asked to indicate whether the name was appropriate, as quickly and accurately possible, by using the right hand's fingers to press one of two adjacent keys marked 'yes' and 'no'. Following this, the subject would initiate a new trial.

2.2. Results

We first obtained descriptive statistics for each subject. Means of percent errors and response times (RTs) were calculated for each combination of the following factors: Pose (contorted versus conventional), presentation ('initial', i.e. presentation 1, versus 'following', i.e. presentations 2–8), and visual field (left versus right). The data were pooled over type of match (yes versus no), since a preliminary analysis showed that this factor did not interact with any of the other factors. Times from trials on which errors occurred were excluded from computation of the RTs, and trials with RTs longer than three standard deviations from each individual's mean RT for that condition were excluded from all the statistics (on average 2% of data were discarded for each subject).

2.2.1. Response times

An analysis of variance with group as the between-subjects factor and pose (contorted versus conventional), presentation (initial versus following presentations) and visual field (left versus right) as within-subjects factors was performed on RTs as the dependent variable. An 'initial' presentation was considered the first time a specific pose of a specific animal was seen. 'Following' presentations were any other presentation of a specific drawing from the second time to the last (i.e. the eighth presentation).

The analysis of variance revealed an interaction between the factors of pose, presentation and visual field, $F(1/36) = 3.93$, $P < 0.05$. Most important, as predicted, subjects evaluated initial presentations of animals seen in contorted poses faster when the pictures were presented in the RVF/LH (mean RT = 1120 ms; SD = 215) than when they were presented in the LVF/RH (mean RT = 1175 ms; SD = 266). Because we had specific predictions about the effects of pose on the initial and following trials, we analyzed each visual field difference in separate

analyses of variance, with group as the between-subjects factor and visual field as the within-subjects factor. The first analysis confirmed the RVF/LH advantage for initial presentations of animals in contorted poses, $F(1/36) = 7.6$; $P < 0.01$. This visual field difference appeared reversed for following presentations of contorted poses (for the LVF/RH, mean RT = 872 ms; SD = 140; for the RVF/LH, mean RT = 886 ms; SD = 146). A separate analysis of variance revealed this to be just a trend towards a LVF/RH advantage, $F(1/36) = 3.3$; $P < 0.08$. Moreover, subjects evaluated following presentations of animals in conventional poses faster in the LVF/RH (mean RT = 799 ms; SD = 116) than in the RVF/LH (mean RT = 817 ms; SD = 117); a separate analysis of variance confirmed such LVF/RH-advantage, $F(1/36) = 8.7$; $P < 0.006$. Finally, there was no statistically reliable visual field difference for the initial presentations of conventional shapes, $F(1/36) = 0.1$; $P < 0.72$. All of these effects are illustrated in Fig. 2.

The main analysis also revealed that animals in conventional poses were identified faster (mean RT = 915 ms; SD = 188) than those in contorted poses (mean RT = 1013 ms; SD = 239), $F(1/36) = 78.5$; $P < 0.0001$. There was also a practice effect, with initial presentations being identified more slowly (mean RT = 1085 ms; SD = 228) than following presentations (mean RT = 843 ms; SD = 134), $F(1/36) = 276$; $P < 0.0001$. We also found that subjects showed a larger practice effect

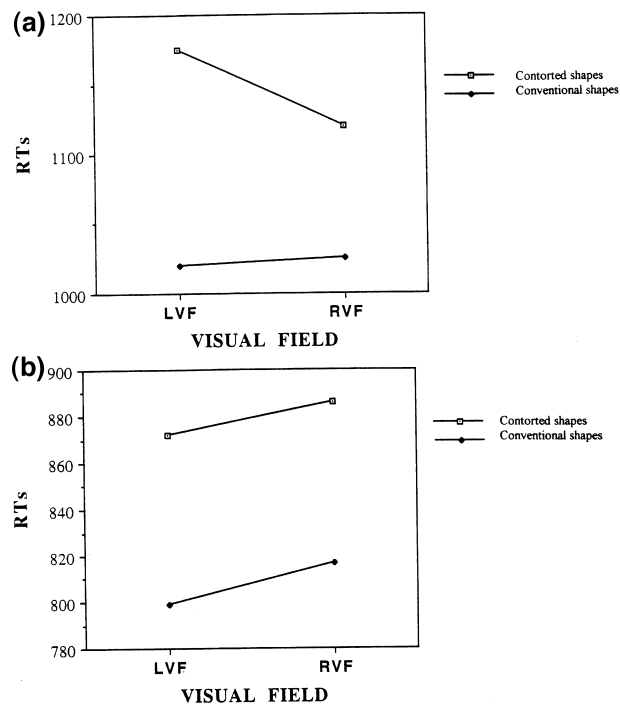


Fig. 2. Experiment 1. Means of response times (RTs) to Contorted and Conventional shapes of animals in the left visual field (LVF) and the right visual field (RVF). (a) Shows the performance for the initial presentation only, whereas (b) shows it for all following presentations (2–8).

(i.e. reduction of RT from initial to following presentations) for contorted poses than for conventional poses, as evident in an interaction of pose and presentation, $F(1/36) = 15.0$; $P < 0.004$.

In addition, subjects generally evaluated the pictures of animals in contorted poses more quickly when they were presented in the RVF/LH (mean RT = 1003 ms; SD = 217) than in the LVF/RH (mean RT = 1024 ms; SD = 260), whereas they evaluated animals in conventional poses generally faster when pictures were presented in the LVF/RH (mean RT = 909 ms; SD = 186) than in the RVF/LH (mean RT = 921 ms; SD = 191), $F(1/36) = 5.7$; $P < 0.02$ for the interaction of visual field and pose. Moreover, subjects were generally faster for initial presentations in the RVF/LH (mean RT = 1073 ms; SD = 209) than the LVF/RH (mean RT = 1097 ms; SD = 238), whereas they were generally faster for following presentations in the LVF/RH (mean RT = 836 ms; SD = 133) than in the RVF/LH (mean RT = 851 ms; SD = 136), $F(1/36) = 9.9$; $P < 0.003$.

The between-subjects factor of group interacted with the factors of pose, presentation and visual field, $F(3/36) = 12.8$; $P < 0.0001$. It appeared that although three of the four groups showed the above-described pattern of findings, one group had no reliable difference between visual fields for the initial presentations of animals in contorted poses. This group may have viewed in the LVF drawings that showed animals in contorted poses that were not as novel as the others, or that had a smaller amount of deformation of the typical animal's shape than those in the other sets. To confirm this hypothesis, we performed a separate ANOVA on drawing sets as the within-subjects factor (thus collapsing data over all the other factors), which revealed that the set of drawings with contorted poses that were seen by the 'anomalous' group initially in the LVF were in general evaluated the fastest by all subjects, $F(3/93) = 2.8$; $P < 0.04$.

According to our theory, once both conventional and contorted poses of animals have been encoded and stored as global shapes after repeated presentations, these forms will be stored in the RH as specific pattern exemplars belonging to a familiar object; therefore, after practice, the initial RVF/LH advantage for contorted forms would disappear and a LVF/RH advantage would appear. Fig. 3 illustrates the average visual field difference (LVF – RVF) for each of the eight presentations, and shows that the RVF/LH advantage is clearly present for presentations of contorted poses (see a) only in the first trial; by the second trial, this difference has already reversed towards a LVF/RH advantage (indicated by a negative difference in RTs). This visual field reversal after the first trial is also clearly absent for presentations of conventional poses (see b).

2.2.2. Error rates

The same analyses were performed on the percent error data. Subjects found the identification of animals in contorted poses (mean % error = 16.9; SD = 9.9) more difficult than that of animals in conventional ones (mean % error = 4.4; SD = 5.7); $F(1/36) = 150$; $P < 0.0001$. In addition, they made more errors in the initial presentations (mean % error = 12.3; SD = 13) than the following ones (mean % error = 9.2; SD = 7.1), $F(1/36) = 15.7$; $P < 0.005$. No other significant findings

surfaced in this analysis, $P > 0.15$ in all cases. Clearly, there were no speed/accuracy trade-offs.

2.3. Discussion

We hypothesized that if an object projects a novel image, which cannot be matched as a single global shape to a stored representation, then individual parts of the pattern and categorical spatial relations among them will be used to identify the object. On the basis of previous research, we assumed that categorical spatial

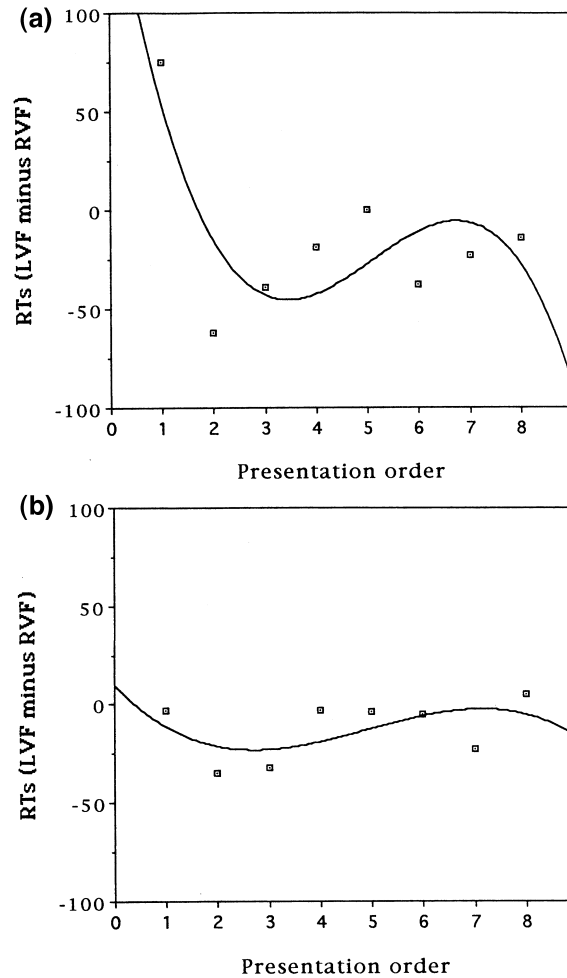


Fig. 3. Experiment 1. Means of differences between response times in the left visual field (LVF) and the right visual field (RVF) are plotted over each trial (1–8). The curved interpolating line between the means is the polynomial fit (order of 3). (a) Illustrates the visual differences for Contorted poses of animals whereas (b) illustrates the visual differences for Conventional.

relations are encoded better by mechanisms in the LH (e.g. Laeng, 1994) whereas specific forms are recognized better by mechanisms in the RH (e.g. Marsolek et al., 1992). As predicted, we found a RVF/LH advantage in response speed when animals in contorted poses were seen for the first time.

The findings also support the idea that the RH has better visual memory storage for specific forms, since there was a LVF/RH advantage for following presentations for both contorted and conventional poses. Additionally, this finding reveals that the LH advantage observed initially could not have been due to encoding words *per se*. In short, the findings are consistent with the hypothesis that the superior ability of the LH in encoding categorical spatial relations or in encoding parts (or both of these factors) plays a role in the initial identification of a novel contorted pose of a flexible shape. But once the overall shape has been encoded and stored, it can be matched directly, which does not rely on mechanisms that are superior in the LH but rather draws on mechanisms that are superior in the RH.

Our hypothesis that the default mechanism for object recognition is a global pattern matching process (see Kosslyn, 1994) implies that subjects should have identified conventional shapes better when they were initially presented to the LVF/RH. However, we observed only a small LVF advantage in response time, which failed to reach statistical significance. One possible explanation for this negative finding could be that some of the conventional views may not have been familiar to every subject. Thus, in the initial presentations of conventional views, subjects may have used a mixture of two hemispheric strategies (i.e. a RH-based global matching or a LH-based part-based matching), depending on the familiarity of each stimulus.

3. Experiment 2

The results from Experiment 1 leave open the issue of whether spatial relations between parts are actually encoded when one recognizes an object that is in a contorted pose. Indeed, an alternative interpretation for our finding of an initial RVF advantage for contorted forms would focus on the LH's superior ability to encode parts (Delis et al., 1986; Robertson and Delis, 1986; Robertson and Lamb, 1991). According to this view, parts of contorted forms could activate representations in associative memory (e.g. tail, whiskers) which, in turn, activate related semantic information about animals that possess such features (e.g. cat, rat, etc.). Thus, in our task, the name could be compared to candidate names linked to the activated semantic information. This hypothesis could account for the initial RVF/LH advantage for recognizing animals in contorted poses, and the disappearance of the RVF advantage after the initial exposure could be explained if the entire shape comes to be encoded.

In this study we investigated the role of encoding categorical spatial information and parts in shape identification. We reasoned as follows: if encoding spatial information is important when one identifies contorted objects, then the speed of encod-

ing such relations should predict the speed of recognizing such pictures. Moreover, if encoding parts is important when one identifies contorted objects, then the speed of encoding parts should predict the speed of recognizing such pictures. However, the results of Experiment 1 led us not to be interested in speed in general, but rather speed of processing when stimuli were presented in the LVF or RVF. Hence, in the following experiment we measured the differences in RT and accuracy between visual fields when subjects encoded categorical spatial information and when they recognized fragments as parts of a whole; following this, we examined how these measures predicted performance in the picture-name matching task of Experiment 1.

In this experiment we test three accounts for the hemispheric asymmetries observed in Experiment 1. First, if they are a result of differences in the ease of encoding parts per se, then we expect a positive correlation between the visual field differences for encoding parts and those for encoding the initial presentations of contorted shapes in picture-name matching. In addition, if this is the only relevant process, then we should not expect visual field asymmetries in encoding categorical spatial information to predict visual field differences for the initial presentations of contorted shapes in picture-name matching. Second, if encoding categorical spatial information is critical, then we expect a positive correlation between the observed asymmetry in encoding categorical spatial information and the visual field differences for initial presentations of contorted shapes in picture-name matching. Third, it is possible that encoding both parts and categorical spatial information play important roles when one names objects in contorted poses. In this case, both measures should predict the differential visual field performance in the picture-naming task. Therefore, we developed two new tasks and administered them to a new group of subjects along with the picture-name matching task used in Experiment 1. The whole-part matching task utilized the same animal figures that were used in the picture-name matching task. These were lateralized and then immediately followed by the presentation, in central vision, of a part of an animal's shape (which could either belong to the same drawing or to a drawing of a different animal); the subject was asked to decide whether the part was in fact a portion of the previously seen picture. The categorical spatial task also used pictures of common animals, but different drawings from those used in the other two tasks. In this task, the figure of an animal was lateralized and then immediately followed by a drawing of the same animal, in central vision, in either the same horizontal orientation (e.g. a rabbit was facing left in both the first and second frame) or in the opposite orientation (e.g. the rabbit was facing left in the first frame, but right in the second one). Studies with brain-damaged patients (see Laeng, 1994) have shown that lesions of the left hemisphere can cause difficulty in judging the left/right orientation of single shapes. Similarly, normal subjects show more efficient performance in judging this type of categorical spatial information when the spatially transformed stimuli are presented in the RVF/LH (Laeng and Peters, 1995). Thus, this task should provide a useful estimate of the degree and polarity of categorical spatial encoding function in normal individuals.

3.1. Method

3.1.1. Subjects

Twelve female and 12 male right-handed students at Harvard University volunteered to serve as paid subjects. Handedness was assessed with the Edinburgh questionnaire. Each subject was assigned to one of two experimental groups, which counterbalanced across subjects the order of the picture-name matching task and the whole-part task, and each of these groups had an identical number of males and females. In addition, these groups were further divided into two groups to counterbalance visual field of presentation of the parts; that is, in each specific presentation, one group viewed the part in the opposite visual field than the other group. Counterbalancing for the picture-name matching task and the whole-part task is important because the same stimuli were used in both tasks (and—as seen in the early shift in hemispheric advantage in the previous experiment—a single exposure can be sufficient for a pose to become familiar). Because the categorical spatial relation task used different stimuli from the other tasks, it was not counterbalanced and was always performed first.

3.1.2. Stimuli and apparatus

Different stimuli were used in the different tasks, as described below.

Picture-name matching task. The stimuli used in Experiment 1 were also used here.

Whole-part matching task. The stimuli used in Experiment 1 were also used here as the ‘whole animal stimuli’. In addition, drawings of parts of the objects were produced by editing out (with Adobe Photoshop) a portion of the animal’s shape, leaving only a part of the animal. Parts were determined according to the following procedure. Twelve additional subjects (also students at Harvard, who did not participate in any of the experimental tasks) were given 64 sheets, each with one of the whole drawings used in the task, and were asked to circle with a pencil a part of the animal that they thought would help them the most in recognizing that particular animal. There was high agreement on the choice of the part (on average eight out of 12 circled the same portions of each drawing). The one part with the highest consensus was then edited out. Each part was fitted and centered (but without changing its lateral orientation within the whole figure or its size) inside a centrally presented box, which was the same size as that used with the whole animals. As in the previous task, for whole presentations, the head of the animal always faced the center of the screen center (i.e. rightward for LVF trials and leftward for RVF ones). Parts extracted from a figure of a different animal in the same set were used as distractors. Both correct parts and distractors appeared equally often in each visual field. Fig. 4 illustrates an example of whole and matching part drawings.

Categorical spatial information encoding task. Different drawings of animals (taken from various sources) than those used in the previous task were collected. These included two different figures of the following animals: Dog, frog, monkey, rabbit, rooster, sheep. The drawings depicted the animals in clear and highly conventional side views. The drawings were digitized as black and white PICT-files,

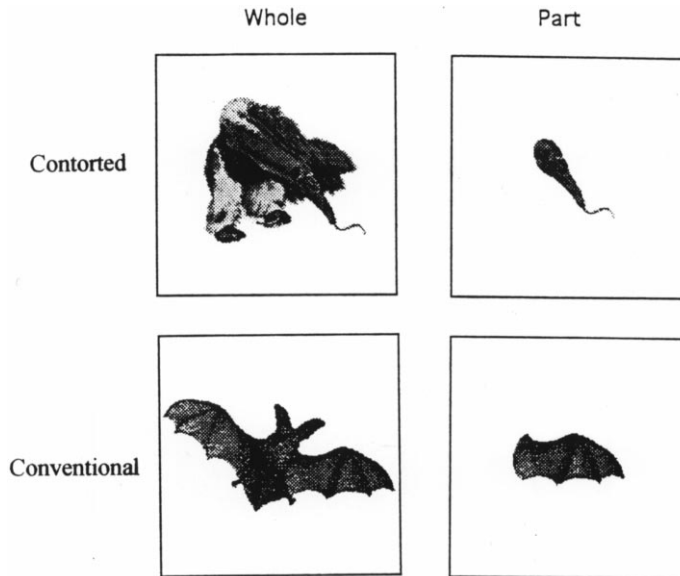


Fig. 4. Experiment 2. Examples of the stimuli used in the part-whole matching task. The left column shows examples of Contorted and Conventional poses of the whole animals, the right column shows just the part to be correctly matched to the same picture of the animal.

and the average size of these stimuli was made as close as possible to the average size of the stimuli used in the other tasks. For half of the trials the figure would change orientation from the first (lateralized) to the second (central) presentation and for half of the trials it would not. Again, in the lateralized presentations the head of the animal always faced towards the center of the screen (i.e. rightward for LVF trials and leftward for RVF ones). Both spatially transformed and non-transformed stimuli appeared equally often in each visual field.

3.1.3. Procedure

For all tasks, no more than three consecutive same-visual-field presentations of any stimulus were allowed or more than three consecutive true or false responses; consecutive trials of a same, or taxonomically closely related, animal's drawing (whole or part) and name also were avoided. Subjects sat upright and placed their heads on a chin-rest located 45 cm away from the computer screen for all three tasks. There were no practice trials for any of the tasks.

Picture-name matching task. The procedure for this task was identical to that of Experiment 1. However, only half of the trials in Experiment 1 were used ($n = 128$); each drawing was seen four times. Moreover, we used only one of the sequences of trials. Because the primary goal of this experiment was to examine how given individuals' visual field differences for the initial identification of contorted shapes correlate with the other measures, we selected one of the three trials sequences that in the previous experiment had revealed a consistent group's RVF/LH advantage for the initial presentations of contorted shapes.

Whole-part matching task. The procedure was the same as in the previous task, except that the subject's job was now to determine whether the part belonged to the previously seen whole figure. Thus, after the 100 ms lateralized presentation of the drawing of an animal in a box, and after a subsequent 50 ms of blank screen, the part of an animal would appear on the screen in a central box and remain in sight until the subjects, using the right hand's fingers, pressed one of two adjacent keys marked 'yes' and 'no' as quickly and accurately as possible. There were a total of 128 trials.

Categorical spatial information task. Subjects were told that they would see two drawings of the same animal in rapid sequence and their task was to determine whether the figures faced the same way. The sequence of events was as follows: (1) the subjects would initiate a trial by pressing the space bar on the computer's keyboard; (2) a small fixation cross appeared at the center of screen, which remained visible for 450 ms; the subjects were asked to gaze directly at the cross and maintain fixation until the end of the trial; (3) a blank screen appeared for 50 ms; (4) a 100 ms lateralized stimulus was presented; (5) a blank screen appeared for 50 ms, followed by a picture of the same animal in central vision; (6) the subject was asked to indicate, as quickly and accurately as possible, whether the animal faced the same way (left/right) in the two frames using the right hand's fingers to press one of two adjacent keys marked 'same' and 'different'. Following this, the screen would go blank and a new trial would begin. There were a total of 48 trials in the task.

3.2. Results

Means of RTs and percent errors were calculated for each combination of the following factors for both the picture-name and the whole-part matching tasks: pose (contorted versus conventional), presentation (presentation 1 vs. presentations 2–4), and visual field (left versus right). Thus, data were pooled over the type of match (correct versus incorrect) and items. For the categorical spatial task, means of RTs and errors were calculated for 'change' and 'no-change' in orientation. For all of these tasks, times from trials on which errors occurred were excluded from computation of the RTs, and trials with RTs longer than three standard deviations from each individual's mean RT were excluded from all the statistics (on average, no more than 2% of data was discarded for each subject in each task). Because preliminary analyses revealed that both the sex and the order of the two matching tasks had no effect, these factors were not included in any of the analyses described below.

3.2.1. Picture-name matching task

Response times. An analysis of variance with pose (contorted versus conventional), presentation (initial versus following presentations) and visual field (left versus right) as the within-subjects factors was performed on RTs as the dependent variable. For initial presentations of animals in contorted poses, the subjects evaluated the pictures faster when they were presented in the RVF/LH (mean RT = 972 ms; SD = 207) than in the LVF/RH (mean RT = 1047 ms; SD = 243), but for following presentations the subjects evaluated these types of pictures faster when

they were presented in the LVF/RH (mean RT = 866 ms; SD = 160) than in the RVF/LH (mean RT = 1014 ms; SD = 197). The analysis of variance revealed that this interaction between the factors of pose, familiarity and visual field was statistically reliable, $F(1/23) = 39.8$; $P < 0.0001$.

Because we had specific predictions about the effects of pose on the initial and following trials, we analyzed each of these sets of data in separate analyses of variance, with group as the between-subjects factor and visual field as the within-subjects factor. These tests confirmed the statistical reliability of both the visual field difference for initial presentations of animals in contorted poses (LVF/RH: mean RT = 1047 ms, SD = 243; RVF/LH: mean RT = 972 ms, SD = 207; $F(23) = 6.9$, $P < 0.01$), and that for following presentations (LVF/RH: mean RT = 866 ms, SD = 160; RVF/LH: mean RT = 1014 ms, SD = 197; $F(23) = 85.6$, $P < 0.0001$). The separate ANOVAs also revealed that the visual field differences for the conventional pose at initial presentations approached significance (LVF/RH: mean RT = 861 ms; SD = 201; RVF/LH: mean RT = 887 ms; SD = 214; $F(23) = 2.9$, $P < 0.09$) whereas it failed to reach statistical reliability for following presentations (LVF/RH: mean RT = 805 ms; SD = 160; RVF/LH: mean RT = 801 ms; SD = 182; $F(23) = 0.1$, $P < 0.72$). Fig. 5 illustrates the interaction of pose, novelty and visual field.

The overall analysis of variance also revealed that subjects recognized animals in conventional poses faster (mean RT = 838 ms; SD = 191) than contorted views (mean RT = 974 ms; SD = 212), $F(1/23) = 99.6$; $P < 0.0001$. Because of practice, subjects responded faster to following presentations (mean RT = 871 ms; SD = 193) than to initial presentations (mean RT = 942 ms; SD = 226), $F(1/23) = 29.5$; $P < 0.0001$. In addition, subjects were faster overall when pictures were presented in the LVF/RH (mean RT = 895 ms; SD = 212) than in the RVF/LH (mean RT = 918 ms; SD = 214), $F(1/23) = 9.3$; $P < 0.006$. Finally, for initial presentations the subjects evaluated the pictures faster when they were presented in the RVF/LH (mean RT = 929 ms; SD = 213) than in the LVF/RH (mean RT = 954 ms; SD = 240); in contrast, they evaluated following presentations faster when pictures were presented in the LVF/RH (mean RT = 835 ms; SD = 161) than in the RVF/LH (mean RT = 907 ms; SD = 216), $F(1/23) = 22.9$; $P < 0.0001$.

Error rates. The same analyses were performed with the error data. This revealed that subjects found animals in contorted poses more difficult to identify (mean % error = 12.3; SD = 14) than animals in conventional poses (mean % error = 9.7; SD = 8); $F(1/23) = 6.1$; $P < 0.02$. More errors were made on the initial presentations (mean % error = 18.1; SD = 11) than the following ones (mean % error = 4; SD = 6), $F(1/23) = 114.4$; $P < 0.0001$. In addition, this effect was greater for the animals in contorted poses (initial presentations: mean % error = 21; SD = 14; following presentations: mean % error = 4; SD = 7) than those in conventional poses (initial presentations: mean % error = 16; SD = 8; following presentations: mean % error = 4; SD = 4), $F(1/23) = 5.9$; $P < 0.02$. Finally, as shown in Fig. 6, less errors were made to animals in contorted poses when they were presented in the RVF/LH (mean % error = 8.5; SD = 11) than in the LVF/RH (mean % error = 16; SD = 16), whereas less errors to animals in conventional poses were made when

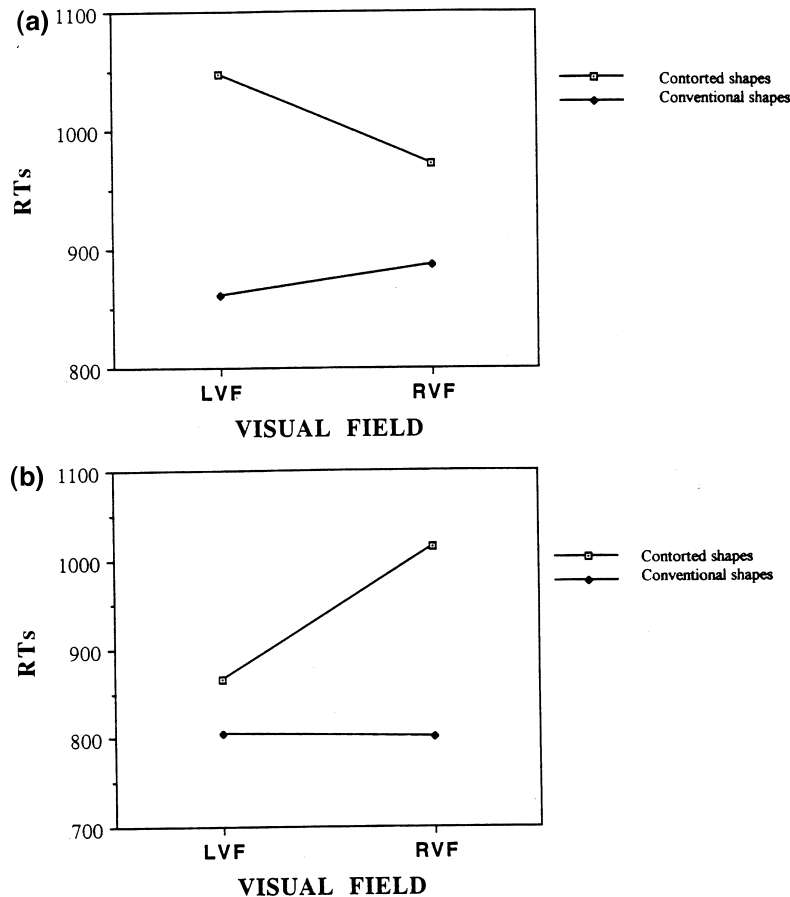


Fig. 5. Experiment 2. Means of response times to Contorted and Conventional shapes of animals in the left visual field (LVF) and the right visual field (RVF) for the picture-name matching task. (A) Shows the performance for the initial presentation only, whereas (B) shows it for all following presentations (2–4).

they were presented in the LVF/RH (mean % error = 8.3; SD = 7) than in the RVF/LH (mean % error = 11; SD = 9), $F(1/23) = 17.8$; $P < 0.0003$.

3.2.2. Whole-part matching task

Response times. An analysis of variance with pose (contorted versus conventional), presentation (initial versus following presentations) and visual field (left versus right) as the within-subjects factors was performed on RTs as the dependent variable. Perhaps the most interesting finding was that the speed in matching parts to whole animals in one visual field depended on their specific pose, $F(1/23) = 6.5$; $P < 0.02$. Subjects were faster in matching parts to contorted poses when the pictures were presented in the RVF/LH (mean RT = 829 ms; SD = 351) than in the LVF/RH (mean RT = 858 ms; SD = 334), $t(23) = 2.1$, $P < 0.4$. Although matches to animals viewed in conventional poses seemed faster when the pictures

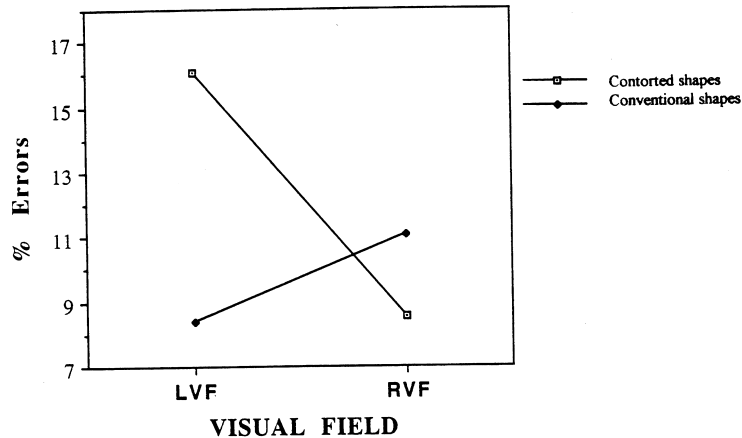


Fig. 6. Experiment 2. Means of the percent error rate for Contorted and Conventional shapes of animals in the left visual field (LVF) and the right visual field (RVF) for the picture-name matching task.

were presented in the LVF/RH (mean RT = 701 ms; SD = 185) than in the RVF/LH (mean RT = 722 ms; SD = 207), the post-hoc test failed to achieve reliability, $t(23) = 0.15$.

Moreover, subjects evaluated initial presentations faster when the pictures appeared in the RVF/LH (mean RT = 783 ms; SD = 318) than in the LVF/RH (mean RT = 821 ms; SD = 318), but had the reverse pattern for following presentations (mean RT in the LVF/RH = 738 ms; SD = 232; mean RT in the RVF/LH = 768 ms; SD = 283), $F(1/23) = 10.6$; $P < 0.004$. Post-hoc tests revealed that the RVF/LH advantage for initial presentations was reliable, $t(23) = 0.01$, and that there was a weak LVF/RH advantage for following presentations, $t(23) = 0.06$.

In addition, we found an interaction between the pose, presentation and visual field, which is illustrated in Fig. 7, $F(1/23) = 6.3$; $P < 0.02$. Subjects evaluated the initial presentations of animals in contorted poses faster when the pictures appeared in the RVF/LH (mean RT = 817 ms; SD = 358) than in the LVF/RH (mean RT = 925 ms; SD = 384), but there was no such RVF/LH advantage for following presentations (mean RT in the LVF/RH = 790 ms; SD = 266; mean RT in the RVF/LH = 840 ms; SD = 351). Furthermore, the subjects showed a tendency to evaluate initial presentations of animals in conventional poses faster in the LVF/RH (mean RT = 716 ms; SD = 191) than in the RVF/LH (mean RT = 748 ms; SD = 237), as well as following presentations (LVF/RH: mean RT = 685 ms; SD = 182; RVF/LH: mean RT = 695 ms; SD = 174). A post-hoc test confirmed that only the RVF/LH advantage for initial presentations of animals in contorted poses was statistically reliable, $t(23) = 3.03$, $P < 0.006$. Indeed, the presence of an initial RVF/LH advantage for matching parts of animals in contorted poses was very robust: Of the 24 subjects, 23 (96%) responded faster to the RVF/LH presentations than to the LVF/RH presentations. Moreover, for 16 of the former 23 subjects (70%) the RVF/LH

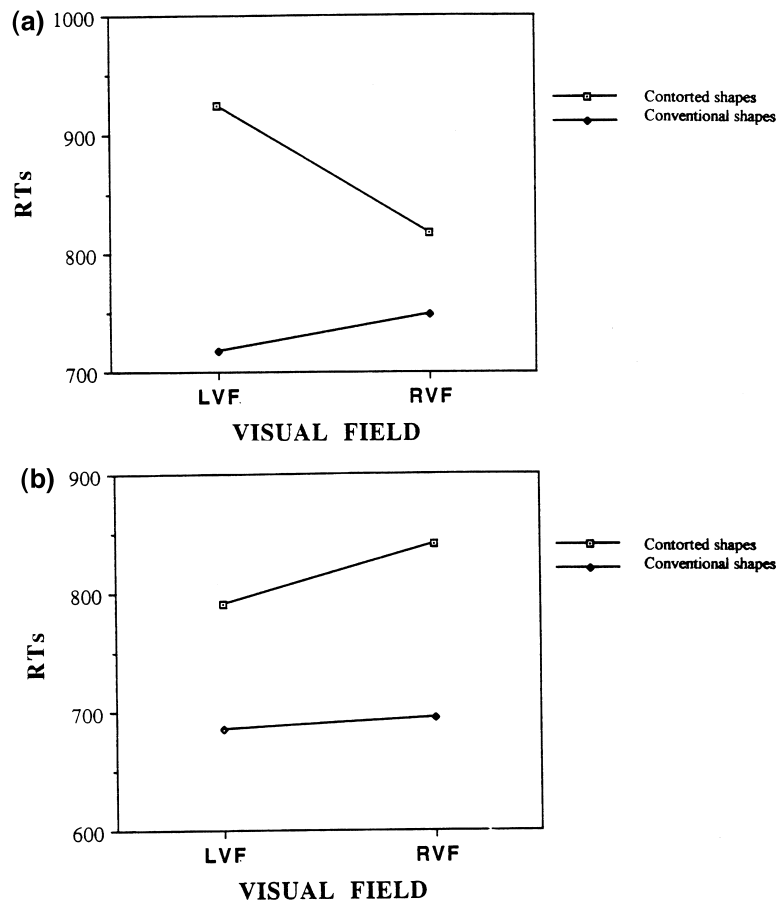


Fig. 7. Experiment 2. Means of response times to Contorted and Conventional shapes of animals in the left visual field (LVF) and the right visual field (RVF) for the whole-part matching task. (A) Shows the performance for the initial presentation only, whereas (B) shows it for all following presentations (2–4).

advantage (LVF > RVF) was greater than 100 ms, whereas the only subject who showed a LVF/RH (LVF < RVF) advantage had only an 88 ms difference.

Finally, this analysis revealed that parts were matched to animals viewed in conventional poses faster (mean RT = 711 ms; SD = 196) than to animals viewed in contorted poses (mean RT = 843 ms; SD = 341), $F(1/23) = 17.5$; $P < 0.0004$, and that subjects improved with practice, taking less time for following presentations (mean RT = 753 ms; SD = 258) than initial presentations (mean RT = 802 ms; SD = 309), $F(1/23) = 12.3$; $P < 0.002$.

Error rates. The same analysis was performed with the error data. Subjects matched parts to whole animals more accurately when the animals were viewed in conventional poses (mean % error = 9.3; SD = 9) than in contorted ones (mean % error = 15.3; SD = 16); $F(1/23) = 17.5$; $P < 0.0004$. More errors were made on the initial presentations (mean % error = 17.6; SD = 14) than the following ones (mean

% error = 6.8; SD = 10), $F(1/23) = 58.4$; $P < 0.0001$. In addition, the interaction of pose and visual field was statistically reliable, $F(1/23) = 7.5$; $P < 0.01$. Subjects evaluated the contorted poses more accurately when drawings were presented in the RVF/LH (mean % error = 13.4; SD = 15) than in the LVF/RH (mean % error = 17.1; SD = 18), but they did not show this effect when evaluating the conventional poses (LVF: mean % error = 7.5; SD = 8; RVF/LH: mean % error = 11; SD = 9). Post-hoc tests revealed that only the visual field difference for the contorted poses approached reliability, $t(23) = 1.9$, $P < 0.06$.

We also found an interaction of pose, presentation and visual field, $F(1/23) = 4.8$; $P < 0.04$. Subjects evaluated contorted poses more accurately for initial presentations in the RVF/LH (mean % error = 18; SD = 16) than in the LVF/RH (mean % error = 26; SD = 18), but not following presentations of contorted poses (LVF/RH: mean % error = 8.5; SD = 13; RVF/LH: mean % error = 8.9; SD = 13). There were less errors for initial presentations of conventional poses in the LVF/RH (mean % error = 11; SD = 10) than in the RVF/LH (mean % error = 16; SD = 9), and there appeared to be less errors for following presentations of conventional poses in the LVF/RH (mean % error = 4; SD = 4) than in the RVF/LH (mean % error = 6; SD = 5). Post-hoc tests revealed that only the RVF/LH advantage for initial presentations of contorted poses was statistically reliable, $t(23) = 2.7$, $P < 0.01$. Because this pattern of effects on the error rates is very similar to the effects previously described for the response times, there was clearly no speed-accuracy trade-off.

3.2.3. Categorical spatial information task

Response times. An analysis of variance with spatial orientation (change versus no change) and visual field (left versus right) as the within-subjects factors was performed on RTs as the dependent variable. This analysis revealed that subjects detected a change in spatial orientation faster than no change (with means of RT = 900 ms, SD = 288, and 967 ms, SD = 307, respectively), $F(1/23) = 7.4$; $P < 0.01$. In addition, subjects were faster overall when the stimulus appeared in the RVF/LH (mean RT = 903 ms; SD = 276) than in the LVF/RH (mean RT = 964 ms; SD = 319), $F(1/23) = 5.6$; $P < 0.03$. The interaction of spatial orientation and visual field failed to achieve significance ($F = 1$). However, post-hoc t -tests revealed that when a categorical spatial transformation occurred (i.e. the animal's left/right orientation changed) there was a RVF/LH advantage (LVF: mean RT = 945 ms; SD = 297; RVF/LH: mean RT = 855 ms; SD = 279), $t(23) = 2.3$, $P < 0.03$; whereas when there was no change in facing direction there was no visual field difference, $t < 1$.

Error rate. An analysis of variance with spatial orientation (change versus no change) and visual field (left versus right) as the within-subjects factors was performed on errors. This analysis revealed no statistically reliable effects or interactions; however, subjects had a tendency to detect a change in spatial orientation more accurately (mean % error = 15; SD = 8) than no change (mean % error = 20; SD = 12), $F(1/23) = 3.5$; $P < 0.07$. Given these error rate findings, there were clearly no speed-accuracy trade-offs.

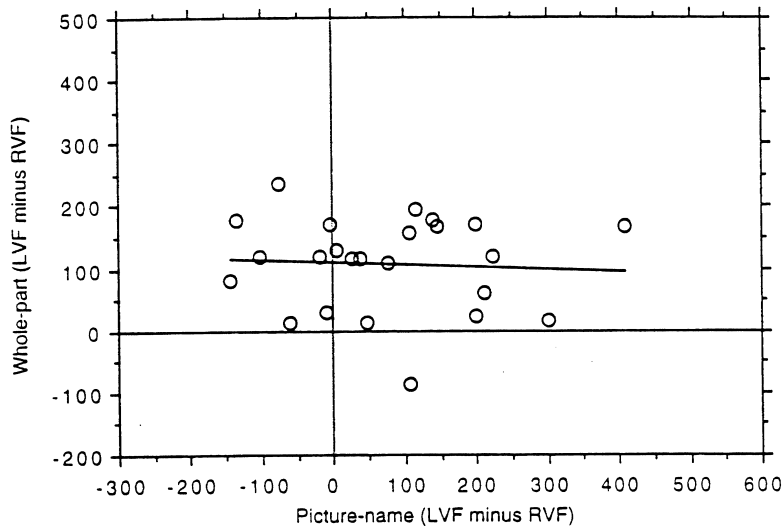


Fig. 8. Experiment 2. Scatterplot of each subject's mean response time visual field difference for the initial presentation of Contorted shapes for the picture-name and whole-part matching tasks. The straight interpolating line between the means is the simple regression.

3.2.4. Linear regressions of visual field differences

Three separate linear regression analyses were performed in order to evaluate whether the observed RVF/LH advantages in both matching a part to the whole and in judging categorical spatial transformations predict the observed RVF/LH advantage for the first identification of a contorted shape of an animal.

Correlation between the whole-part matching and picture-name matching tasks.

For each task, visual field differences in the initial presentation of an animal seen in contorted pose were computed for every subject by subtracting each subject's mean RT in the RVF/LH from each respective mean RT in the LVF/RH. Mean LVF/RH-RVF/LH RTs for each subject for each task were then submitted to a simple regression analysis. As evident in Fig. 8, most subjects had a RVF/LH advantage in the whole-part matching task. However, there was no correlation between these scores and those in the identification task, $R = 0.07$, $\beta = -0.04$, $t(23) = 0.3$, $P < 0.74$. Clearly, a greater RVF/LH advantage in matching parts to the whole in first presentations did not imply an advantage in identifying the contorted pose of an animal by name in this same first presentation.

Correlation between the picture-name matching and the categorical spatial transformation tasks. We next correlated each individual's difference in mean RTs (LVF/RH – RVF/LH) in the categorical spatial transformation task with the visual field difference in the picture matching task for the initial presentation of an animal seen in contorted pose. As evident in Fig. 9, performance on these tasks was in fact related, $R = 0.44$, $\beta = 0.35$, $t(23) = 2.3$, $P < 0.03$. Our overall estimate of each subject's degree and polarity of categorical spatial information encoding predicts the presence of a LH's superiority in identifying novel contorted shapes in these same individuals.

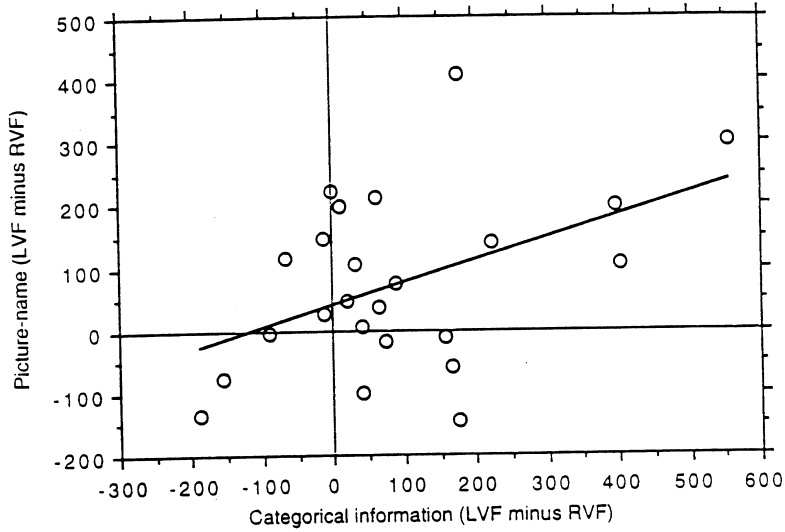


Fig. 9. Experiment 2. Scatterplot of each subject's mean response time visual field difference for the initial presentation of Contorted shapes in the picture-name matching task and for performance in the categorical spatial relation task. The straight interpolating line between the means is the simple regression.

An additional correlation was also performed between each individual's difference in mean RTs (LVF/RH – RVF/LH) in the categorical spatial transformation task with the visual field difference in the picture matching task for the initial

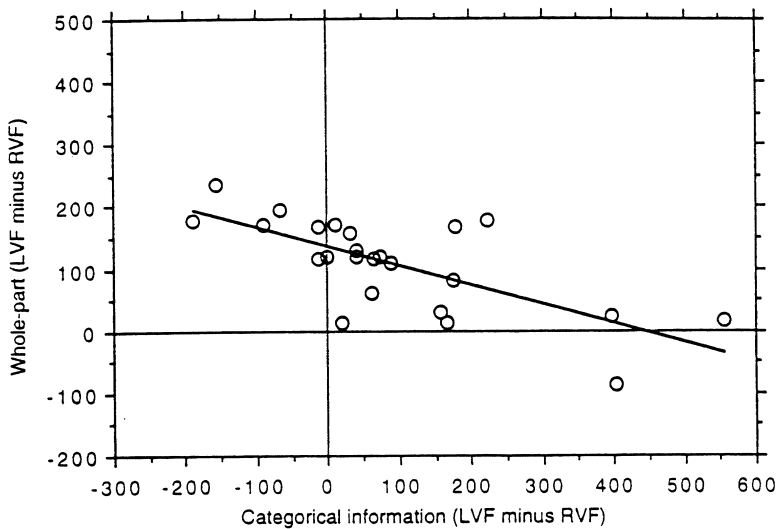


Fig. 10. Experiment 2. Scatterplot of each subject's mean response time visual field difference for the initial presentation of Contorted shapes in the whole-part matching task and the whole performance in the categorical spatial relation task. The straight interpolating line between means is the simple regression.

presentation of an animal seen in a conventional pose. In contrast with the previous regression analysis, we would now expect no correlation between these two scores, because spatial information encodings should be irrelevant for the identification of conventional shapes. And indeed, this analysis revealed no reliable relation between the two variables, $R = 0.21$, $\beta = -0.09$, $t(23) = 1.0$, $P < 0.32$.

Correlation between the whole-part matching and the categorical spatial transformation tasks. The visual field differences in the whole-part matching task (for the initial presentation of an animal seen in contorted pose) and categorical spatial transformation tasks were also correlated. Surprisingly, there was a statistically reliable negative correlation between these two scores, as illustrated in Fig. 10; $R = -0.71$, $\beta = -0.31$, $t(23) = 4.7$, $P < 0.0001$. Considering that all subjects except one responded faster when stimuli were presented in the RVF/LH than in the LVF/RH in the part-whole matching task, the best way to describe this relationship is that a LH's superiority in the spatial transformation task predicted a decrease in the LH's superiority in matching parts to whole of new contorted shapes.

3.3. Discussion

First, we again found that subjects identified animals in contorted poses on the first presentation better when the pictures were presented in the RVF, and hence were encoded initially by the LH. Moreover, we again found that once the specific shapes were memorized, subjects subsequently identified them better when they were presented in the LVF, and hence were encoded initially by the RH. Second, the same effect occurs when a part of the contorted shape is matched to a previously seen global shape. Third, the individual's estimated degree and polarity of hemispheric specialization for categorical spatial relations encoding was related to the LH advantage for identifying novel contorted shapes but had no relationship to this hemisphere ability in identifying conventional shapes.

The latter finding is particularly interesting because it indicates that spatial information is involved in processing novel contorted shapes. If we consider that virtually all of our subjects encoded parts of a contorted shape more efficiently in the RVF/LH, whereas they varied in the laterality estimate of spatial information encoding, it is not surprising that visual field differences for categorical spatial information predicted the RVF/LH's ability to identify a new contorted pose. However, it is worth noting that the ease of encoding parts in the RVF/LH did not, by itself, predict the ease of identifying a new contorted pose in the RVF/LH. As theorized, although parsing is a necessary prerequisite for recovering the object's structural description, one hemisphere's advantage for such operation is clearly not sufficient by itself to guarantee a better same-hemisphere's recognition of contorted shapes. The same hemisphere must also possess an advantage for encoding categorical spatial relations. In this case, the LH apparently can best use a combination of the two types of visual information to access and match structural descriptions.

However, unexpectedly, we found that LH superiority in the spatial task was associated with LH inferiority in matching parts to whole of novel contorted shapes (Fig. 11). One possible explanation for this negative correlation could be that the

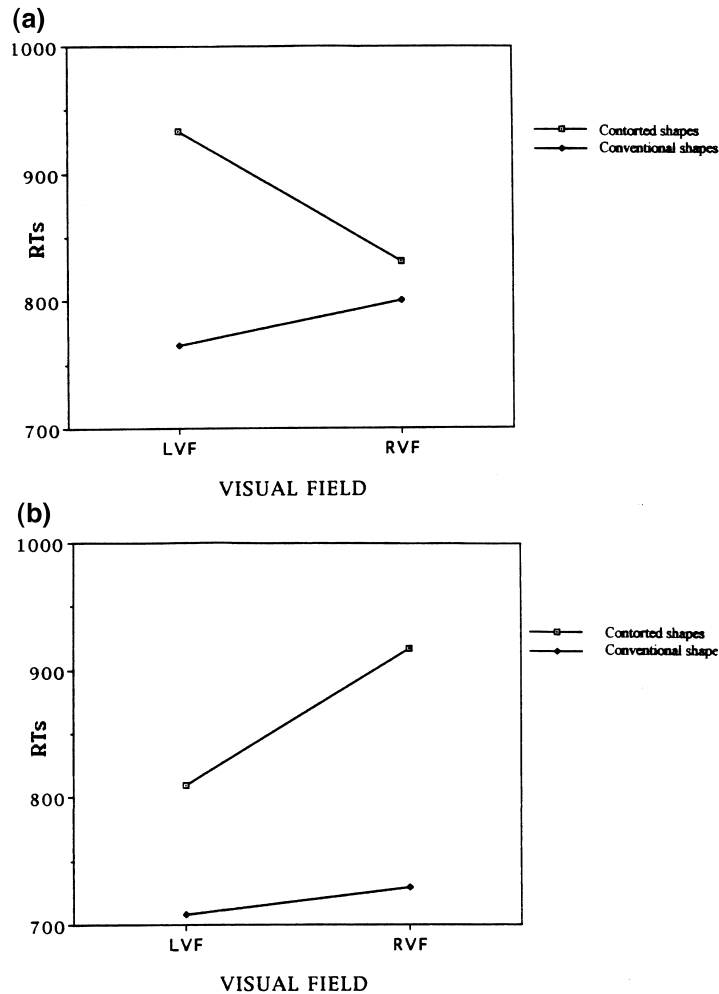


Fig. 11. Experiment 3. Means of response times for Contorted and Conventional shapes of animals in the left visual field (LVF) and the right visual field (RVF) for the picture-name matching task. (A) Shows the performance for the initial presentation only, whereas (B) shows it for all following presentations (2–4).

more resources the LH uses in performing parsing, the less efficient it becomes in performing categorical spatial relations processing. This suggests to us that although parsing and spatial relations computations must cooperate in the same hemisphere to achieve structural descriptions, they can also interfere with each other (possibly because they compete for the same attentional resources and/or mechanisms; cf. Friedman and Polson, 1981). If the subjects who had a moderate advantage in parsing are efficient in encoding categorical spatial relations in the same (left) hemisphere, they will be in turn the most capable of matching structural descriptions of the novel contorted shapes in the LH. And indeed, subjects who had a strong RVF/

LH advantage in the categorical spatial task also showed the clearest RVF/LH advantage in the initial recognition of contorted shapes.

Finally, we found a weak LVF/RH advantage in matching parts of familiar forms (i.e. forms in conventional poses and contorted shapes once they have become familiar). This unexpected finding raises questions about the ability of the whole-part matching task to reveal part processing. In turn, this may suggest that our failure to find a correlation between the part/whole task and the first identification of contorted forms in the RVF/LH reflects the task's failure to properly address the status of part processing in shape identification. Yet, this task also revealed a strong RVF/LH advantage for matching parts to novel contorted forms. One possible interpretation for these results is that the LH's advantage is really for the *extraction* of a part from the whole shape in the initial processing of contorted shapes. Each shape, we claimed, needs to be recognized (made sense from a flashed image) by a process of extracting parts followed by a matching to a structural description. However, such processing would not imply that the RH cannot match efficiently the extracted information about parts to a portion of the contour of a stored (familiar) global shape.

4. Experiment 3

In this experiment we further investigated the relationship between the ability to encode categorical spatial relations in one hemisphere and the ability to identify novel contorted shapes of animals. First of all, we aimed to replicate the positive correlation between the LVF-RVF differences in the categorical spatial relation task and the picture-name matching task. For this purpose we tested a new group of subjects with both tasks. In this replication, we slightly modified the original categorical spatial task by using the same stimuli we used in the picture-name matching task. The main reason for such change was that we had previously used only highly conventional shapes for this task. The argument that categorical spatial relations are essential for identifying contorted shapes would be strengthened if we could demonstrate that the encoding of categorical spatial information of these contorted shapes is related to successfully identifying the object.

4.1. Subjects

Six female and six male right-handed students at the University of Tromsø (Norway) volunteered to participate as paid subjects. Handedness was assessed with a Norwegian translation of the Edinburgh questionnaire.

4.2. Stimuli and apparatus

The same 64 black-and-white line drawings of animals used in the picture-name tasks of the two previous experiments were also used as stimuli for the categorical spatial task. Both tasks were presented on the screen of a Macintosh PowerBook 1400cs/133. The stimuli appeared either 2.7 cm to the left or 2.7 cm to the right of

central fixation and the head of the depicted animal was always toward the screen center. In the picture-name matching task, the animals' names were presented visually as words in Norwegian (block Geneva letters, size 18) appearing 4 cm from the lower center portion of the screen. All sequences of trials and storage of key presses were controlled by MacLab software.

4.2.1. *Design*

Because the two tasks used the same stimuli, in order to avoid the possibility that familiarity with the pictures would affect the initial hemisphere/visual field differences in the picture-name matching task, the subjects always performed this task first. In the new categorical spatial task, each of the 32 pairs of drawings was shown twice in each visual field, half of the time followed by a central presentation of the same drawing in the same orientation and half of the time in the opposite lateral orientation. The sequence of presentation was semi-random, which was random except that animals in the same species could not be presented consecutively and there could be no more than three consecutive presentations of the same response (change/no change in orientation), visual hemifield (left/right), or pose (contorted/conventional). In addition, in order to rule out the possibility that the previously observed hemispheric advantages were an artifact of the delay between the visual and verbal stimuli, the verbal label now followed immediately after the offset of the visual stimulus (i.e., at zero ms delay).

4.2.2. *Procedure*

Subjects sat upright, facing directly the computer screen at a distance of about 45 cm without use of a chin-rest. There were no practice trials. Subjects received the same instructions for each task that we used in the previous experiment. The picture-name matching task was practically identical to the one used in Experiment 2; the only modifications were that (1) the presentation of the lateral visual stimulus was reduced to 75 ms (in order to compensate for possible visual persistence due to the PowerBook's screen passive matrix), (2) verbal labels were presented visually and remained visible until the subject made a key press, and (3) there was no delay between the offset of the visual stimulus and the onset of the verbal label.

4.3. *Results*

Means of RTs and percent errors were calculated in the picture-name matching task for pose (contorted versus conventional), presentation (presentation 1 versus presentations 2-4), and visual field (left versus right). For the categorical spatial task, means of RTs and errors were calculated for change and no-change in orientation. For both tasks, RTs longer than three standard deviations from each individual's mean RTs for that condition were excluded from the statistics (on average, no more than 1.5% of the data was discarded for each subject on both tasks). Since preliminary analyses revealed that both the sex and the type of match (correct versus incorrect) did not interact with any of the other factors, these factors were not included in the analyses described below.

4.3.1. Picture-name matching task

Response times. An analysis of variance with pose (contorted versus conventional), presentation (initial versus following presentations) and visual field (left versus right) as the within-subjects factors was performed on RTs as the dependent variable. We again observed a RVF/LH advantage for initial presentations of animals in contorted poses (LVF: mean RT = 933 ms; SD = 131; RVF/LH: mean RT = 831 ms; SD = 122), as well as the finding that, with following presentations, subjects evaluated these types of pictures faster when they were presented in the LVF/RH (mean RT = 809 ms; SD = 84) than in the RVF/LH (mean RT = 917 ms; SD = 110). The ANOVA revealed that the interaction between the factors of pose, familiarity and visual field was statistically reliable, $F(1/11) = 36.5$; $P < 0.0001$. We also tested each of our specific predictions on the effects of pose on the initial and following trials, by analyzing each of these sets of data in separate analyses of variance, with group as the between-subjects factor and visual field as the within-subjects factor. The separate anovas confirmed the visual field difference for initial presentations of animals in contorted poses, $F(1/11) = 10.7$, $P < 0.007$, and that for following presentations, $F(1/11) = 61.5$, $P < 0.0001$. Importantly, we now observed the expected LVF/RH advantage for the conventional pose at initial presentations (LVF/RH: mean RT = 765 ms; SD = 109; RVF/LH: mean RT = 800 ms; SD = 95; $F(1/11) = 7.3$, $P < 0.02$). The LVF/RH advantage also approached statistical reliability for following presentations of conventional views (LVF/RH: mean RT = 708 ms; SD = 86; RVF/LH: mean RT = 729 ms; SD = 82; $F(1/11) = 3.9$, $P < 0.07$).

The other statistically reliable effects were analogous to those seen in the previous two experiments: (1) subjects recognized animals in conventional poses faster (mean RT = 838 ms; SD = 191) than contorted views (mean RT = 974 ms; SD = 212), $F(1/23) = 99.6$; $P < 0.0001$; (2) they responded faster to following presentations of a drawing (mean RT = 871 ms; SD = 193) than to initial presentations (mean RT = 942 ms; SD = 226), $F(1/23) = 29.5$; $P < 0.0001$; and (3) subjects evaluated the pictures faster, for initial presentations, when they were presented in the RVF/LH (mean RT = 929 ms; SD = 213) than in the LVF/RH (mean RT = 954 ms; SD = 240), but they evaluated following presentations faster when pictures were presented in the LVF/RH (mean RT = 835 ms; SD = 161) than in the RVF/LH (mean RT = 907 ms; SD = 216), $F(1/23) = 22.9$; $P < 0.0001$.

Error rates. The same analysis was performed with the percent error data. More errors were made for animals in contorted poses (mean % error = 13.3; SD = 14.2) than animals in conventional poses (mean % error = 9.5; SD = 7.5), $F(1/11) = 5.8$; $P < 0.03$. More errors were made on the initial presentations (mean % error = 18.9; SD = 11) than the following ones (mean % error = 3.8; SD = 5.4), $F(1/11) = 72$; $P < 0.0001$. In addition, subjects made more errors for animals in contorted poses when they were presented in the LVF/RH (mean % error = 17.9; SD = 17) than when they were presented in the RVF/LH (mean % error = 8.5; SD = 9), $F(1/11) = 35.4$; $P < 0.0001$. No other effects were statistically reliable.

4.3.2. Categorical spatial information task

Response times. An analysis of variance with pose (contorted versus conventional), spatial orientation (change versus no change), and visual field (left versus right) as the within-subjects factors was performed on RTs as the dependent variable. Subjects detected a change in spatial orientation (mean RT = 739 ms, SD = 130) faster than no change (mean RT = 790 ms, SD = 129), $F(1/11) = 14$; $P < 0.003$. Most important, subjects were faster overall when the stimulus appeared in the RVF/LH (mean RT = 745 ms; SD = 128) than in the LVF/RH (mean RT = 784 ms; SD = 134), $F(1/11) = 5.1$; $P < 0.04$. The interaction of spatial orientation and visual field was statistically reliable, $F(1/11) = 6.5$; $P < 0.03$. It appeared that RTs were faster in the RVF/LH (mean RT = 697 ms, SD = 90) than in the LVF/RH (mean RT = 781 ms, SD = 151) when a categorical spatial transformation occurred (i.e. the animal's left/right orientation changed), but there was no such RVF/LH advantage when the two drawings to be matched faced in the same direction (LVF/RH: mean RT = 788 ms, SD = 117; RVF/LH: mean RT = 792 ms, SD = 143). There was no difference in time for the two poses, nor did this factor interact with the others.

Error rates. An analysis of variance with spatial orientation (change versus no change) and visual field (left versus right) as the within-subjects factors was performed on percent errors. This analysis revealed no statistically reliable effects or interactions.

Correlation between the picture-name matching and the categorical spatial information tasks. We correlated each individual's difference in mean RTs (LVF – RVF/LH) for the whole categorical spatial information task with the visual field difference in the picture-name matching task for the initial presentation of an animal seen in a contorted pose. Fig. 12 illustrates that the two scores were, as we expected, positively related, $R = 0.69$, $\beta = 0.42$, $t(11) = 3.01$, $P < 0.01$. Therefore, when the estimate of subjects' degree and polarity of categorical spatial information encoding was based on judgments of contorted animal shapes, we were capable of predicting the presence of a LH superiority in identifying novel contorted shapes. A test of the difference between this correlation coefficient and that obtained in the previous experiment showed that there was no significant difference, Exp 2: $r' = 0.476$, Exp 3: $r' = 0.848$, $z = 0.93$ (Fisher, 1921).

We also performed a correlation between each individual's difference in mean RTs (LVF/RH – RVF/LH) in the categorical spatial transformation task with the visual field difference in the picture matching task for the initial presentation of an animal seen in a conventional pose. Again, as pointed out earlier, we would now expect no correlation between these two scores, since according to our hypotheses the spatial information encodings play no significant role in the identification of conventional shapes. Indeed, the analysis revealed no reliable relation between the two variables, $R = 0.34$, $\beta = -0.05$, $t(23) = 1.1$, $P < 0.29$.

4.4. Discussion

Experiment 3 successfully replicated the major findings of Experiment 2, namely that a LH superiority in encoding categorical spatial relations is related to a LH

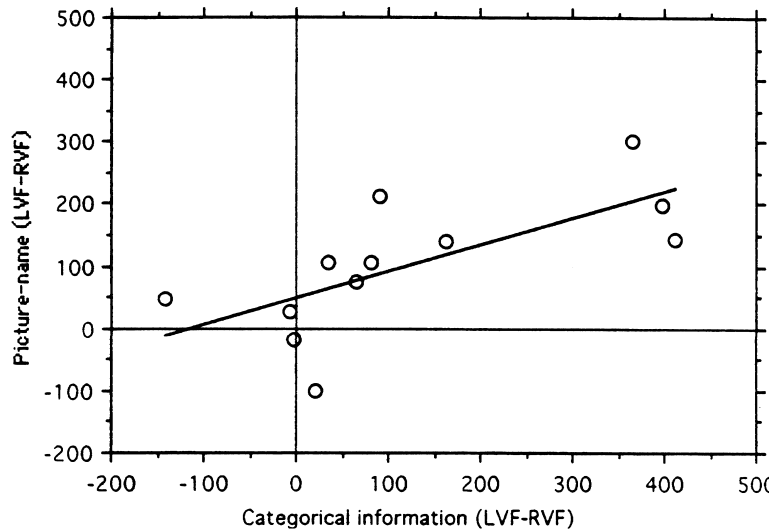


Fig. 12. Experiment 3. Scatterplot of each subject's mean response time visual field differences for the initial presentation of Contorted shapes in the picture-name matching task and for performance in the categorical spatial relation task. The straight interpolating line between the means is the simple regression.

superiority in identifying contorted shapes that are seen for the first time. Thus, the replication strengthens the conclusion that the LH's ability to encode categorical spatial information is indeed related to its ability to identify contorted objects. For the third time we found a RVF/LH advantage in identifying a contorted shape in the first exposure, whereas with following exposures of the same shapes the field advantage switches to a LVF/RH.

In addition, as predicted, in this experiment we found that subjects identified conventional shapes faster in the LVF/RH than in the RVF/LH when they were first shown. We should note that in all three experiments subjects identified initial presentations of conventional shapes faster in the LVF than in the RVF, although in Experiment 1 this difference failed to achieve significance and in Experiment 2 it only approached statistical reliability. We speculate that the conventional views may have occasionally included poses of animals that were unfamiliar to some subjects, which would have weakened the visual field differences.

Finally, we note that despite the fact that we varied several parameters of both encoding tasks from the previous two experiments (i.e. in the picture-name task there was now no delay between picture and name presentations, and in the spatial task both contorted and non-contorted shapes were used), the findings remained virtually identical to those previously observed. It would seem that the effects described here are not limited to a narrow range of experimental manipulations.

5. General discussion

A left-hemisphere advantage was repeatedly found when subjects first encoded

contorted poses of non-rigid objects (animals' bodies in this specific case). However, after the form was identified even a single time, this left-hemisphere's advantage was replaced by a right-hemisphere advantage. A right-hemisphere advantage was also observed for the same non-rigid objects when these were shown in more conventional (possibly already familiar) poses. This unique pattern of results is consistent with the hypothesis that identification of objects that assume unusual shape configurations is achieved by (1) encoding the object's parts separately, (2) encoding the spatial relations among the parts, and (3) matching these encodings to a stored structural description; moreover, (4) once this form has become familiar, its global shape can be directly matched to information stored in memory. In addition, we hypothesized that the left cerebral hemisphere encodes better (1) the parts of an object, which are taken to be essential components of structural descriptions, and also (2) 'categorical' spatial relations among parts, which complete the structural descriptions; furthermore, we hypothesize that (3) the right hemisphere plays a complementary role to the left because it is able to encode better the global form of the object. This set of hypotheses led to the prediction of a left-hemisphere advantage only when one first encodes contorted poses and to the prediction of a right-hemisphere advantage for encoding these same forms once they become familiar or if already familiar (i.e. if shown in conventional poses).

The results from the three experiments reported here confirmed these predictions and, further, revealed that the degree and lateralization of categorical spatial encoding can predict the left hemisphere's initial advantage in the identification of contorted shapes. The latter finding clearly points to the importance of spatial information in processing non-rigid shapes. We conclude from our findings that there are at least two ways to achieve object recognition, one that relies crucially on mechanisms in the LH and one that relies crucially on mechanisms in the RH. Which one will prevail in a given situation depends on multiple factors, including whether a specific view or pose of an object is either familiar or unfamiliar.

Indeed, previous theories of object identification typically have stressed only some of the aspects of processing we have noted here. For example, several theorists have proposed that understanding a picture often requires an iterative perceptual process that begins with selecting salient cues (which could constitute a rough sketch or schema of what the picture may represent) which in turn provide the basis for a guided search, which is under the control of perceptual feedback. Bruner (1957) proposed that object identification begins with a perceptual readiness, characterized by the selection of relevant information and rejection of irrelevant information; this information then serves as a perceptual hypothesis that, via repeated verifications and falsifications, drives additional encoding. According to several current theories of visual perception, parts of objects are always extracted from the visual input, because by necessity (for such theories) objects are nothing more than aggregates of parts (i.e. geometric primitives arranged in a characteristic spatial layout), and parts provide the indices to stored structural descriptions of the object (e.g. Sutherland, 1968; Milner, 1974; Winston, 1975; Marr, 1982; Hoffman and Richards, 1986; Biederman, 1985).

However, there are situations in which object identification may not require that

parts be encoded explicitly. In some cases, objects seen in familiar views may have been stored as templates or a unique set of primitives (e.g. non-accidental properties; Lowe, 1987), which can be matched globally to stored representations. Indeed, there is evidence that object recognition can occur even before figure-ground organization (Peterson, 1994). The global contour of partially occluded objects can also be completed, on the basis of depth cues in the early input, so as to facilitate template matching (Nakayama et al., 1989). These and other considerations (see Kosslyn, 1994) cast doubt on the idea that parsing objects into their natural parts must precede recognition and identification. Indeed, as Ullman (1996) points out, there may not be a single ‘correct’ account of how object recognition and identification are achieved; instead, many of the current approaches may characterize possible methods that are sometimes used by the brain in dealing with different types of objects and object transformations.

Our proposal is that structural descriptions are stored and can guide the system to look for parts in order to match the input and stored visual representations. However, matches will be attempted over the most comprehensive region of the visual input that appears as a perceptually coherent unit, and only when these attempted matches fail will matches to smaller portions be attempted (cf. Navon, 1977; Robertson and Palmer, 1983). Thus, global matching would have priority in visual perception, and recognition by parts through structural description would intervene at a later phase, particularly in ‘visually challenging’ situations.

Our study provides partial support for this view. We proposed that non-rigid forms, even familiar ones like animals, may often present exactly the type of visual challenge described above. This view led us to make clear predictions about normal individuals’ performance in a divided-visual-field task, which were supported by the results. However, the details of our findings suggest that the spatial relations among parts play at least as important a role as the encoding of parts themselves.

Acknowledgements

During the initial phase of this study Bruno Laeng was supported by a JSMF grant (No. 94-34) from the James McDonnell Foundation and The Pew Charitable Trusts and, in the final stage, by internal funds from the University of Tromsø (Norway); Stephen Kosslyn was supported by grant AG 12675 and Jinesh Shah by a grant from the Harvard College Research Program.

References

- Biederman, I., 1985. Human image understanding: recent research and a theory. *Computer Vision, Graphics, and Image Processing* 32, 29–73.
- Bruner, J., 1957. On perceptual readiness. *Psychological Review* 64, 123–152.
- Cave, C.B., Kosslyn, S.M., 1989. The role of parts and spatial relations in object identification. *Perception* 22, 229–248.

- Costin, D., 1988. MacLab: a Macintosh system for psychology labs. *Behavior, Research Methods, and Computers* 20, 197–200.
- Delis, D.C., Robertson, L.C., Efron, R., 1986. Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia* 24, 205–214.
- Driver, J., Baylis, G., 1995. One-side edge assignment in vision. 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science* 4, 201–206.
- Fisher, R.A., 1921. On the probable error of a coefficient of correlation deduced from a small sample. *Metron* 1, 2–32.
- Friedman, A., Polson, M.C., 1981. Hemispheres as independent resource systems: Limited-capacity processing and cerebral specialization. *Journal of Experimental Psychology: Human Perception and Performance* 7, 1031–1058.
- Harter, J., 1979. *Animals*. 1419 copyright-free illustrations of mammals, birds, fish, insects, etc. Dover Publications, New York.
- Hellige, J.B., Michimata, C., 1989. Categorization versus distance: Hemispheric differences for processing spatial information. *Memory and Cognition* 17, 770–776.
- Hoffman, D., Richards, W., 1986. Parts of recognition. In: Pentland, A.P. (Ed.), *From Pixels to Predicates*. Ablex, Norwood, NJ.
- Hoffman, D.D., Singh, M., 1997. Saliency of visual parts. *Cognition* 63, 29–78.
- Kosslyn, S.M., 1987. Seeing and imaging in the cerebral hemispheres: a computational approach. *Psychological Review* 94, 148–175.
- Kosslyn, S.M., 1994. *Image and Brain*. The MIT Press, Cambridge, MA.
- Kosslyn, S.M., Chabris, C.F., Marsolek, C.J., Koenig, O., 1992. Categorical versus coordinate spatial relations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance* 18, 562–577.
- Kosslyn, S.M., Flynn, R.A., Amsterdam, J.B., Wang, G., 1990. Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition* 34, 203–277.
- Kosslyn, S.M., Koenig, O., 1995. *Wet Mind: The New Cognitive Neuroscience*. The Free Press, New York.
- Kosslyn, S.M., Koenig, O., Barrett, A., Cave, C.B., Tang, J., Gabrieli, J.D.E., 1989. Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance* 15, 723–735.
- Laeng, B., 1994. Lateralization of categorical and coordinate spatial functions. A study of unilateral stroke patients. *Journal of Cognitive Neuroscience* 6, 189–203.
- Laeng, B., Peters, M., 1995. Cerebral lateralization for the processing of spatial coordinates and categories in left- and right-handers. *Neuropsychologia* 33, 421–439.
- Laeng, B., Peters, M., McCabe, B., 1998. Memory for locations within regions. Spatial biases and visual hemifield differences. *Memory and Cognition* 26, 97–107.
- Leclerc, G.-L., 1993. *368 animal illustrations from Buffon's 'Natural history'*. Dover Publications, New York.
- Lowe, D.G., 1985. *Perceptual organization and visual recognition*. Kluwer, Boston, MA.
- Lowe, D.G., 1987. The viewpoint consistency constraint. *International Journal of Computer Vision* 1, 57–72.
- Marr, D., 1982. *Vision*. Freeman, San Francisco, CA.
- Marsolek, C.J., Kosslyn, S.M., Squire, L.R., 1992. Form-specific visual priming in the right cerebral hemisphere. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 492–508.
- Marsolek, C.J., Squire, L.R., Kosslyn, S.M., Lulenski, M.E., 1994. Form-specific explicit and implicit memory in the right cerebral hemisphere. *Neuropsychology* 8, 588–597.
- Marsolek, C.J., 1995. Abstract visual-form representations in the left cerebral hemisphere. *Journal of Experimental Psychology: Human Perception and Performance* 21, 375–386.
- Milner, P.M., 1974. A model for visual shape recognition. *Psychological Review* 81, 521–535.
- Nakayama, K., Shimojo, S., Silverman, G.H., 1989. Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18, 550–568.

- Navon, D., 1977. Forest before trees: the precedence of global features in visual perception. *Cognitive Psychology* 9, 353–383.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 5, 127–134.
- Palmer, S., Rosch, E., Chase, P., 1981. Canonical perspective and the perception of objects. In: Long, J., Baddeley, A. (Ed.), *Attention and Performance IX*. Erlbaum, Hillsdale, NJ, pp. 131–151.
- Peterson, M., 1994. Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science* 3, 105–111.
- Richards, W., Dawson, B., Whittington, D., 1988. Encoding contour shape by curvature extrema. In: Richards, W. (Ed.), *Natural Computation*. The MIT Press, Cambridge, MA, pp. 83–98.
- Robertson, L.C., Delis, D.C., 1986. 'Part-whole' processing in unilateral brain-damaged patients: Dysfunctions of hierarchical organization. *Neuropsychologia* 24, 363–370.
- Robertson, L.C., Lamb, M.R., 1991. Neuropsychological contributions to theories of part-whole organization. *Cognitive Psychology* 24, 363–370.
- Robertson, L.C., Palmer, S.E., 1983. Holistic processes in the perception and transformation of disoriented figures. *Journal of Experimental Psychology: Human Perception and Performance* 9, 203–214.
- Saiki, J., Hummel, J.E., 1998. Connectedness and the integration of parts with relations in shape perception. *Journal of Experimental Psychology: Human Perception and Performance* 24, 227–251.
- Sutherland, N.S., 1968. Outline of a theory of visual pattern recognition in animal and man. *Proceedings of the Royal Society of London (B)* 171, 297–317.
- Ullman, S., 1996. *High-level vision. Object recognition and visual cognition*. The MIT Press, Cambridge, MA.
- Winston, P.H., 1975. *The Psychology of Computer Vision*. McGraw-Hill, New York, NY.
- Wolfe, J.M., Bennett, S.C., 1997. Preattentive object files: shapeless bundles of basic features. *Vision Research* 37, 25–43.