Mirror-image relations in category learning

Ingo Rentschler

Institut für Medizinische Psychologie, Universität München, Germany

and

Martin Jüttner

School of Life and Health Sciences - Psychology, Aston University, Birmingham, UK

Reference for citations:

Rentschler, I. & Jüttner, M. (2007). Mirror-image relations in category learning. *Visual Cognition* 15, 211-237.

Please address all correspondence to M. Jüttner, School of Life and Health Sciences - Psychology, Aston University, Aston Triangle, Birmingham B4 7ET, UK. Email: m.juttner@aston.ac.uk

This study was supported by the Deutsche Forschungsgemeinschaft Grant Re 337/10-2, Re 337/12-1 to I.R. and Ju 230/5-1 to M.J.

Abstract

The discrimination of patterns that are mirror-symmetric counterparts of each other is difficult and requires substantial training. We explored whether mirror-image discrimination during expertise acquisition is based on associative learning strategies or involves a representational shift towards configural pattern descriptions that permit to resolve symmetry relations. Subjects were trained to discriminate between sets of unfamiliar grey level patterns in two conditions, which either required the separation of mirror images or not. Both groups were subsequently tested in a 4-class category-learning task employing the same set of stimuli. The results show that subjects who successfully had learned to discriminate between mirror-symmetric counterparts were distinctly faster in the categorization task, indicating a transfer of conceptual knowledge between the two tasks. Additional computer simulations suggest that the development of such symmetry concepts involves the construction of configural, proto-holistic descriptions, in which positions of pattern parts are encoded relative to a spatial frame of reference.

Visual patterns that are mirror-symmetric counterparts of each other are notoriously difficult to distinguish. As already noted by Ernst Mach (1922), children confuse characters and syllables such as p-q and no-on during the acquisition of reading and writing. Although this well-documented difficulty in infancy (see Davidson 1935; Rudel & Teuber, 1963; Bornstein, Gross & Wolf, 1978) is mostly overcome during cognitive development, mirror images continue to be perceived as particularly similar in adulthood. A classical tool to demonstrate and to exploit this phenomenon has been Shepard & Metzler's (1971) mental rotation paradigm. Here the similarity between mirror-image pairs is assumed to enforce a mental alignment of the two target stimuli prior to their mirror-image comparison, which requires more time to complete with increasing angular separation (Shepard & Cooper, 1982).

Mental rotation has also been repeatedly proposed as a general explanatory concept for all object recognition tasks including mirror-image discrimination that involve plane and depth transformations (e.g., Jolicoeur, 1985; Hinton & Parsons, 1988; Tarr & Pinker, 1989; Murray, 1997). However, more recent findings cast some doubt on its relative importance. For example, Lawson & Jolicoeur (1999) demonstrated a distinct, non-monotonic effect of plane rotation on the time required to identify objects, in contrast to the monotonic relation that is typical for tasks involving mirror-image discrimination. This finding, and the fact that practice does reduce plane rotation costs in object naming but not in left/right orientation judgements (Jolicoeur, 1989), indicate that object identification does not necessarily imply the ability to distinguish between mirror images. In fact, neuropsychological evidence shows that mirror-image discrimination may be selectively impaired, leaving object recognition skills spared (Turnbull & McCarthy, 1996; Davidoff & Warrington, 2001, Priftis et al. 2003).

Adopting an evolutionary perspective, Gross & Bornstein (1976) argue that the tendency to confuse mirror images may reflect an adaptive mode of processing rather than a perceptual limitation. As they point out virtually all mirror images in the natural world either arise from profile views of objects that are bilaterally symmetric, or from silhouettes of arbitrary objects seen from opposite sides. In either case, mirror images refer back to the same object in three-dimensional space and should therefore evoke similar response patterns. This notion is compatible with the observation that perceptual priming is unaffected by left-right reflection (Biederman & Cooper, 1991; Lawson & Humphreys, 1998; Fiser & Biederman, 2001), and with neurophysiological evidence for neurons in the inferotemporal cortex of the monkey showing response generalisation over mirror reversal (Rollenhagen & Olson, 2000; Baylis & Driver, 2001). However, the behavioural equivalence of mirror images may depend on the level of categorization involved by a recognition task. Categorization at the basic or entry level, commonly regarded to be the default in neutral contexts (Rosch et al., 1976; Jolicoeur, Gluck & Kosslyn, 1984), may not require a representation that allows the visual system to distinguish a shape from its mirror image - for example, in order to distinguish a shoe from other articles of dress. In contrast, categorization at the subordinate level - for example, a left shoe versus a right shoe - may well do so. Our ability to learn such distinctions to the point where they become almost trivial, i.e. where they attain the status of quasi entry levels within the categorization hierarchy, renders mirror-image discrimination a characteristic feature of visual expertise (Johnson & Mervis, 1997; Tanaka & Taylor, 1991). Examples of expertise requiring mirror-image discrimination skills range from everyday problems in areas in which we are all “experts” (e.g., the distinction between the letters p and q, or b and d) to tasks in highly specialized domains, such as the recall of board positions in chess

(Gobet & Simon, 1996) or the analysis of molecules of different chirality (handedness) in stereochemistry (see e.g. Eliel, Wilen & Doyle, 2001; Yee, 2002).

If mirror-image discrimination is part of visual expertise, the question arises of how such expertise can be acquired, or more specifically, how mirror-image relations become part of the knowledge about object categories during learning. Mirror-image relations between patterns assigned to different categories may affect learning in two different ways: One possibility is that the skill to discriminate between left and right counterparts of mirror-image pairs is acquired via associative learning. This notion can be related to Gross and Bornstein's (1978) aforementioned hypothesis, according to which mirror images are confused because they are interpreted as two views of one object in three-dimensional space. As a consequence, the similarity of mirror-image pairs would arise because such images, albeit being different stimuli, tend to be linked to the same conceptual node in memory (see e.g. Kruschke, 1992). Learning to distinguish between those pairs then would imply breaking up these links and connect them to different nodes, thus allowing a differential response behaviour.

Given the elementary nature of such learning one would expect it to be possible across a wide range of organisms. In support of this prediction research in animal learning shows that many species, including rats (Kinsbourne, 1967; Noonan & Axelrod, 1991), cats (Sutherland, 1963), pigeons (Corballis & Beale, 1967; Hollard & Delius, 1982; Lohmann et al., 1988), and monkeys (Nissen & McCulloch, 1937; Brown & Ettlinger, 1983), although confusing mirror stimuli initially, can be successfully trained to discriminate between them to some extent. Despite caveats such as the dependency of training success on stimulus type and questions as to the limitations of identity-matching in animals (D'Amato, Salmon & Colombo, 1985), differences in the representations at the conceptual level (Premack, 1983) and the problem of response consistency

(Corballis & Beale, 1976), this evidence points at a generic potential to acquire mirror-image discrimination skills. Moreover, in human subjects language might assist the conceptual separation of mirror stimuli in memory. Neuropsychological reports indicate that patients with impaired mirror discrimination may still use a verbal coding strategy to discriminate e.g. a left shoe from a right shoe by explicitly assigning them different response labels (Davidoff & Warrington, 1999).

As a second possibility, the tendency to confuse mirror images could arise at the level of stimulus representation as mirror images necessarily share the same local features and therefore produce similar feature descriptions. Learning to distinguish mirror-image pairs would imply a representational shift, during which local features, or isolated pattern parts, are linked to larger entities within a configural description where the symmetry relations between the two patterns can be resolved. Representational shifts from isolated parts to more holistic formats have been proposed as one of the changes that may emerge during the development of expertise in the recognition of faces and other objects (Farah et al. 1998; Gauthier & Tarr, 2002; Gauthier et al., 2003; for a review, see Palmeri, Wong & Gauthier, 2004).

With regard to mirror-image discrimination, support for the representational-shift hypothesis can be seen in the developmental trajectory of this ability during the first decade of life (Rudel & Teuber, 1963), which parallels that of face recognition. The well-established difficulties of children under the age of 10 to recognize faces (e.g., Mooney, 1957; Carey & Diamond, 1977; Kolb, Wilson & Taylor, 1992; Campbell et al., 1999) have been attributed to a particular encoding of face stimuli that relies on the processing of isolated facial features rather than of their configuration (Carey & Diamond, 1994) although this claim has not gone uncontested (Tanaka et al. 1998). Nevertheless, such correspondences makes configural object processing appear a plau-

sible way in which mirror-image relations may be incorporated into the representation of object categories during expertise acquisition.

The two hypotheses outlined above differ in the way in which they predict generalisation, or transfer, of categorical knowledge involving mirror-image relations. If such relations are mediated by associative learning mechanisms that link stimuli with particular conceptual nodes, then there should be little or no generalization if the same patterns are paired with new labels (nodes) in a subsequent transfer task. In contrast, if learning of mirror-image relations is mediated by representational shifts at the stimulus level, then such shifts, once acquired, should easily transfer to novel tasks in which the same patterns are employed in a different categorization context.

In the present study, we tested these predictions in three category-learning experiments and by means of computer simulation. Our paradigm involved the classification of Compound Gabor stimuli, grey-level patterns that result from the superposition of two sinewave gratings, a fundamental plus its third harmonic (cf. Figure1). Using this particular type of stimulus offers a number of advantages: First, it permits to study the role of mirror-image relations in categorization at a relatively early level. On the one hand, Gabor patterns have been characterized as an elementary stimulus in visual processing (e.g., Watson, Barlow, & Robson, 1983; Rentschler & Caelli, 1990; Westheimer, 1998) whereas they are, on the other hand, perceptually complex enough to stimulate category learning (Rentschler, Jüttner, & Caelli, 1994; Jüttner & Rentschler, 1996, 2000; Jüttner, Langguth & Rentschler, 2004). Second, compound Gabor patches circumvent the problem of prior knowledge, as such stimuli are unfamiliar to naive subjects; hence the acquisition of categories composed of these patterns is completely under experimental control. Finally, and most important in the present context, such patterns permit to control and manipulate mirror-image relations between pattern categories in a systematic manner, as follows.

For the experiments, the patterns were specified within a two-dimensional, evenness-oddness Fourier feature space representing a *continuum of visual shape* where each point uniquely specifies the appearance of a pattern, and patterns with mirror-symmetric coordinates relative to the evenness axis are mirror images of each other (see Figure 1 and Method section for details). Within this continuum, a set of 12 learning patterns was defined (Fig. 1c) forming a square-like configuration of four clusters (I-IV) with three patterns each. In addition of being identical in terms of their spatial frequency composition the four pattern clusters were, owing to the symmetry of the configuration in the generating Fourier space, equivalent in terms of pattern (Fourier) energy and relative spatial phase. The only factor introducing a potential anisotropy at the perceptual level was that the patterns of the cluster pairs I - IV and II - III consisted of mirror images of each other, whereas the cluster pairs I - II and III - IV did not.

In Experiment 1, we investigated whether the existence of such mirror relations would be reflected in a learning task, where subjects were trained to assign the patterns of each of the four clusters into different categories (Fig. 2a). The resulting data also constituted the reference baseline for the subsequent experiments, which addressed the explicit learning of mirror-image relations and their generalisation to a different categorization context. For Experiment 2, we combined the four clusters in a pairwise manner in two different ways that either grouped clusters containing mirror images of each other into the same category (Fig. 2b right) or into different ones (Fig. 2b left). Two further groups of observers were trained in either of these two-class conditions. The same subjects were then tested as to the transfer of their conceptual knowledge in Experiment 3, employing the same 4-class categorization task as in Experiment 1.

To corroborate our findings we also modelled the behavioural data in terms of an evidence-based classification model (Jüttner, Caelli & Rentschler, 1997; Jüttner et al., 2004). Evi-

dence-based classifiers solve a given categorization problem by constructing rules that carry evidence weights for each class alternative. These rules are based on non-relational and relational attributes of object parts defined the image domain. The set of attributes evaluated for rule generation therefore can be regarded as the 'signature' of the underlying conceptual representation. We used the simulations to identify those attributes that were crucial for mirror-image discrimination, and to explore to what extent representations established in Experiment 2 would be compatible with those acquired in Experiment 3, thus providing further support for our behavioural results.

GENERAL METHOD

Apparatus and Materials

The experiments involved the classification of Compound Gabor patterns (Fig. 1a). Such grey-level patterns result from the superposition of two sinewave gratings, a fundamental plus its third harmonic, modulated by a Gaussian aperture. Their intensity profile G(x,y) was defined by

$$G(x,y) = L_0 + \exp\left\{-\frac{1}{\alpha^2}(x^2+y^2)\right\}\left(a\cos(2\pi f_0) + b\cos(2\pi 3 f_0 x + \phi)\right) \quad (1)$$

where L_0 determines the mean luminance, α the space constant of the Gaussian aperture, a the amplitude of the fundamental, b that of the third harmonic, and ϕ the phase angle of the latter. The patterns were generated in a 128 x 128 8-bit pixel format with a linear greylevel-to-luminance function. The aperture parameter α was set to 32 pixels.

Insert Figure 1 here

Pattern variation was restricted to *b* and ϕ. This allowed the use of two-dimensional Fourier feature space with the Cartesian coordinates $\xi = b\cos\phi$ and $\eta = b\sin\phi$. The (ξ,η) feature space provided continuum of visual shape with each point uniquely specifying the appearance of a pattern. Within this feature space, patterns located symmetrically to the ξ - (or 'evenness') axis possess mirror-symmetric luminance profiles (Fig. 1b) whereas patterns located symmetrically to the η- (or 'oddness') axis possess inverted contrast of the third harmonic (see Rentschler et al., 1996). Reflecting a given feature vector (ξ_0, η_0) successively at the ξ- and η-axis (cf. Fig. 1b) leads to a "quadrupole" of patterns with the coordinates (ξ_0,- η_0), (-ξ_0,- η_0) and (-ξ_0, η_0) that are pairwise mirror images of each other but have identical image energy (Fourier power) owing to their equidistance from the origin.

Using this construction principle a learning set of 12 patterns was generated, consisting of four clusters I-IV of 3 patterns each (Fig. 1c). The four cluster means formed a square-like configuration that was centred on the origin of the Fourier feature space. Individual clusters (Experiment 1 and 3) or cluster pairs (Experiment 2) were used to define pattern categories to be learned by the subjects (Fig. 2).

The patterns were displayed on a raster monitor (Barco TVM 3/3.2, P4 phosphor) linked to a digital image processing system (Videograph, LSI 11/73). Space average luminance was kept constant at 60 cd/m^2. The patterns subtended 1.6° at a viewing distance of 165 cm. The spatial frequency of the fundamental was 2.5 c/deg.

Insert Figure 2 here

Subjects

In total, 12 paid observers (Experiment 1: 4 subjects; Experiments 2 and 3: 8 subjects). Their ages ranged between 20 and 30 years, 6 were female, 6 were male. All had normal or corrected-to-normal vision. None of the subjects had any prior experience with psychophysical experiments. They gave their written informed consent to the study after the procedure had been explained to them.

Procedure

Subjects were trained to assign the patterns to their predefined categories using a supervised-learning schedule (Rentschler et al., 1994). The procedure consisted of a variable number of learning units, each of them having two phases, training and recognition test. During the training phase, each pattern was shown three times in random order for 200 ms, followed by a number specifying the class to which the pattern was assigned. The class label was displayed for 1000 ms, with an interstimulus interval of 500 ms relative to the offset of the learning pattern. The test phase of each learning unit served to monitor the learning status of the observer. Here, the patterns were shown once in random order and classified by the subject by pressing the appropriate button on the computer keyboard. No feedback on the correctness of the individual response was provided. However, upon completion of the test phase participants were presented with the over-

all score of correct classifications obtained in the recognition test, which concluded the learning unit. The series of learning units continued until the observer reached the learning criterion of an error-free classification (100% correct) in one recognition test.

Data analysis

Observer performance was assessed in terms of learning duration (i.e., the number of learning units required to reach the learning criterion), and in terms of the mean confusion-error matrix, i.e. the relative frequencies of a correct response averaged across the learning procedure. To visualize the internal class concepts acquired during learning the confusion-error data was analysed in terms of a probabilistic virtual prototype (PVP) model (Rentschler et al., 1994; Jüttner & Rentschler, 1996, 2000). The model is based on the concept of pattern similarity and provides a technique for reconstructing the internal representation of categories that observers develop during learning. This analysis permits inferences about structure and dimensionality of the underlying conceptual space and has been shown to provide, for tasks involving the perceptual classification of Gabor patterns, a more parsimonious description than multidimensional scaling techniques (Unzicker, Jüttner & Rentschler, 1998). Internal representations of pattern classes are modelled as distributions of feature vectors around a mean vector, the so-called virtual class prototype. Human classification behaviour is formally described in terms of a Bayesian classifier that operates on such internal class representations. According to the PVP concept the distance of two virtual prototypes reflects the perceived similarity of the corresponding classes. Error vectors, which relate physical mean vectors of pattern classes to virtual prototypes, are varied to allow a least-square fit between observed and model-predicted classification frequencies.

RESULTS

Experiment 1

The aim of Experiment 1 was to establish to what extent the existence of mirror-image relations between patterns assigned to different categories would affect the acquisition of categorical knowledge. For this purpose, four naive observers (H.E., male; J.I., male; R.U., female; S.I., male) were trained, using the paradigm of supervised learning, to assign the twelve patterns of the learning set to four classes as defined by the four patterns clusters in the generating Fourier feature space (Figure 2A; configuration C0).

Insert Figure 3 here

Figure 3A (left) shows for each subject the relative classification frequencies for each class, cumulated across the entire sequence of learning units. The individual learning time of the observer is indicated by N, which specifies the number of learning units necessary to reach the learning criterion of 100% correct. Individual learning times ranged from 16 to 49 learning units with a mean of 34.2 . Despite the fact that all subjects successfully completed the task the slow learning progress suggests a particular difficulty that can be related to the pattern of confusion errors evident in the classification frequencies. The latter reveals that confusions mainly occur between categories containing mirror images. For example, subject H.E. confounds patterns of

class 1 mainly with those of class 4 but rarely with those of class 2 or class 3. An analogous error pattern is observed, with some individual variation, for the other three observers, and also with regard to the other category pairs involving mirror images, i.e. class 2 vs. class 3, class 3 vs. class 2 and class 4 vs. class 1. To consolidate this observation statistically we performed for each class pairwise comparisons of errors involving the mirror class and the non-mirror class alternative with the same distance in physical feature space: For example, for class 1 the mirror-class alternative would be given by class 4 whereas the non-mirror class alternative would be class 2. In paired t-tests these comparisons proved highly significant across all four classes ($ts(11)>5.77$; $ps<0.001$). The dominance of mirror class over non-mirror class errors is also evident from the group average data shown in Figure 3B (right).

The pronounced asymmetry induced by mirror image relations leads to a distinct distortion of the conceptual space developed during category learning (individual data Fig. 3a right; group average data Fig. 3b right). The symmetry of the square-like class configuration in the generating Fourier feature space (dashed lines) appears broken and distorted towards an elongated, almost one-dimensional arrangement that predominantly extends in parallel to the ξ-axis. Only subject R.U. succeeds to partially separate the mirror images in classes 2 and 3.

Experiment 2

In Experiment 1 we had shown that the existence of mirror image relations between pattern categories had a strong distorting effect on the conceptual space that is developed during category learning. This effect had been established by analysing the confusion error matrices averaged across the entire learning process. In Experiment 2 we focussed on the dynamics of the learning process by addressing the impact of mirror image relations on learning speed. For this purpose,

we combined the four pattern clusters of the learning set in a pairwise manner such that clusters containing mirror images of each other either fell into different categories (condition C1, Fig. 2b right) or into the same one (condition C2, Fig. 2b left). Two further groups of observers (Group 1 and Group 2) were assigned to condition C1 and C2, respectively, and trained to criterion.

Insert Figure 4 here

Fig. 4 shows the learning curves of the two groups, in terms of the percent correct response scores obtained during the test phase of each learning unit, as a function of learning time measured in learning units. Group 1 (observers G.F., female; A.Z., male; H.G., female; G.S., male; cf. Fig. 4 top), which was required to tell apart mirror images during learning, needed an average of 26.2 (range: 14-45) learning units to reach the learning criterion. In contrast, Group 2 (observers J.S. female; O.R., male; P.L., female; G.M., female; cf. Fig. 4 bottom), which was not required to discriminate between mirror images during learning succeeded at an average of 2.75 (range: 2-5) learning units. The difference between the two groups was highly significant ($t(6)=3.92$, $p<0.01$). Despite the fact that in both conditions the two compound clusters defining the target categories had the same distance within the generating Fourier feature space, the existence of mirror image relations in condition C1 induces a dramatic reduction of learning speed by about a factor of 10 relative to C2.

Experiment 3

Experiment 2 had demonstrated that the existence of mirror image relations between patterns of different category membership severely impedes the acquisition of categorical knowledge. Behaviourally this became manifest in a sharp increase in the number of learning units necessary to reach a perfect classification of all patterns in the learning set. The question arises whether the slow learning progress reflects a process by which observers gradually memorize individual patterns by associating them with their appropriate class label, or whether it indicates a shift in the way in which patterns are perceptually interpreted, or more specifically, a shift towards a representational format where mirror-image relations are easier to disambiguate. In the latter case, subjects should be able to generalize their conceptual knowledge in a transfer task employing the same patterns in a different categorization context, whereas they should show little or no such benefit if the former hypothesis is correct. We tested this prediction by requesting the same observers as in Experiment 2 to do a further learning task that was identical to that in Experiment 1. Thus we used the data previously obtained in that experiment as a baseline to compare transfer effects of observers that had either learned to distinguish mirror images in Experiment 2 (Group 1) or not (Group 2).

Insert Figure 5 here

The results for the two groups are summarised in Figures 5 and 6, using the same format as in Figure 3 (cf. Experiment 1). For Group 1 both the cumulative classification frequencies and

the reconstructed conceptual space representations are markedly different from those of the naïve subjects in Experiments I. At the level of the individual data (Figure 5A) there is no indication of a systematic distiortion of the conceptual space due to mirror image confusions. Pairwise comparisons of confusion errors involving the mirror class and the non-mirror class alternative proved non-significant for classes 1, 2 and 4 ($ts(11)<1.49$; $ps>0.16$); only for class 3 there was a significant difference ($t(11)=2.65$, $p<0.05$) between the two error types, which can be related to the exceptional response pattern for class 3 shown by subject H.G. Thus the data suggests a substantial transfer of mirror-image discrimination skills from Experiment 2 to Experiment 3. Further support comes from the group average data (Figure 5B), which shows that the square-like class configuration in Fourier feature space is well preserved in the conceptual space reconstructed from the confusion error data.

Insert Figure 6 here

In contrast, the pattern of results obtained for Group 2, both at individual (Figure 6A) and at group average level (Figure 6B), is strikingly similar to that found for the naïve subjects in Experiment 1 (cf. Figure 3). Again, the individual patterns of confusion errors show a systematic tendency to confuse mirror images assigned to different classes. Pairwise comparisons of errors involving mirror versus non-mirror class alternatives proved significant for all classes ($ts(11)>2.7$; $ps<0.05$). As a consequence, the conceptual space representation reconstructed from the group average data shows a similar distortion towards the ξ-axis as observed in Experiment 1,

although at individual level different configurations may occasionally arise (subject J.S. ; cf. also subject R.A. in Figure 3A).

The advantage of Group 1 relative to Group 2 becomes further evident in learning duration (Figure 7 right). While Group 2 on average learned the patterns after 30.2 (range: 19-48) learning units, Group 1 required only 8.5 (range: 7-11) learning units to reach the learning criterion, leading to a data pattern complementary to that observed in Experiment 2 (Figure 7 left). Again, the difference between the two groups is highly significant ($t(6)=4.13$, $p<0.01$).

Insert Figure 7 here

Simulation Results

The results of Experiment 3 show that subjects who had successfully learned to discriminate between mirror-symmetric counterparts in Experiment 2 generalized this conceptual knowledge to a different categorization context involving the same set of stimuli. This supports the hypothesis that learning to distinguish between mirror images involves a representational shift towards a format in which mirror-image relations are easier to resolve, thus facilitating their integration within categorical knowledge structures. To gain further insight into the nature of the underlying mental representations we modelled human performance in terms of evidence-based pattern classification.

According to the evidence-based systems (EBS) approach to pattern recognition, complex

objects are encoded in terms of parts and their relations that carry evidence weights for each class alternative. Originally developed in the area of machine learning and computer vision (Jain & Hoffman, 1988; Caelli & Dreier, 1994) we have previously demonstrated that the EBS architecture also provides the framework for a process model of human perceptual categorization and generalization (Jüttner et al., 1997, 2004).

An evidence-based classifier first segments a given pattern into its component parts. Each part is characterized by a set of part-specific, or unary, attributes (e.g. size, luminance, area), and each pair of parts is described by a set of relational, or binary, attributes (e.g., distance, angles, contrast). Thus, each part may be formally represented as a vector in a feature space spanned by the various unary attributes, and each pair of parts by a vector in a binary feature space. Within each feature space regions are defined that act as activation regions for rules. An attribute vector falling inside such a region will activate the corresponding rule otherwise the rule remains inactive. This leads to an object representation in terms of a rule activation vector – a vector, the components of which are assigned to the activation states of the individual rules. The activation of a given rule provides a certain amount of evidence for the class membership of the input object. The evidence weights associated with the individual rules and their combinations are implicitly represented within a neural network. Here each input node corresponds to a rule, each output node to a class, and there is one hidden layer. The relative activity of an output node provides a measure of the accumulated class-specific evidence. This activity may be probabilistically interpreted and related to a classification frequency.

In order to ensure internal consistency of our simulations, we constrained the system parameters according to a range that had proved optimal in previous work involving the same type of stimulus material (see Jüttner et al., 1997, 2004): The segmentation stage used a region-

analysis technique that was based on partitioning the image according to connected grey-level regions yielding 3-5 parts per image. The rule-generation stage employed K-means clustering procedure producing a set of 10-14 rules. Furthermore, the classifier was supplied with a reservoir of four unary attributes (position, luminance, aspect ration and size) and three binary attributes (distance, relative size, contrast). We then tested which attribute combinations represented potential solutions of the classification problem, i.e. the neural network could be successfully trained using the backpropagation algorithm to distinguish between the classes.

When used as a framework to describe category learning, each attribute combination represents a state within a search space of possible working hypotheses defined by the set of all possible combinations of unary and binary attributes. Learning speed is determined by the time necessary to find a solution within that search space, i.e., a set of attributes that allows to successfully discriminate between classes. Under the assumptions that the search is exhaustive and that the time needed to evaluate each working hypothesis is constant, learning speed is proportional to N_{FS}/N, where N_{FS} denotes the number of EBS solutions within the search space and N the total numbers of states within that space. In the context of the present experiments this implies the prediction that the number of learning units necessary to solve a classification problem (learning duration) should be proportional to $1/N_{FS.}$

Insert Figure 8 here

Figure 8A (left) shows the EBS-predicted learning durations for the category learning

tasks in Experiment 2. In agreement with the behavioural data fast learning (corresponding to a low number $1/N_{FS}$) is obtained for class configuration C2 (cf. Group 2 in Figure 7 left) that does not require the separation of mirror-symmetric counterparts. In contrast, slow learning (corresponding to a high number $1/N_{FS}$) is obtained for the class configuration C1 (cf. Group 1 in Figure 7 left) that involves the separation of mirror-symmetric counterparts.

Figure 8A, right, plots I_{FS}, where I_{FS} denotes the number of attribute states that allow the solution of the classification problems with two-classes *as well as* that of the classification problem with four classes (C0, cf. Figure 2). The reciprocal of the so-defined cross-task compatibility index is *low* for the classification problem implying a separation of mirror images (C1) and *high* for that without such a separation (C2). The latter results explain the complementary pattern observed for the learning duration of Group 1 and Group 2 in Experiment 3 relative to that in Experiment 2 (Figure 7 right): Most attribute combinations that were solutions in condition C1 (assigned to Group 1 in Experiment 2) but only few of those that were solutions in condition C2 (assigned to Group 2) were simultaneously solutions for the 4-class task in Experiment 3, thus considerably facilitating transfer for Group 1 relative to Group 2.

The relative frequencies of unary and binary attributes within the sets of solutions for condition C1 and C2 are shown in Figure 8B. These histograms may be regarded as signatures of the underlying categorical representations as they indicate the relative importance of the various attributes within the solutions of the classification tasks associated with the two conditions. According to the plot, the signature of C1 differs from that of C2 mainly in that the unary attribute *position* becomes predominant at the expense of the unary attribute *size* for the discrimination of mirror images. The value of the position attribute is measured for each pattern relative to the same reference system such as the fixed display aperture. This attribute therefore attains the qual-

ity of a relational attribute in that its values for a pair of parts imply the distance of its members. However, it should be noted that positional values are signed relative to the origin, whereas those of the regular binary attributes are non-negative by definition. In that sense the attribute *position* becomes crucial for mirror-image discrimination as it mediates the spatial localisation of parts and pairs of parts relative to an external (i.e. non object-centred) frame of reference.

DISCUSSION

Previous work has considered mirror-image-discrimination mainly as an *existing* skill, the presence (or absence) of which being indicative either of the status of cognitive development (Rudel & Teuber, 1963; Bornstein, Gross & Wolf, 1978), or of a particular perceptual deficit within neurological patient populations (Turnbull & McCarthy, 1996; Davidoff & Warrington, 2001, Priftis et al. 2003). The present study transcends this perspective by focussing on the *learning* of mirror-image discrimination skills in normal adult observers, and on the way in which such skills are embedded into the process of pattern categorization - the cognitive backbone of visual perception (Bruner, 1957; Rosch, 1978). We employed a paradigm, in which categories of unfamiliar patterns were defined within a low-dimensional Fourier feature space that allowed to define for each target class two equivalent (in terms of elementary stimulus properties such as spatial frequency, relative spatial phase and Fourier energy) distracter classes that either consisted of mirror images or of non-mirror images of the target class. Our main behavioural findings were twofold: First, the existence of mirror-image relations between patterns assigned to different classes led to a pronounced retardation of category learning (Experiment 2). Second, once observers had learned

to distinguish between mirror images in one categorization task they could transfer this skill to a different task where the same stimuli were used in a different categorization context (Experiment 3), supporting the idea of a representational shift at stimulus level that facilitates the acquisition of mirror-image relations during category learning.

Concerning the former of these results, the difficulty to learn the classification of patterns that are mirror-symmetric counterparts of each other stands in marked contrast to the efficiency and speed of detecting bilaterally symmetric shapes (see Wagemans, 1996). Owing to this efficiency bilateral symmetry relations have been repeatedly assigned to the class of stimulus properties that can be detected pre-attentively (e.g., Barlow and Reeves, 1979; Wolfe and Friedman-Hill, 1992; Locher and Wagemans, 1993). To explain the detection of such relations it has been assumed that a potential axis of symmetry is selected within a target pattern that permits a more detailed evaluation by way of a piecewise comparison of the corresponding pattern halves (Palmer and Hemenway, 1978). Alternatively, a bootstrapping mechanism has been proposed that allows the propagation of local pairwise groupings along a unique direction within a coherent global structure (Wagemans, Van Gool, Swinnen & Van Horebeek, 1993). These concepts have in common that establishing a perceptual frame of reference *within* the target pattern is essential for the detection of bilateral symmetry. The fact that the latter task can be accomplished very rapidly indicates that such a frame of reference is obtained via the *quasi* parallel filtering and grouping operations of early visual processing (e.g., Locher and Wagemans, 1993).

In contrast, the computational analysis of our learning data indicates that mirror-image discimination requires a structural representational format where pattern components are encoded relative to a spatial frame of reference that is *external* to the target pattern. In principle, such a reference frame can be specified either in observer-centered (egocentric) or scene-based

(allocentric) coordinates. Whereas our simulations do not permit us to distinguish between these two possibilities, independent evidence favours the later alternative. By analysing physical rotation patterns during comparisions of shape (including mirror images) at different orientations Hinton & Parsons (1988) showed that observers rely more on a scene-based rather than a viewer-centered representation. In neuropsychology, the occurrence of agnosias for mirror stimuli has been related to a failure to assign a suitable frame of reference to observed objects (Turnbull et al. 1997), and demonstrated for a case of specific impairment in spatial tasks relying on allocentric coordinates (Priftis et al. 2003). In conjunction with our results this suggests that category representations involving the separation of mirror-images involve positional relations relative to a scene-based frame of reference.

Our EBS simulations also provide answers to the questions of *why* the learning of such representations is so slow, and *why* their transfer to a different categorization context is relatively easy – our second major finding in the present study. Within the context of EBS, pattern category learning can be described as a successive testing of working hypotheses that in case of the categorization of compound Gabor stimuli has received support through the psychometric analysis of the confusion error data during the learning process (Jüttner & Rentschler, 1996; Unzicker et al., 1999). Formally, each working hypothesis corresponds to the selection of a subset of attributes that define a reference system for describing pattern parts and their relations. Once chosen, the elaboration of such a working hypothesis will include the formation of rules and the tuning of evidence weights. Eventually, the elaboration process either results in a successful categorization, or the current working hypothesis is rejected and replaced by a different one. The simulations show that given a certain reservoir of non-relational and relational attributes, subsets of attributes that allow the separation of mirror classes (condition C1 in Experiment 2) are relatively sparse,

whereas those allowing the discrimination of non-mirror classes (condition C2) are much more numerous. As a consequence, the search for a solution within the search space of all possible attribute combinations will be much more time-consuming in the former case relative to the latter – in agreement with the behavioural data. Furthermore, the simulations reveal that working hypotheses (subsets of attributes) that prove successful in condition C1, but not those that are successful in C2, tend to constitute simultaneous solutions of condition C0 (Experiment 3) once they have been "elaborated" (in terms of a re-adjustment of rules and evidence weights) accordingly. Without further assumptions the model predicts for Experiment 3 a pattern of learning times that is complementary to that of Experiment 2, and thus provides a parsimonious description of the data sets of the two experiments.

Evidence-based classifiers provide an explicit link between physical and internal representation, as image segmentation, attribute extraction and rule generation are entirely defined within the image domain. Furthermore, structural information is preserved by describing patterns in terms of component parts and their unary (part-specific) and binary (part-relational) attributes. This representational format contrasts with that of established psychometric approaches to categorization such as the Generalized Context Model (Nosofsky, 1986, 1991) or General Recognition Theory (Ashby, 1989; Ashby & Maddox, 1993). These models generally represent objects or patterns as single points within a multidimensional psychological space, the metric of which is determined by perceived similarity via multidimensional scaling (MDS). However, similarity-based approaches necessarily fall short to provide any explanation of the difficulty of mirror-image discrimination because they remain tacit as to *what* makes mirror stimuli look so similar. Despite recent successful attempts to apply psychometric classification techniques directly to physical parameter space rather than similarity space (Op de Beeck, Wagemans & Vogels, 2001;

Peters, Gabbiani & Koch, 2003) the difficulties in relating perceived similarity to physical stimulus properties have been abundant, and historically were part of the motivation for the development of MDS (Shepard, 1987). Specifically, for compound gratings as used in the present study MDS dimensions only partly correlate with the dimensions of the generating Fourier feature space (Kahana & Bennett, 1994), and confusion errors are predicted neither by pixel-wise pattern correlation (Jüttner et al., 1997) nor by image representations based on Laplacian pyramids or 2D curvature (Rentschler et al., 1996). The EBS approach circumvents these problems by relating classification behaviour to a representation based on the components that constitute perceptual pattern structure (see also Jüttner, 2005).

While in this respect evidence-based classification adopts a more low-level perspective than traditional psychometric categorization models it assumes a more high-level stance than physiologically inspired approaches, such as the HMAX model of Riesenhuber & Poggio (1999). HMAX operates directly in image space and consists of alternating layers of linear (S) and non-linear (C) units that perform a hierarchical decomposition of the input image into features defined by the S units. Crucially, C units employ a nonlinear maximum operation to pool over afferents tuned to different positions and scales thus achieving invariance to translation and size. Such a decomposition could be conceived as a pre-processing front end to an evidence-based classifier to detect the presence of pattern components. However, the spatial pooling performed by the C units makes the model *per se* less adequate to explain the discrimination of mirror-patterns that differ in the position of their local features. In fact, HMAX simulations yield similar confusion patterns for pseudo-mirror views of depth-rotated paper clip objects as observed for neurons in the inferotemporal cortex of the monkey (Riesenhuber & Poggio, 1999).

The evidence-based classification approach underlying our computer simulations has been validated in a number of previous studies of classification learning employing the same type of stimulus material (see Jüttner et al., 1997, 2004). Formally, this approach belongs to the class of so-called part-based recognition systems originally developed in machine vision for the recognition of complex objects in complex scenes (see Ballard & Brown, 1982; Caelli & Bischof, 1997). In general, evidence-based classifiers will produce only "attribute-indexed" representations, i.e., they ignore the explicit associations between attributes and pattern parts. Such representations would be sufficient for the distinction of classes involving patterns that are not mirror-symmetric counterparts of each other but they necessarily fail to separate classes involving mirror images, which are characterized by the same sets of unary and binary attributes. To achieve the latter task, attributes need be associated with the parts to which they refer. In our simulations this association is re-established by the use of the attribute *position*, which uniquely indexes parts by their spatial coordinates. The resulting representations attain the additional quality of being "part-indexed" and allow for more powerful but computationally more expensive processing strategies (such as graph-matching, see Bunke, 2000) capable to discriminate objects that are mirror-symmetric counterparts. The emerging prevalence of the position attribute in tasks involving mirror-image discrimination therefore indicates a *qualitative* difference with regard to the underlying representations of pattern categories.

From a phenomenological perspective, part-indexed representations can be regarded as one possible realization of a "holistic" format, in which pattern parts become connected to each other in a unique, non-interchangeable way – in contrast to attribute-indexed representations where this uniqueness is not guaranteed. Learning object representations that are capable of resolving mirror-image relations therefore suggests a shift towards a *proto-holistic* format in which

individual parts form larger constituents, or fragments, within patterns. Such fragments have been shown to be sufficient to support categorization at an intermediate level (Ullman et al., 2002) but could be part of a hierarchy of representations of increasing complexity to support also judgements at expert level (Palmeri et al., 2004). In that sense the learning of mirror-image discrimination skills might call upon mechanisms similar to those that have been proposed for the acquisition of perceptual expertise in the recognition of faces (Farah et al., 1998) and other objects (Gauthier et al., 2003).

REFERENCES

Ashby, F. G. (1989). Stochastic general recognition theory. In: D. Vickers & P. L. Smith (Eds.) *Human Information Processing: Measures, Mechanisms, and Models* (pp. 435-457). Amsterdam: Elsevier.

Ashby, F. G. & Maddox, W. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37,* 372-400.

Ballard, D. H., & Brown, C. M. (1982). *Computer Vision*. Englewood Cliffs, NJ: Prentice Hall.

Barlow, H. B., & Reeves, B. C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Research, 19,* 783 – 793.

Baylis, G. C., & Driver, J. (2001). Shape-coding in IT cells generalises over contrast and mirror reversal, but not figure-ground reversal. *Nature Neuroscience, 4,* 937-942.

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

Bornstein, M. H., Gross, Ch. G., & Wolf, J. Z. (1978). Perceptual similarity of mirror images in

infancy. *Cognition, 6,* 89 – 116.

Bruner, J. (1957). On perceptual readiness. *Psychological Review, 64,* 123-152.

Bunke, H. (2000). Graph matching for visual object recognition. *Spatial Vision, 13,* 335-340.

Caelli, T., & Bischof, W. F. (1997). *Machine learning and image interpretation.* New York: Plenum Press.

Caelli, T., & Dreier, A. (1994). Variations on the evidence-based object recognition theme. *Pattern Recognition, 27,* 1231-1248.

Campbell, R., Walker, J., Benson, P.J., Wallace, S., Coleman, M., Michelotti, J., & Baron-Cohen, S. (1999). When does the inner-face advantage in familiar face recognition arise and why? *Visual Cognition, 6,* 197-216.

Corballis, M. C. & Beale, I. L. (1967). Interocular transfer following simultaneous discrimination of mirror-image stimuli. *Psychonomic Science, 9,* 605-606.

Corballis, M., & Beale, I. L. (1976). *The psychology of left and right.* Hillsdale, NJ: Erlbaum.

Brown, J. V., & Ettlinger, G. (1983). Intermanual transfer of mirror-image discrimination by monkeys. *Quarterly Journal of Experimental Psychology, 35B,* 119-124.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195,* 312-314.

Carey, S., & Diamond R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition,* 1, 253-274.

D'Amato, M. R., Salmon, D. P., & Colombo, M. (1985). Extent and limits of the matching concept in monkeys (Cebus apella). *Journal of Experimental Psychology: Animal Behavior Processes, 11,* 35-51.

Davidoff, J., & Warrington, E. K. (1999). The bare bones of object recognition: Implications of a

case of object recognition impairment. *Neuropsychologia, 37,* 279-292.

Davidoff, J., & Warrington, E. K. (2001). A particular difficulty in discriminating between mirror images. *Neuropsychologia, 39,* 1022-1036.

Davidson, H. P. (1935). A study of the confusing letters b, d, p, q. *Journal of Genetic Psychology, 47,* 458-468.

Eliel, E. L., Wilen, S. H., & Doyle, M. P. (2001). *Basic organic stereochemistry.* New York: John Wiley.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. W. (1998). What is 'special' about face perception? *Psychological Review, 105,* 482-498.

Fiser, J., & Biederman, I. (2001). Invariance of long-term visual priming to scale, reflection, translation, and hemisphere. *Vision Research, 41,* 221-234.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6,* 428-432.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 431-446.

Gobet, F. & Simon, H. A. (1996). Recall of Random and Distorted Positions: Implications for the theory of expertise. *Memory & Cognition, 24,* 493-503.

Gross, C.G. & Bornstein, M.H. (1978). Left and right in science and art. *Leonardo, 11,* 29-38.

Hinton, G. E., & Parsons, L. M. (1988). Scene-based and viewer-centered representations for comparing shapes. *Cognition, 30,* 1-35.

Hollard, V. D., & Delius, J. D. (1982). Rotational invariance in visual pattern recognition by pigeons and humans. *Science, 218,* 804-806.

Jain, A.K., & Hoffman, R. (1988). Evidence-based recognition of 3-D objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 10,* 783-802.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126,* 248-277.

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

Jolicoeur, P. (1989). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology, 42,* 461-478.

Jolicoeur, P., Gluck, M., & Kosslyn, S. M. (1984). Pictures and names: making the connection. *Cognitive Psychology, 16,* 243-275.

Jüttner, M., & Rentschler, I. (1996). Reduced perceptual dimensionality in extrafoveal vision. *Vision Research 36,* 1007-1022.

Jüttner, M., & Rentschler, I. (2000). Scale invariant superiority of foveal vision in perceptual categorization. *European Journal of Neuroscience, 12,* 353-359.

Jüttner, M., Caelli, T., & Rentschler, I. (1997). Evidence-based pattern recognition: a structural approach to human perceptual learning and generalization. *Journal of Mathematical Psychology, 41,* 244-258.

Jüttner, M., Langguth, B., & Rentschler, I. (2004). The impact of context on pattern category learning and representation. *Visual Cognition, 11,* 921-945.

Jüttner, M. (2005) Part-based strategies for visual categorisation and object recognition. In: N. Osaka, I. Rentschler & I. Biederman (Eds.) *Object Recognition, Attention & Action* (in press). Springer, Tokyo.

Kahana, M. J., & Bennett, P. J. (1994) Classification and perceived similarity of compound grat-

ings that differ in relative spatial phase. *Perception & Psychophysics 55,* 642-656.

Kinsbourne, M. (1967). Sameness-difference judgements and the discrimination of obliques in the rat. *Psychonomic Science, 7,* 183-184.

Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain and Cognition, 20,* 74-84.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22-44.

Lawson R., & Humphreys, G. W. (1998). View-specific effects of depth rotation and foreshortening on the initial recognition and priming of familiar objects. *Perception & Psychophysics, 60,* 1052-1066.

Lawson, R., & Jolicoeur, P. (1999) The effects of prior experience on recognition thresholds for plane-disoriented pictures of familiar objects. *Memory & Cognition, 27,* 751-758.

Locher, P., & Wagemans, J. (1993). The effects of element type and spatial grouping on symmetry detection. *Perception, 22,* 565 – 587.

Lohmann, A., Delius, J. D., Hollard, V. D., & Friesel, M. F. (1988). Discrimination of shape reflections and shape orientations by *Columba livia*. *Journal of Comparative Psychology, 102,* 3 – 13.

Mach, E. (1922). *Die Analyse der Empfindungen* (9th ed.). Jena: G. Fischer.

Mooney, C. M. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology,* 11, 219-226.

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

Nissen, H. W., & McCulloch, T. L. (1937). Equated and non equated stimulus situations in discrimination learning by chimpanzees. I. Comparisons with unlimited response. *Journal of Comparative Psychology, 23,* 165-189.

Noonan, M., & Axelrod, S. (1991). Acquisition of left-right response differentiation in the rat following section of the corpus callosum. *Behavioural Brain Research, 46,* 135-143.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General 115,* 39-57.

Nosofsky, R. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 3-27.

Op de Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterised shapes. *Nature Neuroscience, 4,* 1244-1252.

Palmer, S. E., & Hemenway, K. (1978). Orientation and symmetry: Effects of multiple, rotational, and near symmetries. *Journal Experimental Psychology: Human Perception and Performance, 15,* 635 – 649.

Palmeri, T. J., Wong, A. C-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8,* 378-386.

Peters, R. J., Gabbiani, F., & Koch, C. (2003) Human visual object categorization can be described by models with low memory capacity. *Vision Research, 43,* 2265-2280.

Premack, D. (1983). The codes of man and beast. *Behavioral and Brain Sciences, 6,* 125-167.

Priftis, K., Rusconi, E., Umilta, C., & Zorzi, M. (2003). Pure agnosia for mirror stimuli after right inferior parietal lesion. *Brain, 126,* 908-919.

Rentschler, I., & Caelli, T. (1990). Visual representations in the brain: Inferences from psycho-

physical research. In: H. Haken & K. Stadler (Eds.) *Synergetics of Cognition,* Springer Series in Synergetics, Vol. 45 (pp. 233-248). Berlin: Springer.

Rentschler, I., Jüttner, M., & Caelli, T. (1994). Probabilistic analysis of human supervised learning and classification. *Vision Research, 34,* 669-687.

Rentschler, I., Barth, E., Caelli, T., Zetzsche, C. & Jüttner, M. (1996). On the generalization of symmetry relations in visual pattern classification. In: C.W. Tyler (Ed.) *Human Symmetry Perception* (pp. 237-264). Utrecht: VSP.

Riesenhuber, M., & Poggio, T. (1999) Hierarchical models of object recognition in the cortex. *Nature Neurosciences, 2,* 1019-1025.

Rollenhagen, J. E., & Olson, C. R. (2000). Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science, 287,* 1506-1508.

Rosch, E. (1978). Principles of categorization. In: E. Rosch & B. Lloyd (Eds.) *Cognition and Categorization* (pp. 27-48). Hillsdale, NJ: Erlbaum.

Rudel, R. G., & Teuber, H. L. (1963). Discrimination of direction of line in children. *Journal of Comparative and Pysiological Psychology, 56,* 892-898.

Rosch, E., Mervis, C., Gray, W., Johnson, D., & Boyes-Braem, P. (1976) Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science, 237,* 1317-1323.

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

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

Sutherland, N. S. (1963). Cat's ability to discriminate oblique rectangles. *Science, 139,* 209-210.

Tanaka, J.W., Kay, J.B., Grinnell, E., Stansfield, B., & Szechter,L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition, 5,* 479-496.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457-482.

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

Turnbull, O. H., & McCarthy, R. A. (1996). Failure to discriminate between mirror-image objects: A case of viewpoint-independent object recognition? *Neurocase, 2,* 63-71.

Turnbull, O. H., Beschin, N., & Della Sala, S. (1997). Agnosia for object orientation: implications for theories of object recognition. *Neuropsychologia, 35,* 153-163.

Ullman, S., Vidal-Naquet, M., & Sali, E. (2002). Visual features of intermediate complexity and their use in classification. *Nature Neuroscience, 5,* 682-687.

Unzicker, A., Jüttner, M., & Rentschler, I. (1998). Similarity-based models of human visual recognition. *Vision Research, 38,* 2289-2305.

Unzicker, A., Jüttner, M., & Rentschler, I. (1999). Modelling the dynamics of visual classification behaviour. *Mathematical Social Sciences,* 38, 295-313.

Wagemans, J. (1996). Detection of visual symmetries. In: C.W. Tyler (Ed.) *Human symmetry perception* (pp. 237-264). Utrecht: VSP.

Wagemans, J., Van Gool, L., Swinnen, V., & Van Horebeek, J. (1993). Higher-order structure in regularity detection. *Vision Research, 33,* 1067 – 1088.

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature, 302,*

419-422.

Westheimer, G. (1998). Lines and Gabor functions compared as spatial visual stimuli. *Vision Research, 38,* 487-491.

Wolfe, J. M., & Friedman-Hill, S. R. (1992). On the role of symmetry in visual search. *Psychological Science, 3,* 194-198.

Yee, G. T. (2002). Through the looking glass and what Alice ate there. *Journal of Chemical Education,* 79, 569–571.

FIGURE CAPTIONS

Figure 1. Greylevel representation **(a)** and corresponding luminance profiles **(b)** of iso-energy compound Gabor patterns used for category learning. The patterns were composed of two gratings, a fundamental spatial frequency component of fixed amplitude and cosine phase with a third harmonic of variable amplitude *b* and phase φ within a Gaussian aperture. A metric feature space with the Cartesian even, $\xi = b \cos \varphi$, and odd, $\eta = b \sin \varphi$, co-ordinates was used for pattern representation. Reflecting a given feature vector (ξ_0, η_0) successively at the ξ- and η-axis leads to a "quadrupole" of patterns that are pairwise mirror images of each other but have identical image energy (Fourier power) owing to their equidistance from the origin. **(c)** Using this construction principle a learning set of 12 patterns was generated, consisting of four clusters I-IV of 3 patterns each (small symbols). The four cluster means (illustrated in A, B but not part of the learning set) formed a square-like configuration that was centred on the origin of the Fourier feature space. Scale: One unit corresponds to 20 cd/m^2 amplitude relatively to D.C.

Figure 2. (a) Individual clusters (Condition C0) or **(b)** cluster pairs (Condition C1 and C2) of the learning set shown in Fig. 1c were used to define pattern categories to be learned by the subjects. Note that in (B) clusters with mirror images of each other were either grouped into the different classes (C1) or into the same class (C2). Symbols are used to denote each cluster: cluster 1: black circles; cluster 2: black squares; cluster 3: open squares; cluster 4: open circles. Same symbol shape refers to clusters containing mirror patterns of each other. Large symbols and dotted lines indicate the cluster means and are not part of the learning set.

Figure 3 Classification of the learning set into four classes (Condition C0) by four naive observers. **(a, left)** Relative classification frequencies for each class (symbols as in Fig. 2) cumulatesd over learning units. Initials of subjects and the number *N* of learning units to 100% classification as insets. **(a, right)** Virtual prototype solution derived from the observed classification probabilities. For comparison the square-like configuration of the four mean pattern vectors of the learning set (cf. Fig. 2a) is indicated by the dotted square. The variable *e* denotes the root mean squared (RMS) error between observed and predicted classification probabilities. **(b)** Same as (a) but classification frequencies collapsed over observers.

Figure 4. Mean percent correct classification as a function of the number of learning units (learning curves) for the four subjects of Group 1 (Condition C1 in Fig. 2b) and of Group 2 (Condition C2).

Figure 5. Classification of the learning set into four classes (Condition C0) by the four observers of Group 1 who were pre-trained with two-class condition C1. Data format as in Fig. 3.

Figure 6. Classification of the learning set into four classes (Condition C0) by the four observers of Group 2 who were pre-trained with two-class condition C2. Data format as in Fig. 3.

Figure 7. Mean learning duration (number of learning units to criterion) in Experiment 2 (two-classes configurations) and in Experiment 3 (four-classes configuration) for Group 1 and Group 2.

Figure 8. Simulation results. **(a)** EBS-predicted relative learning durations for Experiment 2 (two-class conditions C1 and C2, left) and for Experiment 3 (four-class condition C0, right) with observers being pre-trained in either C1 or C2. N_{FS} denotes the number of attribute solution states within the EBS search space. I_{FS} is the number of attribute solution states that are simultaneous solutions of the classification tasks involving two classes (conditions C1 or C2) as well as that with four classes (C0). Note that the complementary learning time patterns for Experiment 2 and Experiment 3 closely match the behavioural data shown in Fig. 7. **(b)** Relative frequencies of unary (*u.P*: position, *u.S*: size, *u.I*: luminance, *u.A*: aspect ratio) and binary (*b.D*: distance, *b.S*: relative size, *b.C*: contrast) attributes within the EBS solutions for the classification tasks C1 and C2. Note that the attribute *position* attains a predominant role for condition C1, which involves the discrimination of mirror patterns.

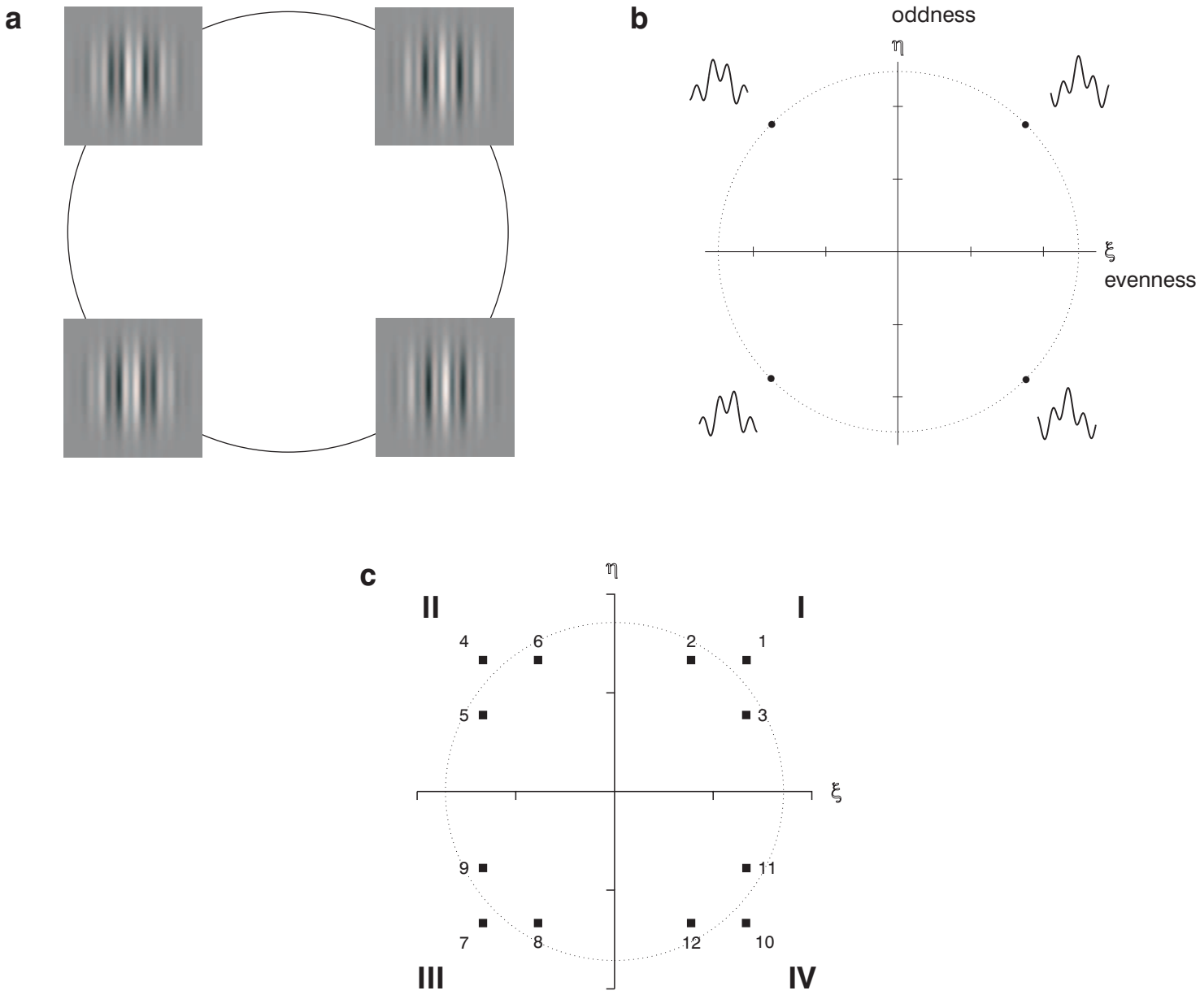

Fig. 1

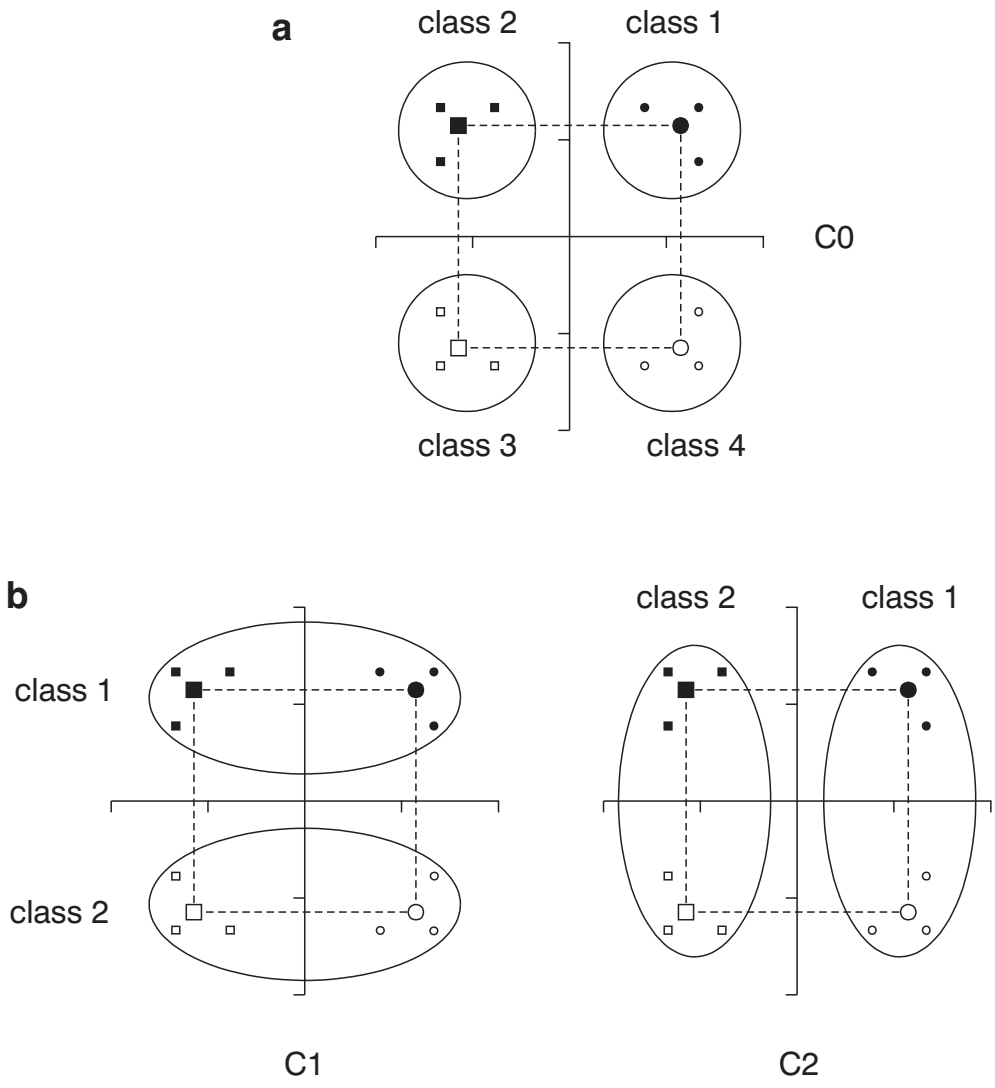
a
class 2
class 1
C0
class 3
class 4
b
class 1
class 2
C1
class 2
class 1
C2

Fig. 2

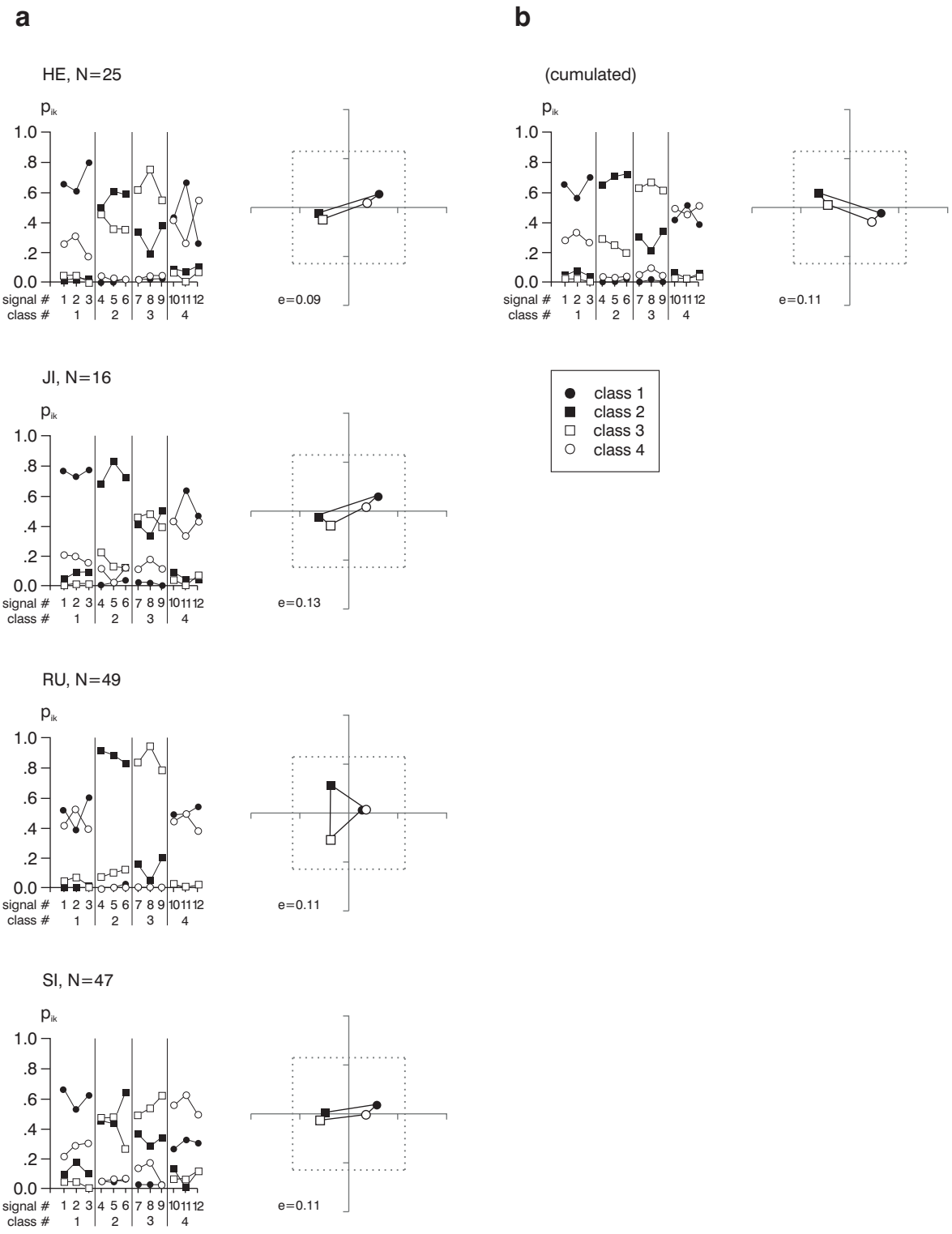
a
b
HE, N=25
(cumulated)
JI, N=16
RU, N=49
SI, N=47
p_{ik}
1.0
.8
.6
.4
.2
0.0
signal # 1 2 3 4 5 6 7 8 9 10 11 12
class # 1 2 3 4
e=0.09
e=0.11
e=0.13
e=0.11
e=0.11
class 1
class 2
class 3
class 4

Fig. 3

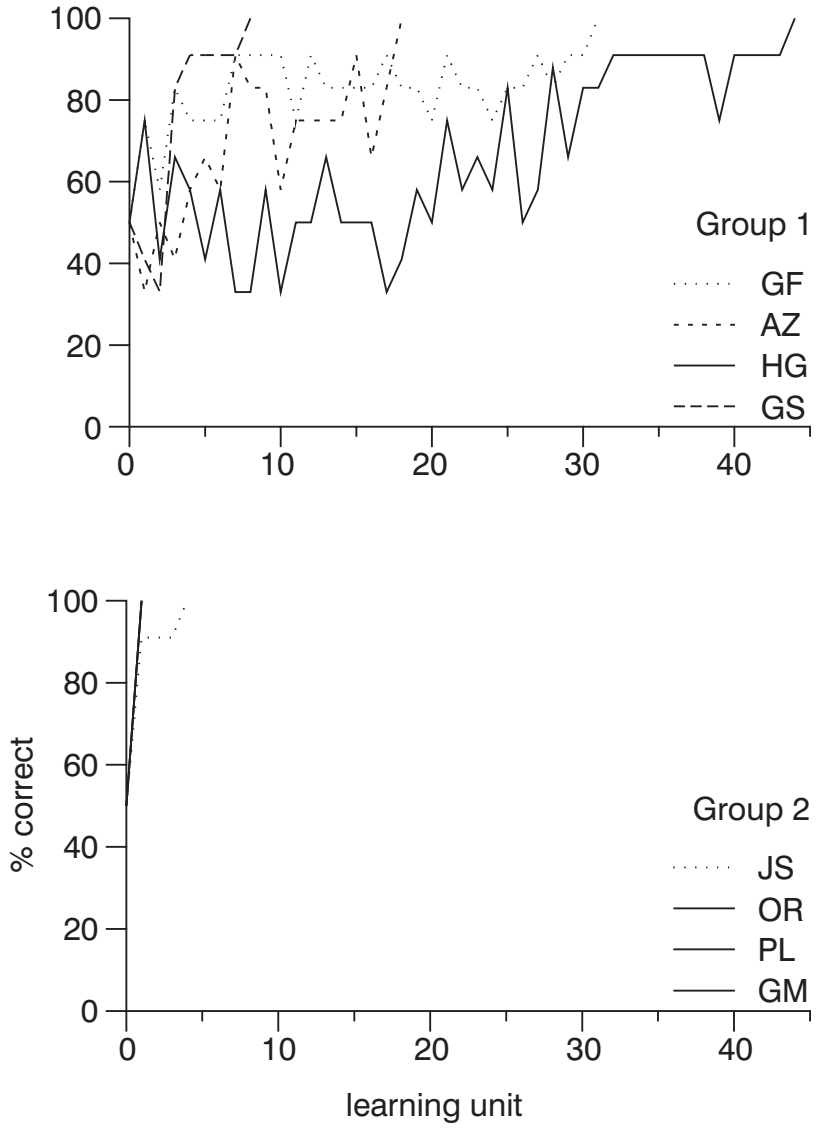
100
80
60
40
20
0
0
10
20
30
40
Group 1
GF
AZ
HG
GS
100
80
60
40
20
0
% correct
0
10
20
30
40
learning unit
Group 2
JS
OR
PL
GM

Fig. 4

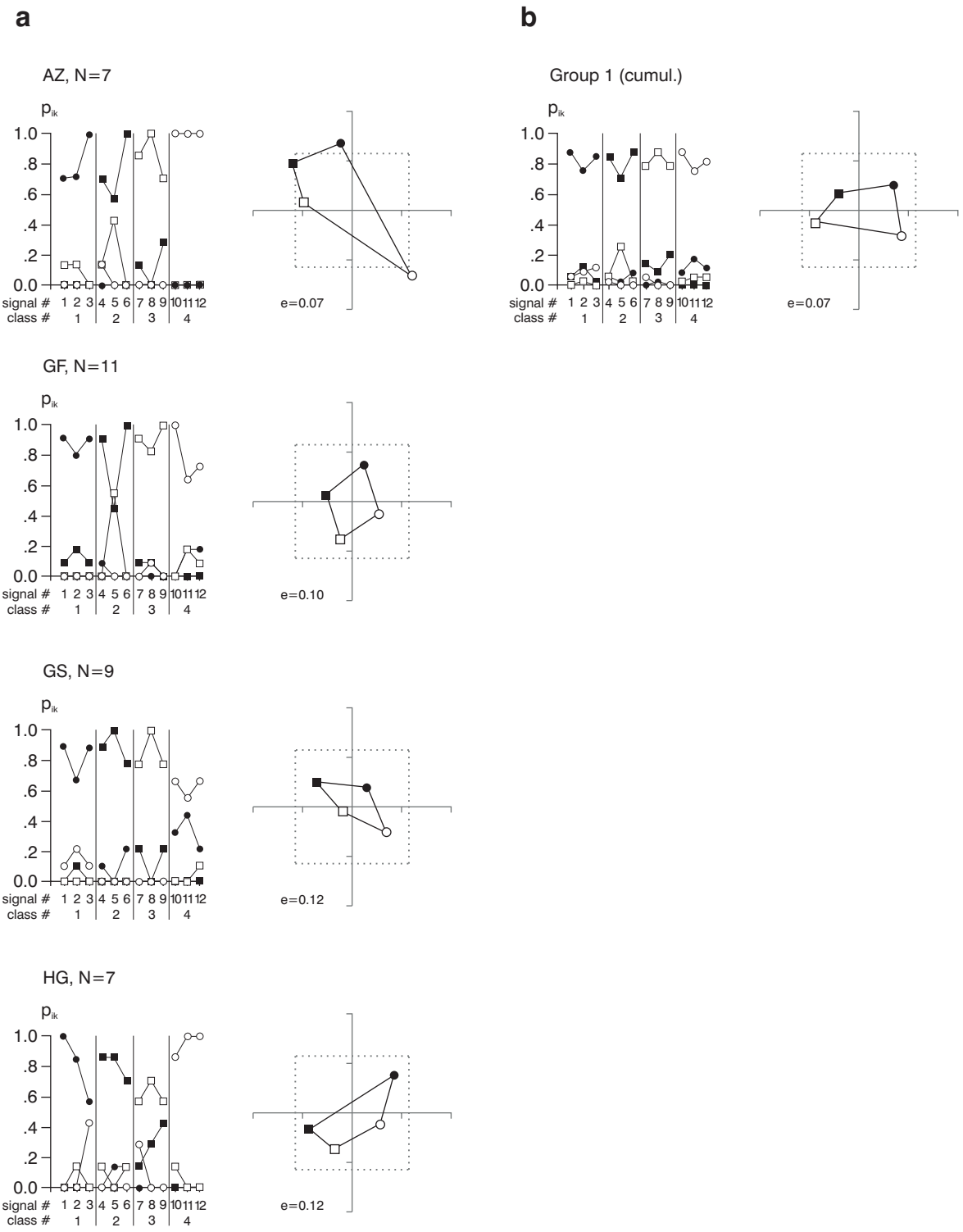
a
b
AZ, N=7
Group 1 (cumul.)
GF, N=11
GS, N=9
HG, N=7
p_{ik}
1.0
.8
.6
.4
.2
0.0
signal # 1 2 3 4 5 6 7 8 9 10 11 12
class # 1 2 3 4
e=0.07
e=0.10
e=0.12

Fig. 5

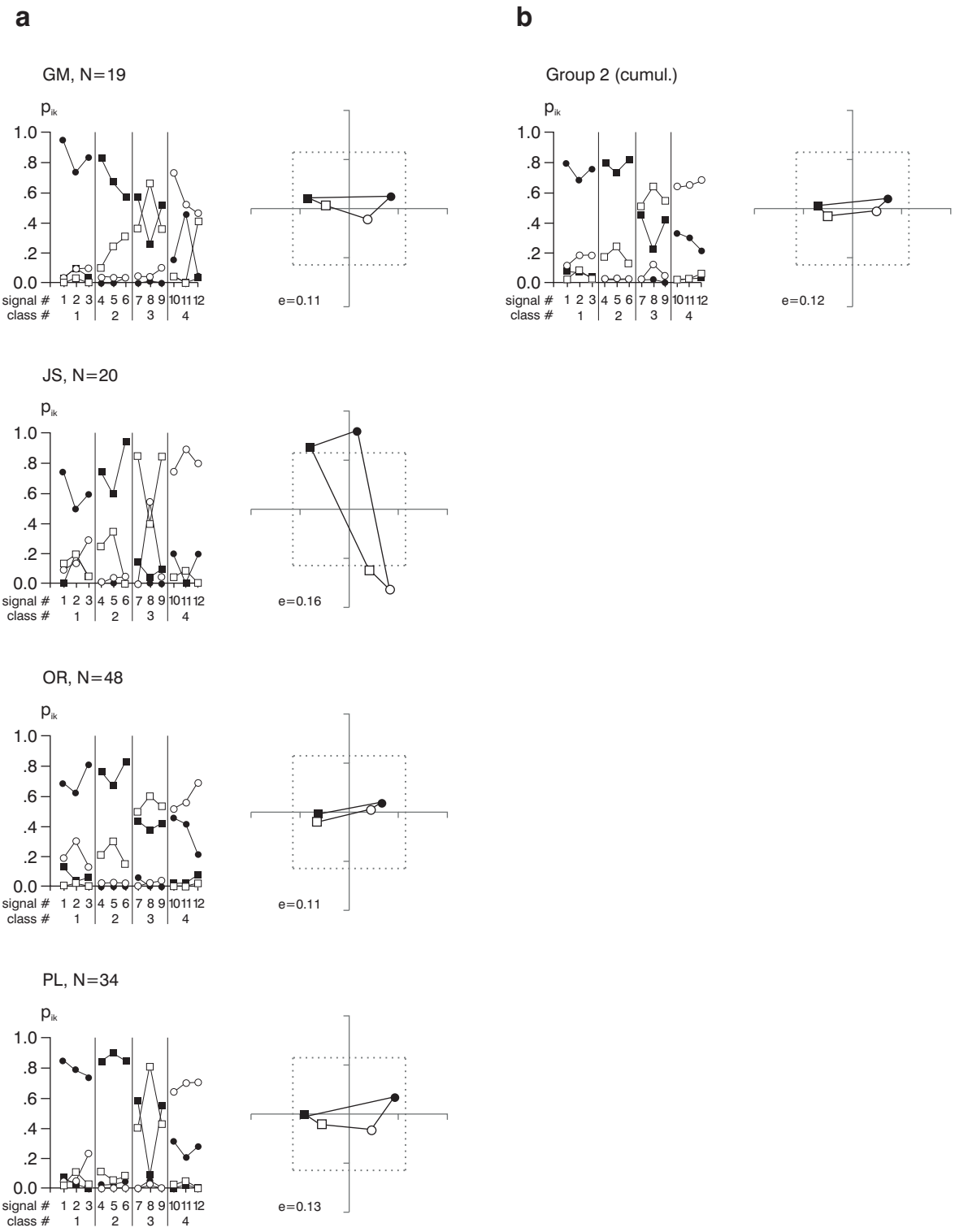
a
b
GM, N=19
Group 2 (cumul.)
JS, N=20
OR, N=48
PL, N=34
p_{ik}
1.0
.8
.6
.4
.2
0.0
signal # 1 2 3 4 5 6 7 8 9 10 11 12
class # 1 2 3 4
e=0.11
e=0.12
e=0.16
e=0.11
e=0.13

Fig. 6

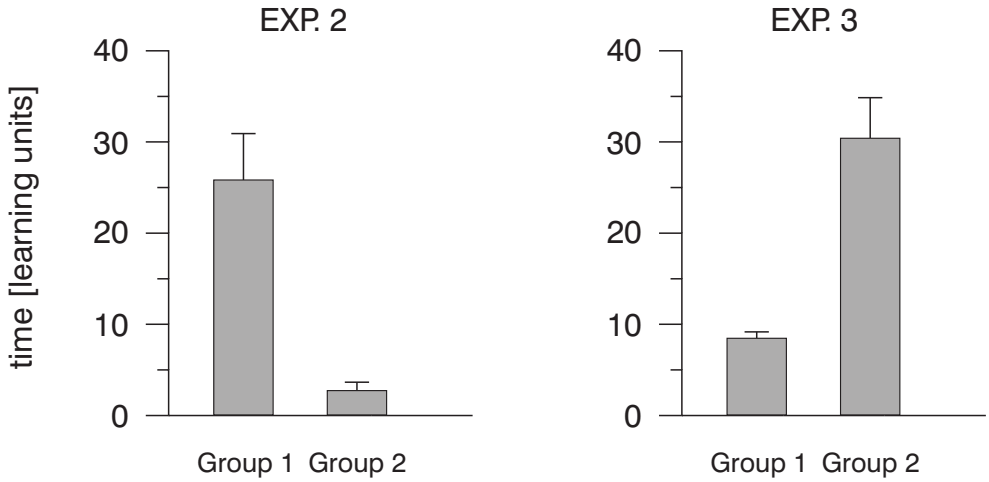
EXP. 2
EXP. 3
time [learning units]
40
30
20
10
0
Group 1
Group 2
Group 1
Group 2

Fig. 7

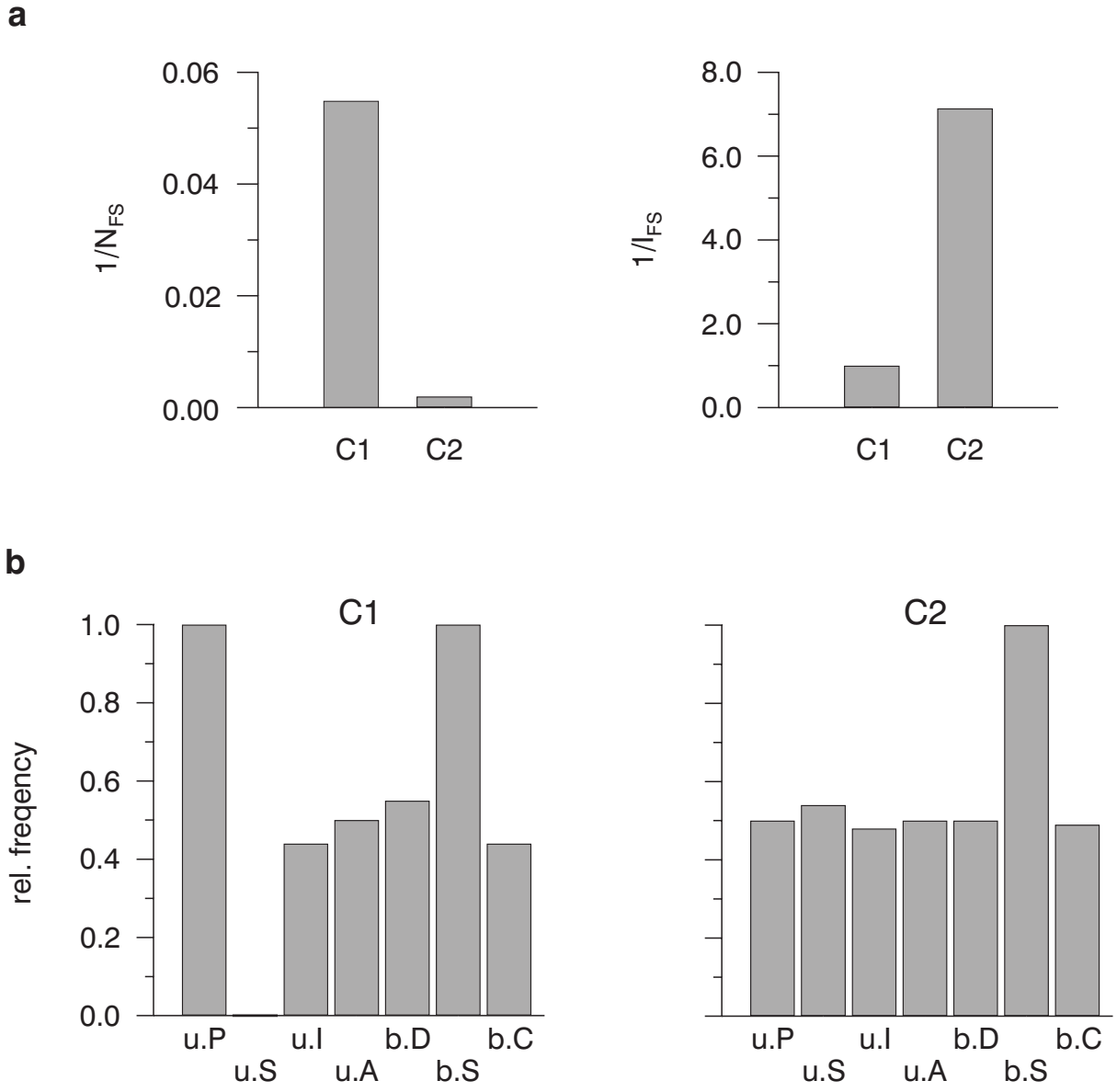
a
0.06
0.04
0.02
0.00
1/N_FS
C1
C2
8.0
6.0
4.0
2.0
0.0
1/I_FS
C1
C2
b
C1
C2
1.0
0.8
0.6
0.4
0.2
0.0
rel. freqency
u.P
u.S
u.I
u.A
b.D
b.S
b.C

Fig. 8