

Category learning induces position invariance of pattern recognition across the visual field

Martin Jüttner^{1,*} and Ingo Rentschler²

Reference for citations:

Jüttner, M. & Rentschler, I. (2008) Category learning induces position invariance of pattern recognition across the visual field. *Proceedings of the Royal Society B* 275, 403-410.

¹Psychology - School of Life & Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, UK

²Institut für Medizinische Psychologie, Universität München, Goethestrasse 31, D-80336 München, Germany

* Author for correspondence: (m.juttner@aston.ac.uk)

Keywords: position invariance; categorization; learning; pattern recognition; peripheral vision

Abstract

Human object recognition is considered to be largely invariant to translation across the visual field. However, the origin of this invariance to positional changes has remained elusive, since numerous studies found that the ability to discriminate between visual patterns develops in a largely location-specific manner, with only a limited transfer to novel visual field positions. In order to reconcile these contradicting observations we traced the acquisition of categories of unfamiliar grey-level patterns within an interleaved learning and testing paradigm that involved either the same or different retinal locations. Our results show that position invariance is an emergent property of category learning. Pattern categories acquired over several hours at a fixed location in either the peripheral or central visual field gradually become accessible at new locations without any position-specific feedback. Furthermore, categories of novel patterns presented in the left hemifield are distinctly faster learnt and better generalized to other locations than those learnt in the right hemifield. Our results suggest that during learning initially position-specific representations of categories based on spatial pattern structure become encoded in a relational, position-invariant format. Such representational shifts may provide a generic mechanism to achieve perceptual invariance in object recognition.

1. INTRODUCTION

Our ability to recognize familiar objects is surprisingly robust against displacements of such objects within the visual field (Ellis *et al.* 1989; Biederman & Cooper 1991; Stankiewicz & Hummel 2002). However, the explanation of this phenomenon of position invariance has proved difficult since psychophysical studies have found pattern-discrimination learning to be largely location-specific, with only a limited potential for transfer to novel visual field positions (Foster & Kahn 1985; Nazir & O'Regan 1990; Dill & Fahle 1997, 1998). These seemingly contradictory observations at the behavioural level are not easily reconciled by neurophysiological findings. Neurons in the inferotemporal cortex (IT), an area of crucial importance for object recognition in primates, typically show large receptive fields, a property that has previously been linked to translational response invariance (Rolls 1992). On the other hand, more recent evidence indicates that response behaviour of IT cells is affected by other factors as well and may show a far greater sensitivity to retinal position than predicted by receptive field size (DiCarlo & Maunsell 2003; Rolls *et al.* 2003). Such modulation effects and the fact that objects may be represented by the combined activity of neuronal populations (Haxby *et al.* 2001) make it difficult to predict the effect of stimulus displacement on the overall response of the visual system on the basis of receptive field properties alone.

Our approach to resolve this debate is based on the notion that object recognition properly relies on previously acquired stimulus categories (Rosch 1978), and on evidence that learning involves changes in the internal representation of categories (Schyns *et al.* 1998; Rentschler & Jüttner 2007) that may also affect the invariance properties of such representations (Jüttner *et al.* 2004). To maximize the potential impact of category learning on the development of position-invariant object recognition we took advantage of the fact that for the discrimination of unfamiliar (i.e. unlearned) structure-only stimuli, i.e. patterns that only differ in the spatial arrangement of their constituent parts rather than the shape of those parts, position invariance is broken (Dill & Edelman 2001). On this basis we designed classes of unfamiliar patterns that were defined by the spatial composition of their constituent parts. Our stimuli were Compound Gabor gratings, two-dimensional grey level patterns with a well-defined, one-dimensional part structure in terms of bright and dark bars along their horizontal symmetry axis. The use of such patterns allowed us to confine positional changes to one-dimensional displacements along the horizontal meridian across the visual field. Furthermore, Compound-Gabor gratings represent an elementary

stimulus type in early visual processing (Watson *et al.* 1983; Westheimer 1998) that is perceptually highly unfamiliar, thus stimulating learning while minimizing confounding effects of prior knowledge (Jüttner & Rentschler 1996, 2000).

We used a set of fifteen Compound Gabor patterns that were defined within a two-dimensional Fourier feature space (figure 1*a*). This feature space allowed to define stimuli in a low dimensional “form continuum”, within which each point uniquely defines the appearance of a pattern and clusters of points are used to define classes to be learned by the subject. Observers were trained to classify the patterns employing a paradigm of interleaved learning and testing that involved either the same or different retinal locations during the learning and testing phase of each learning unit (figure 2*a*). The experiments consisted of an induction stage, during which learning and testing involved the same retinal location, and two transfer stages, during which the test location was shifted to a novel position whereas the learning location remained unchanged. Using this paradigm we explored to what extent categorical pattern knowledge during learning could be transferred from the left visual field (LVF) to the right visual field (RVF), the RVF to the LVF, and from extrafoveal to foveal vision, or vice versa.

2. MATERIAL AND METHODS

(a) Subjects

Thirty paid observers (age range 19 to 32 years) participated in the study, with ten subjects (5 female, 5 male) being assigned to each of the three experiments reported. All participants were right-handed and had normal or corrected-to-normal vision. None of them had any prior experience with psychophysical experiments. All gave their informed consent prior to the study.

(b) Stimuli

A set of fifteen compound Gabor gratings, consisting of a fundamental plus its third harmonic within a Gaussian aperture, served as learning patterns (figure 1). The stimuli were generated on a computer (Research Machines PC; Matrox Millennium G450 graphics) and displayed on 17-inch monitor (EIZO F56; spatial resolution of 1024 x 768 pixel; refresh rate 75 Hz). Space average luminance was kept constant at 60 cd/m². The fundamental of the Gabor stimuli had a spatial frequency was 2.4 cycles/deg and an amplitude of 40 cd/m². The patterns subtended 1.7 deg at a viewing distance of 101 cm when seen foveally. In the 3 deg off-axis viewing conditions, the

Category learning induces position invariance of recognition

stimulus size was re-scaled to 2.7 deg according to cortical magnification (Rovamo & Virsu 1979). Eccentricity was measured between the fixation point and the centre of stimulus patterns.

(c) Interleaved learning and testing

The experiments used an interleaved learning and testing procedure (Jüttner & Rentschler 1996, 2000) that was split into a variable number of learning units. Each of them consisted of a learning (L) phase and a test (T) phase (figure 2a). During the learning phase, the patterns of the learning set were successively presented for 200 ms. Each pattern was shown three times in random order, and each presentation was followed, after an interstimulus interval of 500 ms, by a number, that was displayed for 1000 ms and specified the category of the pattern. The test phase employed the same temporal parameters for stimulus presentation. However, each pattern was shown only once and subjects had to indicate the category of each stimulus by pressing the corresponding key. No feedback on the correctness of the response was given. The series of learning units, i.e. the alternating sequence of learning and testing, continued until either the observer had achieved a criterion of 100% correct in a test phase, or had passed 40 learning units without achieving it.

Each experiment was divided into three stages, one induction and two transfer stages (figure 2b). During induction (IN) the viewing locations during the learning (L) phases and test (T) phases of each learning unit were identical; during the two transfer stages (T1 and T2), the viewing locations during learning and testing became dissociated from each other. Participants proceeded through these stages in one-hour sessions (approximately 10 learning units per session) on consecutive days. Upon completion of the induction stage or transfer stage T1 at least two learning units of the following stage were performed within the current session to include the transition between consecutive stages.

Throughout all experiments subjects had to fixate a central fixation point on the computer display. Patterns were presented on the horizontal meridian either in the left visual field (LVF; eccentricity of -3 deg relative to the fixation target), in the right visual field (RVF; eccentricity +3 deg), or centrally (eccentricity 0 deg). Viewing always was binocular. The short stimulus duration of 200 ms ensured a presentation to a consistent retinal location and discouraged saccadic eye movements. Eye movements were not monitored. Pilot experiments with controlled eye position in the context of previous studies employing similar stimuli and viewing conditions (Jüttner & Rentschler 1996) showed that for the above learning paradigm fixation errors or

Category learning induces position invariance of recognition

erroneous saccades are infrequent (less than 3% of trials) due to the highly repetitive nature of the task.

Experiment 1 contrasted two conditions, in which the subjects learnt the patterns (during the L phase of each learning unit) either in the RVF (C1) or in the LVF (C2). During the induction stage of the experiment the subjects were also tested (during the T phase) at these positions. After completion of this stage subjects entered transfer stage T1, in which the test location was moved to the mirror-symmetric location in the contralateral field, i.e. the LVF in C1 and the RVF in C2. Finally, in stage T2 the test location was moved to the central position for both groups.

Experiment 2 compared two conditions, in which subjects always learnt the patterns (during the L phase of each learning unit) in direct view. During the induction stage of this experiment, both groups were also tested at the central location. During T1 the test location in condition C1 was moved to the RVF, while it was shifted to the LVF in condition C2. Finally, during T2 the test location was moved to the mirror-symmetric position in the contralateral field, i.e. into the LVF in C1 and into the RVF in C2.

In Experiment 3 the same sequence of test locations was used as in Experiment 2. In contrast to the latter, however, shifts of the test location were always accompanied by identical shifts of the learning location. Thus learning and testing always occurred at the same position throughout the experiment.

(d) Data analysis

Observer performance was assessed in terms of learning time (i.e., the number of learning units required to reach the learning criterion), and in terms of the relative frequencies of a correct response during the three stages of each experiment.

In order to obtain robust estimates for the transfer across stages time-normalized learning curves were derived for each subject from the series of percent-correct scores obtained during the test phase of each learning unit. This was done by computing averages for each decile of learning units (or quartile, if the learning duration was less than 10 units). The first and last of these averages were used as anchor values to assess the transition between induction stage and transfer stage T1, and between transfer stage T1 and transfer stage T2. The same anchor values served to evaluate learning rate, defined by the ratio of the difference between the anchors within each

stage and the absolute number of intervening learning units.

To track the development of internal class representations during learning individual confusion-error matrices representing the average for the first and last decile (quartile) of learning units were computed. The group means of these matrices were analysed in terms of a probabilistic virtual prototype model (Jüttner & Rentschler 1996, 2000). The model provides a technique to visualize changes in the similarity structure and dimensionality of the conceptual space during learning and has been shown to yield, for tasks involving the perceptual classification of Gabor patterns, a more parsimonious description than multidimensional scaling (Unzicker *et al.* 1998). Internal representations of pattern classes are modelled as distributions of feature vectors around a mean vector, the so-called virtual class prototype, and human classification is described in terms of a Bayesian classifier operating on such representations. Distances between virtual prototypes reflect the perceived similarity between the corresponding class concepts and are varied to minimize the mean squared error between observed and model-predicted classification frequencies.

3. RESULTS

(a) *Experiment 1*

In Experiment 1 subjects saw the patterns either in the right visual field (condition C1) or in the left visual field (condition C2) during the learning phase of each learning unit (figure 2*b*, top). They were subsequently tested at the same location (during the induction stage IN), the mirror-symmetric location in the contralateral field (during transfer stage T1) and the fovea (during transfer stage T2). In each stage observers were trained to criterion.

Figure 3*a* shows the individual, time-normalized learning curves as well as the mean classification performance in each learning condition at the beginning and end of the three stages. Following the increase during the induction stage, performance significantly drops ($F(1,8)=8.17$, $P<0.05$, repeated-measurement ANOVA, mixed design, simple contrasts) as the test location is moved to the mirror-symmetric position in the contralateral field at the beginning of T1. However, as the training at the original location continues, performance at the test location is gradually restored indicating an increased robustness to spatial displacement. Shifting the test location to the fovea in T2 produced no significant ($F(1,8)=1.05$, $P=0.99$) change in recognition accuracy suggesting a perfect transfer of classification performance. There were highly

Category learning induces position invariance of recognition

significant improvements of performance both during T1 ($F(1,8)=12.10$, $P<0.01$) and T2 ($F(1,8)=51.50$, $P<0.001$). These improvements were achieved in absence of any feedback at the new test locations, thus ruling out any explanation in terms of position-specific learning.

Generally, learning the patterns in the left visual field (LVF) led to a better transfer than learning in the right visual field (RVF). The effect of learning condition on classification performance was highly significant both for the transition between IN and T1 ($F(1,8)=5.71$, $P<0.001$) and for the transition between T1 and T2 ($F(1,8)=20.91$, $P<0.001$), whereas the interaction Transition \times Condition was not ($P>0.1$). The difference between the two learning conditions becomes prominent in learning time, indicated by the number of learning units needed to reach the criterion (figure 3c). Subjects learning the patterns in the LVF (C2) were significantly faster than those learning the patterns in the RVF (C1) to learn the patterns in the new test locations in transfer stage T1 ($t(8)=3.72$, $p<0.05$) and T2 ($t(8)= 2.72$, $p<0.05$).

To track the conceptual space, i.e. the perceived similarity structure of the pattern categories, during learning we reconstructed internal class representations from the confusion matrices using a probabilistic virtual prototype model (Jüttner & Rentschler 1996, 2000). Compared to the two-dimensional configuration of the class means in the defining physical feature space (cf. figure 1a), the configurations of the virtual prototypes appear degenerated to almost one-dimensional arrangements at the beginning of T1 (solid triangles in figure 3b). This deformation, which is more strongly pronounced in condition C1 than in C2, indicates a reduced perceptual dimensionality of the conceptual space at the new test location. However, further learning restores conceptual space relative to the configuration in the defining feature space (dashed triangles in figure 3b), which is preserved in T2, again indicating the increasing degree of shift invariance of internal class concepts.

Experiment 2

In Experiment 1 testing in foveal vision always succeeded testing in extrafoveal vision in both conditions. Although patterns in extrafoveal viewing were re-scaled in size in order to compensate for differences in terms of spatial resolution (Rovamo & Virsu 1979), the performance increase in T2 relative to T1 may be related to a scale-invariant advantage of foveal vision for pattern categorization (Jüttner & Rentschler 2000) rather than an increasing degree of shift invariance of the internal class representations. Experiment 2 therefore employed a

Category learning induces position invariance of recognition

complementary presentation sequence, where subjects learned the patterns in foveal view during the induction stage, before being tested for spatial generalisation in extrafoveal vision, both in the LVF and in the RVF, in either order (figure 2*b*, middle).

Owing to the foveal viewing, recognition accuracy in Experiment 2 rapidly increases during induction, which is evident already in the first decile of learning units during that stage (figure 4*a*). Moving the test location into the extrafoveal visual field (T1) led to a sharp initial drop in classification performance. However, as the foveal learning continued, performance at the test location was fully restored. Shifting the test location to the mirror-symmetric location in the contralateral field (T2) still produced a significant ($F(1,8)=22.73$, $P<0.001$) drop in recognition. However, despite the large shift performance levels at the beginning of T2 are significantly higher than at the beginning of T1 ($F(1,8)=12.98$, $P<0.01$) indicating the increase of position invariance of internal representations of pattern categories.

Learning condition had no significant effect on classification performance in Experiment 2. However, it had some effect on learning time in stage T2 (figure 4*b*, top). Patterns that after the foveal induction stage had been transferred to the LVF were significantly faster learned ($t(8)=3.32$, $P<0.05$) when subsequently transferred to the RVF (condition C2) than when shown in the reverse order (condition C1). However, no significant difference between learning conditions was observed during T1, which suggests that even a relatively brief foveal familiarisation distinctly reduces visual field asymmetry in pattern category learning.

Experiment 3

The results of Experiment 1 and 2 indicate that, in addition to the increasing degree of position invariance *at the beginning* of the two transfer stages (notably T2), there is a significant improvement of recognition performance *during* both transfer stages, even though subjects receive no feedback at the novel test locations. To assess the magnitude of this feedback-free learning effect, a control experiment was performed in which the sequence of test locations was identical to that in Experiment 2 but accompanied by simultaneous shifts of learning location, thus providing a feedback-driven reference condition. Accordingly, learning and testing in Experiment 3 first occurred in foveal and then in extrafoveal view, both in the LVF and in the RVF, in either order (figure 2*b*, bottom). Overall learning duration in Experiments 2 and 3 (figure 4*b*, top) show similar patterns. The slight reduction in the latter relative to the former can be

related to the distinct carry-over effects in Experiment 3 at the beginning of T1 and T2, which exceed the transfer observed in Experiment 2 (figure 4a). However, a comparison of the average learning rates in T1 and T2 (figure 4b, bottom) showed no significant differences between Experiment 2 and 3 ($P_s > 0.73$). Thus the feedback-free improvement of recognition performance in Experiment 2 induced by category learning at a fixed location in the visual field proceeds at a rate no slower than the improvement observed with position-specific feedback. No left-right asymmetry was observed in Experiment 3, in accordance with the learning data for the induction stage in Experiment 1 and 2.

4. DISCUSSION

Our results indicate that knowledge about pattern categories acquired at one particular location in the visual field gradually becomes available at other locations. This allows us to reconcile apparently contradictory findings of, on the one hand, a limited translation invariance of human object recognition observed in tasks involving pattern discrimination (Foster & Kahn 1985; Nazir & O'Regan 1990; Dill & Fahle 1997, 1998), and the robustness of recognition against spatial displacements found for familiar objects (Ellis *et al.* 1989; Biederman & Cooper 1991; Stankiewicz & Hummel 2002) on the other hand. In concordance with the former we observed an initial drop of performance as learning and test location became dissociated from each other (cf. beginning of transfer stage T1 in figures 3a and 4a). However, as the learning at the original location continued, performance at the test location was gradually restored to criterion level and showed increased robustness against further displacement (in transfer stage T2).

Our experiments differ from earlier work, which has focussed on the position invariance of pattern discrimination in same-different matching tasks that either avoided learning (Foster & Kahn 1985; Dill & Fahle 1998, Dill & Edelman 2001) or restricted learning processes to typical durations of less than one hour (Nazir & O'Regan 1990; Dill & Fahle 1997). In contrast, we employed a paradigm of long-term category learning that involved three pattern classes defined by multiple exemplars and extended over several days for up to eleven hours (median: 6.8 hours and 3.5 hours in Exp. 1 and 2, respectively) of total learning time. Throughout Experiment 1 and 2 subjects only received feedback about category membership at the original learning location but not at the test location. Learning progress at the test location therefore was not a practice effect owing to position-specific feedback. The translational invariance observed for category learning

in the present experiments stands in marked contrast to the positional specificity that has been observed for a range of other perceptual learning tasks (Karni & Sagi 1991; Shiu & Pashler 1992; Fahle *et al.* 1995), even though recent evidence suggests that category learning may affect even early perceptual processing (Notman *et al.*, 2005).

Further evidence for the crucial role of category learning in the acquisition of position invariance comes from the advantage of the LVF relative to the RVF for the learning and generalisation of novel patterns. As our stimulus patterns, both within and between categories, only differed in terms of the spatial relationships between their constituting part components, the results are compatible with previously reported behavioural dissociations showing a LVF advantage for the processing of metric coordinate representations (Kosslyn *et al.* 1989) and exemplar-specific encoding of pattern categories (Marsolek 1999). This asymmetry can be related to the predominant activation of right prefrontal and parietal areas reported in neuroimaging studies for tasks involving visual reasoning (Seger *et al.*, 2000) and visuospatial working memory (Jonides *et al.* 1993; Smith & Jonides 1997). In contrast, a left dorsolateral prefrontal activation has been found for tasks involving analytic problem solving (Smith & Jonides 1997) and formal reasoning (Wharton & Grafman 1998). For visual learning, this activation shows a distinct dependency on learning status (Seger *et al.* 2000) and might indicate a left-hemispheric specialisation for the formation of abstract categories (Marsolek 1999). The bilateral stimulation mediated by the foveal learning in Experiment 2 may facilitate the recruitment of predominantly left-hemispheric mechanisms involved in category abstraction and their interaction with the image-based processing of individual category exemplars predominantly located in the right-hemisphere, thus yielding an attenuation of behavioural lateralisation with increasing learning progress. Consistent with this interpretation, the entry-level categorization of familiar patterns or objects often shows no visual-field asymmetry in normal subjects (e.g., Peterzell *et al.* 1989; Biederman & Cooper 1991).

Insight into how category learning induces position invariance in object recognition is provided by computer simulations performed in earlier work involving the same type of stimulus material (Jüttner *et al.* 1997, 2004; Rentschler & Jüttner 2005). These simulations suggest that category learning of Compound Gabor gratings relies on production rules that combine multiple attributes representing either properties of individual pattern parts or those of part relations. The distinction between two attribute formats allows to account for variations in the degree of

perceptual invariance. In the present experiments, a part-specific encoding of visual field position of individual pattern components would yield rules that are highly location specific, whereas an encoding of relative position for adjacent components would produce rules that are translation invariant. These different ways of encoding positional information may have a correspondence in the systematic change of receptive field properties along the higher stages of the ventral visual pathway in primates involving area V4, the posterior (TEO) and the anterior (TE) region of the inferotemporal cortex. Cells in TEO and V4 have smaller receptive fields and a preference for simple patterns (Kobatake & Tanaka 1994), and thus may serve the extraction of part-specific information. In contrast, cells within TE tend to have large receptive fields that often include the fovea (Ito *et al.* 1995). Nevertheless, such cells show a far greater sensitivity to retinal position than predicted by receptive field size (Rolls *et al.* 2003). Their preference for complex configural patterns rather than isolated pattern components (Tanaka 1996) suggests that these cells may play an important role in the representation of part-relational information.

A shift in the format of positional information during category acquisition would then become manifest in an emerging position invariance of visual recognition without requiring any position-specific feedback. Adaptive use of multiple features during category learning has been previously demonstrated behaviourally in humans (Nosofsky 1986; Schyns *et al.* 1998; Op de Beeck *et al.* 2003) and, at a behavioural as well as neurophysiological level, in monkeys (Freedman *et al.* 2001; Sigala & Logothetis 2002). For foveal viewing we have recently shown that flexible use of position-indexing during learning can explain the acquisition and generalisation of mirror-image categories (Rentschler & Jüttner 2007). Similar considerations with regard to appearance-related attributes may account for other phenomena of perceptual invariance, such as against changes in contrast polarity (Jüttner *et al.* 2004).

Current approaches to object recognition and understanding generally assume that position invariance is achieved at an early level of visual processing and reflects an automatic, adaptive response to the spatiotemporal statistics of the visual environment (e.g. Wallis & Rolls 1997; Riesenhuber & Poggio, 1999; Wiskott & Sejnowski 2002; Cox *et al.*, 2005). Against this background our findings add a novel perspective, as they demonstrate that invariance to positional changes is also a by-product of the top-down structuring of the visual world imposed by the process of category acquisition. In this way, position invariance induced by category learning might act complementary to invariance mechanisms of more limited scope, which may

be active at early and intermediate levels of feature processing and result from a conjunctive sampling of the visual field (Riesenhuber & Poggio 1999) or partial generalizations built upon past sensory experience (Ullman & Soloviev 1999).

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) and the Economic & Social Research Council (ESRC).

REFERENCES

- Biederman, I. & Cooper E. E. 1991 Evidence for complete translational and reflectional invariance in visual object priming. *Perception* **20**, 585-593.
- Cox, D. D., Meier, P., Oertelt, N. & DiCarlo, J. J. 2005 >Breaking= position-invariant object recognition. *Nat. Neurosci.* **8**, 1145-1147.
- DiCarlo, J.J. & Maunsell, J.H.R. 2003 Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *J. Neurophysiol.* **89**, 3264-3278.
- Dill, M. & Edelman, S. 2001 Imperfect invariance to object translation in the discrimination of complex shapes. *Perception* **30**, 707-724.
- Dill, M. & Fahle, M. 1997 The role of visual field position in pattern-discrimination learning. *Proc. R. Soc. Lond. B* **264**, 1031-1036.
- Dill, M. & Fahle, M. 1998 Limited translation invariance of human visual pattern recognition. *Percept. Psychophys.* **60**, 65-81.
- Ellis, R., Allport, D. A., Humphreys, G. W. & Collis, J. 1989 Varieties of object constancy. *Q. J. Exp. Psychol. A* **41**, 775-796.
- Fahle, M., Edelman, S. & Poggio, T. 1995 Fast perceptual learning in hyperacuity. *Vision Res.* **35**, 3003-301.
- Foster, D. H. & Kahn, J. I. 1985 Internal representations and operations in the visual comparison of transformed patterns: effects of point-inversion, positional symmetry, and separation. *Biol. Cybern.* **51**, 305-312.

- Freedman, D., Riesenhuber, M., Poggio, T. & Miller, E. 2001 Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* **291**, 312-316.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Shouten, J.L. & Pietrini, P. 2001 Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425-2430.
- Ito, M., Tamura, H., Fujita, I. & Tanaka, K. Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J. Neurophysiol.* **73**, 218-226 (1995).
- Jonides, J., Smith, E. E., Koeppe, R. R., Awh, E., Minoshima, S. & Mintun M. A. 1993 Spatial working memory in humans as revealed by PET. *Nature* **363**, 623-625.
- Jüttner, M. & Rentschler, I. 1996 Reduced perceptual dimensionality in extrafoveal vision. *Vision Res.* **36**, 1007-1021.
- Jüttner, M. & Rentschler, I. 2000 Scale invariant superiority of foveal vision in perceptual categorization. *Eur. J. Neurosci.* **12**, 353-359.
- Jüttner, M., Caelli, T. & Rentschler, I. 1997 Evidence-based pattern classification: A structural approach to human perceptual learning and generalization. *J. Math. Psychol.* **41**, 244-259.
- Jüttner, M., Langguth, B. & Rentschler, I. 2004 The impact of context on pattern category learning and representation. *Vis. Cogn.* **11**, 921-945.
- Karni, A. & Sagi, D. 1991 Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proc. Natl. Acad. Sci. USA* **88**, 4966-4970.
- Kobatake, E. & Tanaka, K. 1994 Neural selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* **71**, 856-867.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. Tang, J. & Gabrieli, J. D. E. 1989 Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *J. Exp. Psychol. Human* **15**, 723-735.
- Marsolek, C. J. 1999 Dissociable neural subsystems underlie abstract and specific object recognition. *Psychol. Sci.* **10**, 111-118.
- Nazir, T. & O'Regan, J. K. 1990 Some results on translation invariance in the human visual system. *Spatial Vision* **5**, 81-100.
- Nosofsky, R. N. 1986 Attention, similarity and the identification-categorization relationship. *J. Exp. Psychol. Gen.* **115**, 39-61.
- Notman, L. A., Sowden, P. T. & Özgen E. 2005 The nature of learned categorical perception

Category learning induces position invariance of recognition

- effects: a psychophysical approach. *Cognition* **95**, B1-B14.
- Op de Beeck, H., Wagemans, J. & Vogels, R. 2003 The effect of category learning on the representation of shape: dimensions can be biased, but not differentiated. *J. Exp. Psychol. Gen.* **132**, 491-511.
- Peterzell, D. H., Harvey, L. O. JR. & Hardyck, C. D. 1989 Spatial frequencies and the cerebral hemispheres: Contrast sensitivity, visible persistence, and letter classification. *Percept. Psychophys.* **46**, 443-455.
- Rentschler, I. & Jüttner, M. 2007 Mirror-image relations in category learning. *Vis. Cogn.* **15**, 211-237.
- Riesenhuber, M. & Poggio, T. 1999 Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**, 1019-1025.
- Rolls E.T. 1992 Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philos. T. Roy. Soc. B* **335**, 11-21.
- Rolls, E.T., Aggelopoulos, N.C. & Zheng, F. (2003) The receptive fields of inferior temporal cortex neurons in natural scenes. *J. Neurosci.* **23**, 339-348.
- Rosch, E. 1978 Principles of Categorization. In *Cognition and Categorization* (eds. E. Rosch, E. & B. Lloyd), pp. 27-48. Hillsdale NJ: Erlbaum.
- Rovamo, J. & Virsu, V. 1979 An estimation and application of the human cortical magnification factor. *Exp. Brain Res.* **37**, 495-510.
- Schyns, P. G., Goldstone, L. & Thibaut, J. P. 1998 The development of features in object concepts. *Behav. Brain Sci.* **21**, 1-53.
- Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H. & Gabrieli, J. D. E. 2000 Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia* **38**, 1316-1324.
- Shiu, L.-P. & Pashler, H. 1992 Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept. Psychophys.* **52**, 582-588.
- Sigala, N. & Logothetis, N. 2002 Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* **415**, 318-320.
- Smith, E. E. & Jonides, J. 1997 Working memory: a view from neuroimaging. *Cognitive Psychol.* **33**, 5-42.
- Stankiewicz, B. & Hummel, J.E. 2002 Automatic priming for translation- and scale-invariant

representations of object shape. *Vis. Cogn.* **9**, 719-739.

Tanaka, K. 1996 Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19**, 109-139.

Ullman, S. & Soloviev, S. 1999 Computation of pattern invariance in brain-like structures. *Neural Networks* **12**, 1021-1036.

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

Wallis, G. & Rolls, E. T. 1997 Invariant face and object recognition in the visual system. *Prog. Neurobiol.* **51**, 167-194.

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 Res.* **38**, 487-491.

Wharton, C. M. & Grafman, J. 1998 Deductive reasoning and the brain. *Trends Cogn. Sci.* **2**, 54-59.

Wiskott, L. & Sejnowski, T. J. 2002 Slow feature analysis: unsupervised learning of invariances. *Neural Comput.* **14**, 715-770.

Figure captions

Figure 1. (a) A set of fifteen compound Gabor gratings, consisting of a fundamental plus its third harmonic within a Gaussian aperture, served as learning stimuli in the experiments. Stimulus variation was restricted to the amplitude b and phase angle φ of the third harmonic. Thus the configural structure of each greylevel pattern along its horizontal axis was uniquely determined by its coordinates $\xi=b \cos \varphi$ and $\eta= b \sin \varphi$ in a two dimensional Fourier feature space. Within this feature space, the fifteen learning stimuli formed three cluster of equal variance defining three classes (1,2,3) to be learned by the subject. Scale: 1 unit = 20 cd/m². (b) Illustration of the actual greylevel representations of the patterns.

Figure 2. (a) Interleaved learning and testing schedule. Subjects were trained to criterion in a series of learning units, each having a learning phase (L), during which patterns of the learning set were randomly presented followed by their corresponding class label, and a test phase (T), during which observers had to categorize each stimulus. (b) In Experiment 1 (top) and 2 (middle), viewing location during the learning phase of each learning unit was kept constant, whereas the location during the test phase was systematically varied between locations in the left visual field (LVF), the right visual field (RVF), and the central position. For example (see inset), in condition C2 of Experiment 1 the patterns were always presented in the LVF during the learning phase of each learning unit, whereas testing successively involved locations in the LVF (during the induction stage IN), the RVF (transfer stage T1) and the centre (transfer stage T2). Experiment 3 (bottom) replicated the sequence of test locations in Experiment 2, however learning and testing always occurred at the same position.

Figure 3. (a) Asymmetric effects of extrafoveal pattern category learning in the right (condition C1) and left (C2) visual field on position-invariance. Individual time-normalized learning curves of ten subjects derived from the average percent-correct scores across each decile of learning units. Symbols show group means for each condition in the first and last decile of each stage. The horizontal dashed line indicates chance level. Note the dissociation of the curves in the two conditions, even though performance in both groups improves in the absence of location-specific feedback during T1 and T2. (b) Visualization of the similarity structure between internal class

representations in terms of a virtual-prototype model (see Material and Methods). As indicated each prototype configuration corresponds to the group mean of the first or last decile of the learning curves shown in (a), with e denoting the root of the mean squared error between observed and model-predicted classification frequencies. Compared to the configuration of the class means in physical feature space (dashed triangle, cf. figure 1), the virtual prototype configurations for subjects in condition C1 appear particularly degenerated, mainly due to an insufficient separation of class 2 and class 3. (c) Mean number of learning units (LUs) to reach the learning criterion in the three stages of the experiment. Subjects learning the stimuli in the left visual field (C2) show a significantly better transfer to the new locations tested in T1 and T2. Error bars indicate ± 1 s.e. of the mean.

Figure 4. (a) Spatial generalisation of foveal pattern category learning in extrafoveal vision. Coloured symbols show group means of classification performance in the first and last decile of learning units for the two learning conditions in Experiment 2. Percent-correct scores show a distinct drop at the beginning of T2. However, extrafoveal classification accuracy becomes increasingly less susceptible to positional changes as foveal learning continues, as a comparison with the corresponding performance scores of a reference condition with no separation of learning and test location (Experiment 3, black and white symbols) demonstrates. (b) Mean learning time (top) and mean learning rate (bottom) in Experiment 2 and 3. Note that the feedback-free learning during T1 and T2 in Experiment 2 proceeds at the same rate as the feedback-driven learning in Experiment 3. Error bars indicate ± 1 s.e. of the mean.

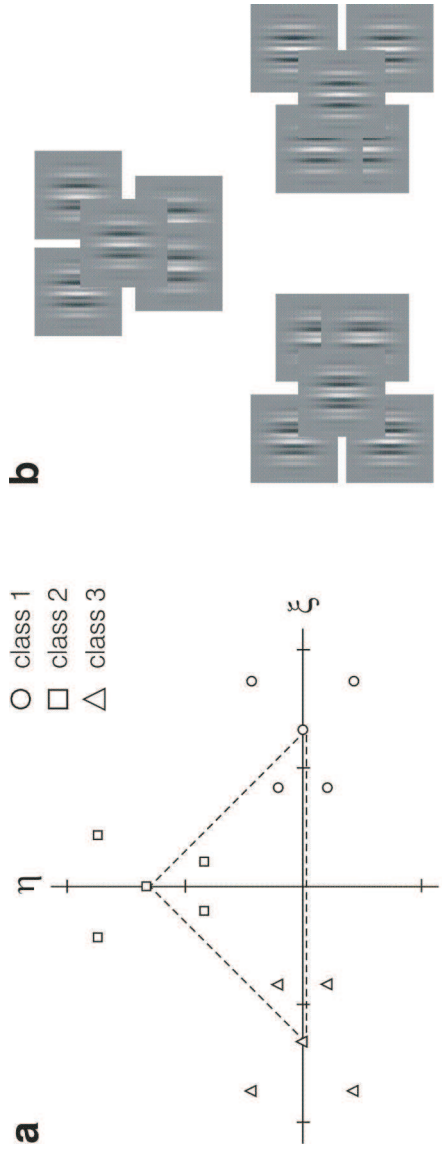
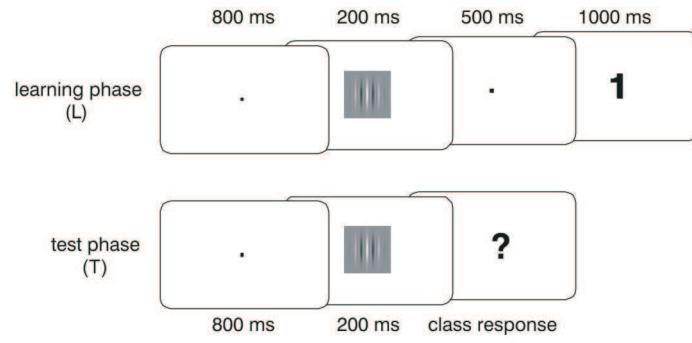


Fig. 1

a



b

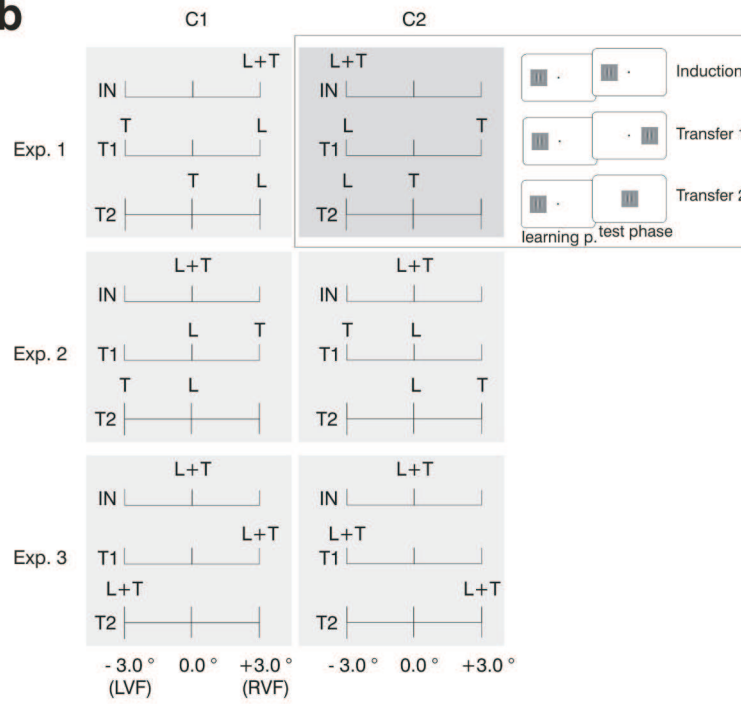


Fig. 2

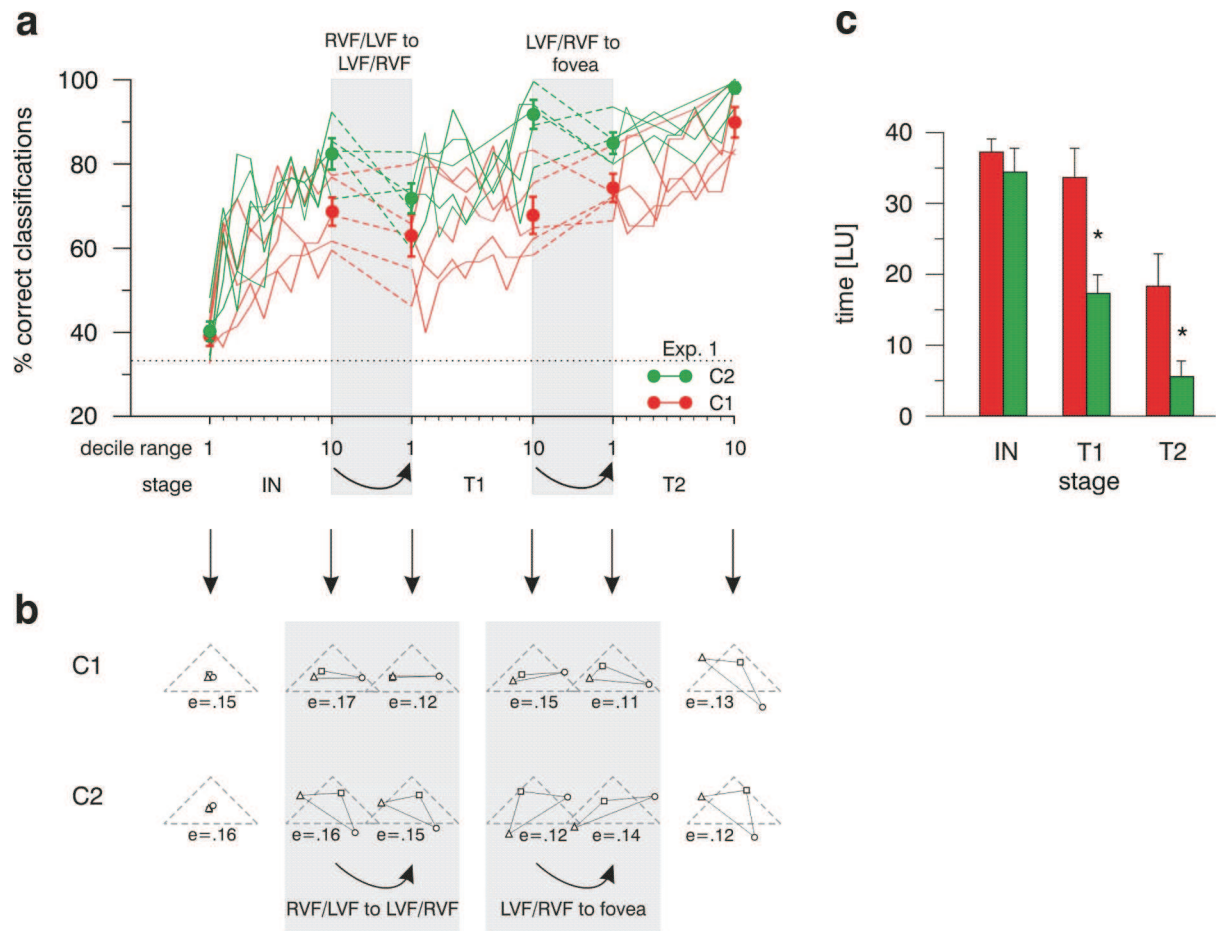


Fig. 3

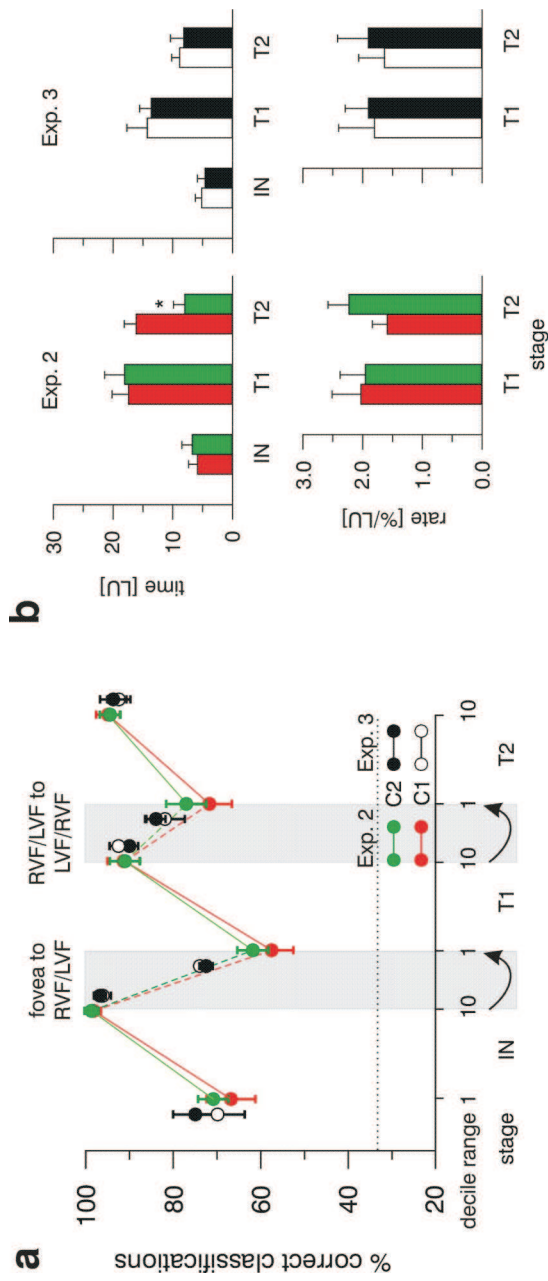


Fig. 4