

# Visual crowding and category specific deficits for pictorial stimuli: A neural network model

**Tim M. Gale** <sup>1,2,3</sup>     **D. John Done** <sup>1</sup>     **Ray J. Frank** <sup>3</sup>

1. Division of Psychology, University of Hertfordshire, UK

2. QEII Hospital, Welwyn Garden City, UK

3. Division of Computer Science, University of Hertfordshire, UK

Requests for reprints should be addressed to Dr. T.M. Gale, Division of Computer Science, University of Hertfordshire, College Lane, Hatfield, Herts, AL10 9AB, UK. Telephone +44(0)1707 365383, Fax +44(0)1707 335168, e-mail [t.gale@herts.ac.uk](mailto:t.gale@herts.ac.uk)

Some of this work formed part of the first author's Ph.D. thesis and was subsequently supported by a post-doctoral fellowship in the Divisions of Psychology and Computer Science at The University of Hertfordshire. The authors are grateful to: Deborah Wilson for image collection and pre-processing; Koulla Asimakapoulou for normative data collection; Neil Davey, Keith Laws, Stella George, Bruce Hajilou and Steve Hunt for commenting on earlier drafts. Finally, we would like to thank David Plaut and 2 anonymous reviewers for their insightful contributions to this paper.

## ABSTRACT

This paper describes a series of modular neural network simulations of visual object processing. In a departure from much previous work in this domain, the model described here comprises both supervised and unsupervised modules and processes *real* pictorial representations of items from different object categories. The unsupervised module carries out bottom-up encoding of visual stimuli, thereby developing a 'perceptual' representation of each presented picture. The supervised component then classifies each perceptual representation according to a target semantic category. Model performance was assessed (i) during learning, (ii) under generalisation to novel instances, and (iii) after lesion damage at different stages of processing. Strong category effects were observed throughout the different experiments, with living things and musical instruments eliciting greater recognition failures relative to other categories. This pattern derives from within-category similarity effects at the level of perceptual representation and our data support the view that visual crowding can be a potentially important factor in the emergence of some category specific impairments. The data also accord with the Cascade model of object recognition since increased competition between *perceptual representations* resulted in category-specific impairments even when the locus of damage was within the *semantic* component of the model. Some strengths and limitations of this modelling approach are discussed and the results are evaluated against some other accounts of category specific recognition failure.

## INTRODUCTION

One of the most consistently reported findings from neuropsychological work on visual object recognition is the emergence of category specific recognition deficits after *different* forms of neurological impairment (e.g. Basso, Capitani and Laiacona, 1988; Farah, McMullen and Meyer, 1991; Farah, Meyer and McMullen, 1996; Gonnerman, Andersen, Devlin, Kempler and Seidenberg, 1997; Humphreys, Riddoch and Quinlan, 1988; Warrington and McCarthy, 1987). Category specific deficits are characterised by a sparing of recognition and naming abilities for certain classes of object in contrast to a marked deficit for others. The most frequently reported *broad* pattern of impairment is a difficulty in naming pictures of living things relative to non-living things (e.g. Arguin, Bub and Dudek, 1996; Sartori and Job, 1988; Warrington and Shallice, 1984). The reverse dissociation has also been documented but is more unusual (e.g. Hillis and Caramazza, 1991; Sacchett and Humphreys, 1992; Turnbull and Laws, 2000; Warrington and McCarthy, 1983). Although the living/ non-living distinction has been the focus of many studies, it is notable that category deficits are rarely this pure: neurological patients with purported living thing deficits often present with similar impairments for musical instruments whilst those with non-living thing deficits often have difficulty naming body-parts (Gainotti, Silveri, Daniele and Giustolisi, 1995; Parkin and Stewart, 1993; but see Caramazza and Shelton, 1998 for an unusually pure deficit in naming animals).

It has been argued that apparent category effects may emerge from inadequate control of variables known to predict object recognition performance. Funnell and Sheridan (1992) and Stewart, Parkin and Hunkin (1992) both pointed out that living things, relative to non-living things, are less familiar, more visually complex and have lower word frequency. Since these three variables are strong predictors of naming performance, it is plausible that a general reduction in cognitive efficiency might give rise to marked anomia for stimuli that are visually complex, conceptually unfamiliar and infrequently referred to (i.e. living things). However, whilst this explanation has gained some support it cannot explain the reversed pattern of deficit, nor the persistence of category effects under careful control of predictor variables (e.g. Farah et al., 1991; 1996; Laiacona, Barbarotto and Capitani, 1993; Mauri, Daum, Sartori, Riesch and Birbaumer, 1994; Sartori, Job and Coltheart, 1993; Sheridan and

Humphreys, 1993). Thus, although experimental artefact may underlie some reported cases, a great many must undoubtedly be genuine.

Category-specific impairments may be informative about the functional architecture for representing knowledge in the brain and there has been extensive debate on their implications for theories of object processing and organisation of semantic memory. Most advances in this area have been made under neuropsychological investigation (e.g. Hillis and Caramazza, 1991; Laiacona, Capitani, and Barbarotto, 1997; Warrington and Shallice, 1984) although other approaches have also made substantial inroads, including electrophysiological studies in primates (see Logothetis and Sheinberg, 1996; Tanaka, 1996; Yamane, Kaji and Kawano, 1988) and visual object recognition studies in normal adults (e.g. Gaffan and Heywood, 1993; Laws and Neve, 1999; McRae, de Sa and Seidenberg, 1997). More recently, there have been important contributions from studies of connectionist neuropsychology (e.g. Devlin, Gonnerman, Andersen and Seidenberg, 1998; Durrant-Peatfield, Tyler, Moss and Levy, 1997; Humphreys, Lamote and Lloyd-Jones, 1995; Small, Hart, Nguyen and Gordon, 1995). It is the latter approach upon which this paper will principally focus although some *consistencies* between data from neural network models and other approaches will also be discussed.

Reviews and critiques of candidate explanations for category specific deficits are widespread in the neuropsychological literature (e.g. Caramazza, 1998; Caramazza, Hillis, Leek and Miozzo, 1994; Coltheart, Inglis, Cupples, Michie, Bates and Budd, 1998; Farah et al., 1996; Humphreys and Forde, in press; Moss, Tyler, Durrant-Peatfield and Bunn, 1998) so here we will limit our discussion only to 'emergent property' accounts (see Caramazza, 1998) in which connectionist models have had the greatest influence. Emergent property accounts include the sensory/ functional theory (SFT) of Warrington and Shallice (1984), the organised unitary contents hypothesis (OUCH) of Caramazza, Hillis, Rapp and Romani (1990) and the cascade model/ structural description hypothesis of Humphreys et al., (1988). We will firstly discuss SFT and OUCH since they both account for category specific impairments at the level of *semantic* representation.

Warrington and Shallice (1984) argued the case for a multi-modal semantic store comprising perceptual and functional sub-systems. Category specificity was accounted for by positing that semantic representations of living things and non-living things are *predominantly* specified in terms of visual/ perceptual and functional/ associative properties respectively. In such a framework, it is plausible that damage within a single module could induce recognition failure for one broad class of item. In the later OUCH model of Caramazza et al. (1990) there was no reliance upon modality specificity but, rather, it was assumed that items belonging to 'natural kind' classes share a comparatively higher number of attributes (e.g. moves, has eyes, eats, etc.) and that strongly correlated properties (e.g. flight, feathers, beak) are represented in adjacent substrate. Under these assumptions, properties of living things would tend to be close together which would render stored knowledge of living things more prone to catastrophic effects of localised neural damage.

Several connectionist models of emergent category specific phenomena have followed assumptions about semantic organisation deriving from SFT. However, some have also included aspects of OUCH. The model described by Farah and McClelland (1991) was a pure simulation of SFT. A vector of binary features represented either pictorial or verbal stimuli. These inputs were associated with semantic output targets in a feed-forward network. Semantic space was partitioned into perceptual and functional semantics and, based upon information reported in dictionary definitions of item names, living things were represented predominantly by perceptual attributes whilst non-living things were represented more evenly across both feature types. Simulated lesioning produced deficits for living things when damage was located in perceptual semantics, even when connectivity loss was minimal. Lesioning of functional semantics elicited non-living thing deficits, but only at moderate to severe levels. Thus, category effects in the model derived from the ratio of perceptual to functional features specifying a given item and the anatomical separation of sensory and functional semantics. A particular strength of this model in articulating the SFT theory is that it predicts the relative frequencies of living and non-living deficits within a single architecture. However, whilst it succeeds on grounds of parsimony, its explanation is limited to semantic dysfunction since it cannot simulate aspects of pre-semantic processing (cf. Humphreys et al. 1995).

In a later model of semantic activation described by Devlin et al. (1998), locally represented semantic features were either distinctive (activated by few stimuli) or shared (activated by many stimuli). Consistent with OUCH and the work of McRae et al. (1997), the ratio of shared to distinctive features was higher for living things resulting in an increased level of correlated activity between units representing natural kinds. Classification errors emerged for non-living things after low levels of semantic layer lesioning but high intercorrelations between the majority of natural kind features afforded living things protection against relatively minor connectivity loss. However, when lesioning was increased, a relative deficit for living things was observed which derived from a catastrophic loss of correlated feature information. Durrant-Peatfield et al. (1997; see also Tyler et al., 2000) also utilised the principle of correlated features *combined* with a sensory/ functional distinction in a hybrid model of SFT and OUCH. A motivation behind this work was the more recent finding that living thing deficits are not *always* accompanied by selective deficits for perceptual properties (Laiacona et al., 1993; 1997; Sheridan and Humphreys 1993), an empirical finding which is not easily explained under SFT alone. For living things, there was high correlated activity between units representing shared perceptual features (e.g. 'wings') and shared functional features (e.g. 'flight'). For non-living things, on the other hand, correlated activity was high between units representing distinctive perceptual properties and distinctive functional properties. Following lesioning, identification errors and intra-category confusions were greatest for living things, yet superordinate categorisation was preserved. Loss of perceptual properties varied in kind, rather than degree: whereas distinctive perceptual features were susceptible for living things, shared perceptual features were more susceptible for artefacts. Lesioning severity was also an important factor: at 60% or lower severity the impairment was greatest for living things but, at higher levels, the converse was true (cf. Devlin et al. 1998).

The three models discussed so far can all account for living *and* non-living deficits although they do so as a result of *different* assumptions about semantic representation. Whilst certain assumptions are necessary to test any theoretical account, the same assumptions may also limit a model's ability to objectively explore the target domain. As pointed out by Reeke and Edelman (1988), any model relying on experimenter-imposed representation risks being 'homuncular', whereby it can exhibit similar behaviour to a human yet fail to capture important elements of underlying processing in the

target domain (Reeke and Sporns, 1993; see also Perry, 1999). A potential problem with nameable attribute representations is that inter-item similarity is constrained in an artificial way depending on the choice of items and features which are represented. For example, the animals *dog* and *mouse* have identical representations if specified across the set of attributes: <has fur>, <is a mammal>, <lives indoors>, <has a tail>, <has 4 legs>; however, they differ markedly if encoded over the set: <is small>, <eats grain>, <barks>, <eats meat>, <is a pet>. Thus, in any model where behaviour is critically dependent upon similarity between items, the choice of features used to encode a diverse range of items has high explanatory power for emergent behaviour. This may not be problematic if properties of the artificial representations hold for real representations but, without strong evidence to support such a comparison, any model must be treated with at least some scepticism. For example, assumptions about the nature of feature-based representation in both the Devlin et al. (1998) and Durrant-Peatfield et al. (1997) models generate *contradictory* patterns of performance after lesioning.<sup>1</sup>

All three models discussed so far simulate processing in the semantic system and do not explore other architectures which purportedly contribute to full visual object recognition (e.g. Humphreys et al. 1988). Whilst this does not diminish their ability to elucidate specific theoretical accounts of category specificity (i.e. SFT, OUCH or both), it does limit their ability to test whether non-semantic processes may also contribute to such deficits. The third emergent property account, however, proposes that factors operating at the level of structural description (i.e pre-semantic, perceptual-categorical representation) can underlie some category-specific impairments (e.g. Forde, Francis, Riddoch, Rumiaty and Humphreys, 1997; Humphreys et al., 1988; Riddoch and Humphreys, 1987a; Sartori, Job, Miozzo, Zago and Marchiori, 1993). The central tenet with this theory is that structural descriptions of living things embody greater item-to-item similarity and are more visually complex than those of non-living things. These characteristics render living things more prone to confusion when processing capacity is constrained. This explanation is able to account for the most commonly observed pattern of category impairment (i.e. living thing deficits) but not the reverse pattern and is not, therefore, mutually exclusive to some of the accounts discussed earlier. Moreover, it only comfortably accounts

---

<sup>1</sup> Such anomalous model performance may also be explained by the finding that critical behaviours of some connectionist models (e.g. Devlin et al., 1998) are susceptible to small manipulations in architecture (Perry, 1999).

for recognition difficulties with pictorial stimuli, since tasks involving recognition of verbal stimuli (i.e. object names, definitions, etc) may not require interrogation of structural representations. The structural description account does predict greater confusion in discrimination learning and slower reaction times in speeded recognition tasks with normal subjects, since difficulties in resolving structural descriptions for living things arise, at least in part, as a result of stimulus characteristics (e.g. Humphreys et al., 1988; Lloyd-Jones and Humphreys, 1997). In an influential study, Gaffan and Heywood (1993) trained monkeys to discriminate between pairs of pictures selected from the corpus of Snodgrass and Vanderwart (1980). Error rates during training were three times higher when the pairs comprised living things relative to non-living things. Furthermore the disparity in error rate grew logarithmically as set size increased from 8 stimuli at the beginning to 128 by the end, suggesting that the finding is not artefactual. Using the same stimulus set again, Gaffan and Heywood (1993) demonstrated a similar pattern of performance limitation in speeded, degraded visual recognition tasks with normal human subjects and argued that ‘perceptual crowdedness’ offered the best explanation for their data. Perceptual crowding occurs when the range of partial features extracted from a stimulus object, and then bound together to form an internal structural description, is no longer sufficient to differentiate between visually similar items. During learning, perceptual crowding may become apparent only when the size of the stimulus set exceeds a critical boundary. In the natural world one might expect it to occur within certain biological categories (e.g. birds, fish, butterflies) where expertise is required to represent each member in a multidimensional space with a dimensionality that is sufficient to allow inter-item separation (Edelman, 1998). Following brain damage, category-specific deficits may emerge for perceptually crowded categories when restriction is placed on the number of dimensions available to represent objects. This hypothesis was tested by Humphreys et al. (1995) in a connectionist simulation of the Cascade model utilising an interactive activation (IAC) model first described by McClelland and Rumelhart (1981; 1988). The network consisted of 3 layers, each of 20 units. ‘Visual’ stimuli comprised 20 input vectors representing 10 animals and 10 items of clothing. Similarity ratings for the real-life referents were derived from ‘visual’ comparisons made by judges and these were captured in the model by representing correlated patterns of activity across the input units, which represented the structural description system. The inputs then connected to a semantic layer with localised representation (i.e. one unit per object) and



this layer, in turn, was connected to a layer of name units which coded the identity of each input. At the same level a separate set of two super-ordinate units connected to the preceding layer to simulate category decision (i.e. 'animal' versus 'clothing'). Following simulated lesioning, identification error rates and reaction times were higher for structurally similar inputs (animals), irrespective of lesion site, although at higher lesion severity the 'visually' similar and dissimilar inputs produced comparable error rates. Thus, with a novel method of representing visual similarity in their model, Humphreys et al. (1995) demonstrated a strong association between structural similarity and impaired recognition for living things. However, a number of concerns arise with this model: firstly, although between-item visual similarity ratings were collected from uninterested parties, it is probable that such judgements also reflect semantic proximity due to difficulty in disambiguating visual and semantic similarity in visual representations (Dixon, Bub and Arguin, 1997). For example, knowing a-priori that a duck and kingfisher are both birds may exaggerate perception of their structural similarity even though they have dissimilar shapes. Secondly, as Dixon et al. (1997) also point out, the kind of visual similarity that normal subjects ascribe to a set of objects may not be the sort that is crucial emergence of living thing deficits. Finally, representing a small number of stimuli (20 objects from 2 categories) in a model of relatively large size (62 units) is unlikely to provide sufficient generalisation performance (Baum and Haussler, 1989; Hinton, 1989) thereby restricting the model's potential to capture underlying processing in the simulated domain. Thus, although this model demonstrates a potential role for visual similarity in emerging living thing deficits, it would be informative to explore the idea further with a larger set of stimuli in which (i) the number of object categories is increased, (ii) the visual similarity between exemplars is captured in a way which does not also reflect semantic proximity and, (iii) other groups of item with special significance in the category specific debate (i.e. musical instruments) are also included.

So far, we have discussed four different connectionist models which offer alternative ways of conceptualising category-specific recognition failure. Three of these models simulate dysfunction within the semantic system and, although impairments are accounted for by quite different theoretical assumptions, none of these models rule out the possibility that pre-semantic representations may also play an important role in *some* impairments. On the other hand, the Humphreys et al. (1995) model

demonstrates that pre-semantic representations *may* underlie deficits for natural kinds but does not rule out the possibility that some cases might be better explained by models of semantic function. Nonetheless, given that perceptual representations are activated prior to semantic representations in the visual object recognition process (e.g. Humphreys et al. 1988), it is important to establish their potential contribution to category biases in order that these effects can be accounted for in models of the semantic system.

In this paper we investigate the contribution of perceptual representations to emergent category specificity in a series of simulations utilising an *unsupervised* neural network. All connectionist models of category specificity described so far have used supervised training procedures or, in the case of Humphreys et al. (1995), the IAC model. In these models, representations are largely determined by the modeller and correct mapping between input and output is determined by a ‘tutor’ learning algorithm which makes corrections by assigning ‘blame’ to different units for their relative role in producing an erroneous output (or, in IAC, by setting connections between nodes beforehand). However, there is evidence that human perceptual representations can arise through self organisation. For example, very young infants can form superordinate and basic level ‘perceptual’ categories in the absence of language ability and without feedback from the environment (e.g. Behl Chadha, 1996; Behl-Chadha, Eimas and Quinn, 1995; Eimas and Quinn, 1994; Eimas, Quinn and Cowan, 1994; Quinn and Eimas, 1986). In light of these findings, unsupervised neural networks *may* have more ecological validity in modelling the acquisition of perceptual-categorical information. In a further departure from some other models, we have not used nameable features to represent visual characteristics of items in this model. This is because it is difficult to prove that any given set of features adequately reflects the kind of information extracted from visual stimuli during object recognition and is of sufficient size and content to provide an accurate reflection of between-item visual similarity. Instead, we present real images as input to our model so that categorisation can be achieved only by analysing the topography of visual information within each presented picture. Our investigation is restricted to exploring the final emergent property account for living thing category specific deficits - namely the structural description/ visual crowding hypothesis. Although we utilise semantic representations in some simulations, we do so only to investigate the mapping between perceptual representations and

semantic knowledge and we make no assumptions about categorical differences in the nature of semantic representation. Our model should provide a stringent test of perceptual crowding theory since any purported visual similarity between stimuli should be captured by the self-organising component of the model. This will permit investigation of between category differences in the nature of emergent perceptual representations and whether these are sufficient to explain *any* patterns of category specific impairment.

## GENERAL METHOD

This section describes aspects of methodology that are common to all experiments reported in this paper. These include (i) the way in which training data was assembled and (ii) the various neural architectures used to simulate components of the visual object recognition process.

### Training Patterns: Image Choice and Preparation

The training exemplars were derived from the superordinate categories of animals, musical instruments, clothing and furniture and were chosen to fall within the tripartite hierarchy of increasing specificity originally proposed by Eleanor Rosch and colleagues (Rosch, 1973; Rosch, 1975; Rosch and Mervis, 1975; Rosch, Mervis, Gray, Johnson and Boyes Braem, 1976; see also Collins and Loftus, 1975; Collins and Quillian, 1969; Mervis and Greco, 1984; Tversky and Hemenway, 1984). This allowed exemplars to be grouped at (i) the superordinate level (e.g. animals, furniture), (ii) the basic level (e.g. fishes, birds, clocks, chairs), and (iii) the subordinate level (e.g. pikes, bullfinches, alarm clocks, office chairs). This categorical structure was imposed not because we believed it would accurately reflect the real world of objects but simply because it ensured that each superordinate was represented by an equal number of exemplars and characterised by the same degree of object specificity. Seven basic level categories were chosen for each superordinate as follows: *Animals* - snakes, frogs, fishes, spiders, deer, mice, and birds; *Furniture* - chairs, beds, chests, clocks, lamps, tables and wardrobes; *Musical Instruments* - guitars, violins, pianos, saxophones, electronic keyboards, flutes and drums; *Clothing* - jackets, shorts, trousers, gloves, boots, shoes and t-shirts. These basic level categories were chosen to reflect a representative range of visual diversity within each superordinate

so as not to deliberately bias any category towards visual crowding. Thus, to give some examples, the category of animals comprised basic level representatives from different genera (e.g. amphibians, mammals, fish, birds, etc.) and the category of musical instruments comprised string, wind and percussive types (e.g. piano, drum, guitar, flute). All basic level categories were listed in Battig and Montague (1969) and most, excluding a minority of musical instruments and clothing items, were listed in Snodgrass and Vanderwart (1980). To minimise experimenter bias, the task of choosing images was given to someone who was not part of the investigative team and who was naive to the rationale of the study. The guidelines given to this person were to select either colour photographs or very high quality coloured drawings which depicted items in a *typical* orientation, preferably with no foreshortening of the principal axis (see Palmer, Rosch and Chase, 1981). Selected images were required to depict different, but not atypical, examples of each basic level category. In total, 140 images were selected (35 for each superordinate).

Each image was scanned into an Apple Power Macintosh™ computer at a resolution of 100dpi and was then edited to (i) remove background detail, (ii) convert from colour to 8-bit greyscale and (iii) reduce in size such that each object's maximal dimension fitted exactly within a 50 by 50 pixel grid. Such pre-processing removed relative size differentials between objects and also invoked the assumption that objects are perceived in the central visual field.<sup>2</sup> Although these assumptions may limit the full biological plausibility of the model, they do not diminish its informativeness to the category specific debate since the same assumptions are common to other tests utilising static pictorial stimuli. Some example images of the basic level category 'clock' are displayed in figure 1 and the full set of all 140 images is displayed in appendix A.

*INSERT FIGURE 1 HERE*

In order to ensure that visual properties of our selected images were broadly consistent with those used in other neuropsychological tests, 45 first-year undergraduate psychology students were asked to rate each image for its visual complexity using exactly the same scale and definitions devised by

---

<sup>2</sup> Although it is, in principle, feasible to create an architecture capable of recognising objects when rotated or enlarged, this would require considerable additional pre-processing resources (e.g. Fuchs and Haken, 1988) and is not of concern to our investigations.

Snodgrass and Vanderwart (1980). Visual complexity ratings for items were averaged across subjects and then across all images within each superordinate. The mean ( $\pm$  SD) visual complexity ratings for each category was as follows: animals 3.14 ( $\pm$  0.47); furniture 2.72 ( $\pm$  0.51); musical instruments 2.71 ( $\pm$  0.51); clothing 2.67 ( $\pm$  0.46). A between groups ANOVA demonstrated a significant effect of category ( $F$  [3, 139] = 7.167,  $p$  < 0.0005) and pair-wise comparisons (using Bonferroni) revealed significant differences between animals and all other categories. In the Snodgrass and Vanderwart (1980) corpus a similar distribution of visual complexity is evident in all categories except musical instruments (animals 3.83 ( $\pm$  0.5); furniture 2.81 ( $\pm$  0.69); musical instruments 4.03 ( $\pm$  0.54); clothing 2.71 ( $\pm$  0.64)). The discrepancy in visual complexity between our musical instrument images and those of Snodgrass and Vanderwart (1980) may arise because our stimuli are reduced-size greyscale representations that cannot afford such accurate depiction of fine-grained detail inherent in certain line drawn musical instrument exemplars (e.g. flute). To ensure further that our images were realistic representations of their referent objects we asked 10 subjects to name each picture. Nearly all images were accurately named (mean 97.3%, range 84-100) at the basic or subordinate level except for some of the saxophones and flutes which were mistaken for other brass or woodwind instruments. In recognition terms then, we contend that our image set is not unrepresentative of the type of stimuli used in other studies.

Before any experiments began, two additional control measures were introduced. Firstly, a left-right inversion of each image was created to overcome the fact that some asymmetrical objects (e.g. fishes) tended always to be depicted in the same orientation (e.g. with head to the left and tail to the right), whereas others (e.g. chairs, snakes) had more variable orientations. Secondly, for each left-right inversion, a novel image was created by shifting the distribution of greyscale values (apart from those of the white background) by  $\pm 20\%$ . Images whose originals were subjectively judged to be dark were made lighter, whilst images that were judged to be light were made darker. These manipulations generated 4 versions of each subordinate exemplar, increasing the total image set size to 560. Figure 2 depicts the light/ dark variations and left-right inversions for a single image.

*INSERT FIGURE 2 HERE*

### **The Modular Artificial Neural Network**

The full connectionist model used within this set of studies comprised two modules: an *unsupervised* visual processing module (VPM) and a *supervised* categorisation module (CM). These were trained independently of each other and will be described separately. Broadly speaking, the VPM was designed to perform bottom-up encoding of pictorial stimuli to simulate the level of perceptual categorisation achievable in the absence of lexical-semantic information (e.g. Behl Chadha, 1996; Behl-Chadha, et al., 1995; Eimas and Quinn, 1994; Eimas et al., 1994; Quinn and Eimas, 1986). The CM was designed to perform top-down encoding (supervised object categorisation) of the representations that had developed in the VPM in order to simulate classification (linguistic or otherwise) of each item.

#### ***The Visual Processing Module (VPM)***

The centrepiece of the VPM was based on Kohonen's self organising feature map (see Kohonen, 1982a; 1982b; 1988) - an unsupervised neural network using competitive learning. In Kohonen's original model, pattern classification was achieved by activation of a single 'winning' unit in a two-dimensional array of competing output units. In our *modified* self organising feature map (SOFM), however, pattern classification was distributed across all units in the grid, thereby generating a *contoured* internal representation of each input pattern. The 'winning unit' in this model (i.e. that unit which had the highest level of activation for a given pattern) was still highly important in the representation of an input pattern but, unlike Kohonen's original model, did not have *exclusive* diagnostic capability. The SOFM used within the VPM was structurally and functionally very similar to that described by Schynns (1991) and an overview is provided here (see also figure 3 for an illustration). To clarify the use of terminology at an early stage, the abbreviation VPM denotes the entire visual processing module - which includes the visual input vector, the SOFM and their adjoining weight matrix. The term SOFM refers only to the map of output units.

The VPM simulations in these experiments were characterised by an  $n$ -dimensional input vector ( $A$ ), where  $n$  was the number of pixel values in each training image - in this case 2500 (50 pixels square). The input vector ( $A$ ) was fully connected to a SOFM of 10 units square. In a VPM of this size (i.e. a 10 by 10 unit SOFM connected to a 2500 pixel input vector) there were a quarter of a million connections

with each having its own independent weight value. Each output unit ( $o_i$ ) in the SOFM was connected to the input vector with an  $n$ -dimensional weight vector ( $W_i$ ). Each input vector ( $A_i$ ) comprised 2500 integer values with each value representing 8-bit greyscale information (i.e. the level of 'greyness' between 0 and 255) and spatial location of the corresponding pixel in the input pattern.

*INSERT FIGURE 3 HERE*

At each iteration, the activation values across the SOFM were computed by comparing the euclidean distance between the input vector ( $A$ ) and the weight vector ( $W_i$ ) for each output unit. The output unit with the lowest euclidean distance from the input vector (and hence the highest output value) was regarded as the winner (see appendix B for a formal description). The output map was computed using a transfer function that returned a value in the range of 0 to 255 where 0 represented a unit whose weight vector was far away from the input vector and 255 represented a unit whose weight vector was identical. The training rule updated the weights of the winning unit, and also those of a neighbourhood surrounding the winner, moving them closer to the input vector (see appendix B). The neighbourhood was characterised by a Gaussian function such that weight vectors of units *closer* to the winner tended to approximate the input vector to a greater degree than those further away. It is such correlated zones of activation which facilitate self-organisation (Kohonen, 1982a; Schynns, 1991). As training time increased, the neighbourhood size decreased linearly such that, eventually, only the winning unit's weight vector was updated. Initially, to ensure global order, each unit's neighbourhood was larger than one half of the output map but this was reduced over training time. Similarly, to ensure learning stability, the training rate also declined over training time. In these simulations, VPMs were trained for 1500 epochs with an initial neighbourhood of 7 units square and an initial learning rate of 0.5. Presentation of training set exemplars was randomised within each epoch. The neighbourhood size and learning rate decreased linearly (to 1 and 0 respectively) every 250 and 150 epochs respectively. It will be helpful to the reader to keep in mind that a SOFM surface is wrapped around on itself such that opposite sides and corners are actually in close proximity. For illustration purposes (fig. 3), a SOFM surface is most easily depicted as a 2-dimensional map but, dynamically, it is more akin to the surface of a sphere. In our model, each output unit could take a value between 0 and 255, which made it possible to view the whole surface as an 8-bit (i.e. 256 possible values)

greyscale contour map with regions of high and low activity. Regions of low activity tended more towards values of 0 (black) and regions of high activity more towards values of 255 (white). Two examples of such contour maps are shown in figure 4. Moreover, because each output unit's *weight vector* gradually evolved towards a state which *approximated* a particular input vector, or group of vectors, the *weight vector* for any output unit could be viewed as a 2500 pixel greyscale image. This kind of depiction is informative about the training pattern, or group of training patterns, to which a particular output unit is maximally responsive (see Luckman et al., 1995).

*INSERT FIGURE 4 HERE*

In summary then, the VPM *reduces the dimensionality* of the input vector but *preserves training pattern topology*. Images that are visually similar will, therefore, tend to generate similar contour maps. Preservation of topology means that spatial relationships between areas of a picture and, indeed, between different pictures will be preserved in VPM representations. These characteristics are not incongruent with some properties of structural descriptions although it must be noted that similar VPM representations will arise because items are globally visually similar rather than because they share similar part structures. Moreover, given that the precise representational characteristics of structural descriptions are open to debate and difficult to specify (e.g. Biederman, 1983; Humphreys et al., 1988; Marr, 1982; Marr and Nishihara, 1978), direct comparison between structural descriptions and the perceptual representations in these simulations would be premature.

### ***The Categorisation Module (CM)***

The CM was realised by a multi-layered perceptron with partial recurrence, trained with the back-propagation algorithm (Rumelhart, Hinton and Williams, 1986a; 1986b). A typical recurrent network comprises input, hidden and output layers plus an additional layer of units which 'clean up' the output, creating stable attractor states. This is to say that the euclidean distance between a novel representation and the training set representation to which it is most similar is reduced by a cycle of activity through the clean-up layer. In this way, a novel pattern can fall into the attractor basin of a training set pattern endowing the model with a memory for some, if not all, of the training patterns which it has experienced. A comprehensive description of attractor networks can be found in both



Hinton and Shallice (1991) and Plaut and Shallice (1993). In our model, the size of the CM input layer was determined by the size of SOFM within the preceding VPM. In the majority of experiments this was 100 units (i.e. 10 by 10 output units in the VPM). The CM input layer fed into a layer of hidden units which, in turn, connected to a layer of 32 output units. The output layer had bi-directional connections to a layer of 10 ‘clean-up’ units which allowed attractor states to develop in the output space. The size of the output layer was initially determined by the number of object categories to be encoded using local feature representation (see Hinton, McClelland and Rumelhart, 1986 for a discussion of local vs. distributed representations). For most experiments there were 4 superordinate categories and 28 basic level categories (7 for each superordinate) which necessitated 32 outputs, of which, only 2 should be active for each training item (i.e. one superordinate and one basic level unit). In a later simulation the output layers encoded distributed patterns representing semantic features and microfeatures across the same 32 output units. The exact nature of the output representation is discussed within each experiment. The number of units in the hidden layer was always the minimum necessary for consistent, accurate learning of all training patterns.<sup>3</sup> Each CM layer was fully interconnected with the subsequent layer but there were no intra-layer connections.<sup>4</sup> All CM simulations were run on the t-learn simulator program (Plunkett and Elman, 1997) using a learning rate of 0.01 and a momentum of 0.95.

### ***Model Summary***

To summarise, we implemented a modular architecture comprising a visual processing module (VPM) and a categorisation module (CM). The VPM performs dimensionality reduction with topological preservation on a series of greyscale images and the resulting representations form inputs to the CM which is subsequently trained to perform object classification. Figure 5 depicts the full model showing the interaction between VPM and CM and the various layers within each module.

*INSERT FIGURE 5 HERE*

---

<sup>3</sup> Utilising a minimal hidden layer size is important since it maximises the generalisation performance of the model for the task in which it has been trained (Hinton, 1989).

<sup>4</sup> Intra-layer connections were used in the output layer when target outputs were distributed rather than local (expt. 5)

## THE EXPERIMENTAL INVESTIGATIONS

### Experiment 1

#### *Introduction*

The first experiment explored the representations of visual stimuli which develop in the VPM, with particular focus on whether emerging contour maps provide any evidence for categorical distinctions at a pre-semantic level of representation. At this early stage, the response profile of an output unit was considered only when it was the most highly activated ('winning') unit in the SOFM. Although representations were distributed across the whole SOFM surface, the winning unit is still partially diagnostic since its weight vector tends to approximate the pixel map of one or more exemplars in the training set. The fact that there were only 100 SOFM units and 560 exemplars necessitated that some units would be winners to *more* than one pattern. In any self organising network, the distinctiveness of a training pattern determines whether it will share a similar representation with other patterns. Highly distinctive patterns excite exclusive regions of the SOFM whereas less distinctive patterns tend to excite the same regions (Luckman et al., 1995; Schynns, 1991). This, in turn, determines the probability that a given unit is a 'winning' unit for *more than* one pattern. If natural kind categories are more visually crowded (e.g. Damasio, 1990; Gaffan and Heywood, 1993; Humphreys et al., 1988; Tranel, Logan, Frank and Damasio, 1997) it follows that fewer units should respond *maximally* to animals since there will be greater redundancy of visual information within this class. For non-living things, by contrast, representation should tend to be at an exemplar level because there is a purported lower level of perceptual overlap in these categories. In order to formalise these hypotheses into testable predictions, the following scenarios are proposed: if the level of visual crowding within each superordinate is similar, the expectation is that exemplars, or amalgamations of exemplars, from the 4 superordinates will *each* excite 25% of winning units; however, if the levels of visual crowdedness are not similar between these categories, the expectation is to see unequal proportions of 'winning' units for each category. If there is variation in the VPM's representation of categories, this can be due only to the topography of visual information in the training images since no explicit categorical data is provided in the input.

### ***Method***

Ten VPM simulations were run, each utilising a SOFM of size 10 by 10 units and each being trained with all 560 pictures. Each SOFM started with a random configuration of initial connection weights. After training, the final weight vector (i.e. the 2500 weight values) for each unit in all 10 SOFM's was output as a 50 by 50 pixel image in order to visualise the type of input for which each unit was maximally responsive. The resulting images were then laid out according to the surface grid structure of the SOFM to which they corresponded. Thus a set of 10 by 10 weight vector plots was derived for each VPM. An example of the weight vector images from a VPM simulation can be seen in Figure 6.

*INSERT FIGURE 6 HERE*

It is notable that some images in figure 6 depict clearly identifiable training stimuli (row 2, column 6) whilst others depict amalgamations of stimuli (row 3, column 2). Furthermore, some training exemplars are barely visible, if at all (e.g. there are no clearly discernible frogs in the weight matrix depictions for this VPM despite the fact that it experienced all 20 exemplars of frogs in training). This pattern was not confined only to the VPM in figure 6 but was evident in the other nine also. To test whether such representational biases were consistent, we assessed the response profile of every unit in each SOFM and recorded the type of image which generated maximum activity. In this way the frequency of 'winning units' for each superordinate category could be assessed for each VPM.

### ***Results***

Table 1 displays, for each VPM, the number of winning units for each superordinate. The different VPM simulations produced highly consistent frequency distributions.

*INSERT TABLE 1 HERE*

Furniture and clothing exemplars were depicted more frequently amongst the weight vector plots than animals and musical instruments (repeated measures ANOVA  $F [3, 39] = 408, p < 0.0001$ , with post-hoc, Bonferroni corrected, Scheffe  $F$  tests demonstrating significance for all pair-wise comparisons except animals vs. musical instruments). Although living things were, as predicted by

visual crowding theory, less likely to cause maximal activation in a high number of units, the same was also true for musical instruments. So, purely with respect to winning unit activity, some form of category bias is evident which can arise only on the basis of visual information.

### ***Discussion***

These data cannot be accounted for by any of the so-called 'artefact' variables which predict recognition accuracy in humans: item familiarity carries no explanatory power here since each training pattern was presented an equal number of times; visual complexity can neither explain the pattern of results because it was only rated as being significantly higher within the category of animal stimuli (the ratings obtained for musical instruments, furniture and clothing were all similar). That animals and musical instruments should generate fewer states of maximal activation offers some preliminary evidence of greater visual redundancy within these categories. However, whilst the structural description hypothesis would presumably define visual redundancy in terms of shared parts and part configurations, the redundancy in this model is attributable to global visual similarity (e.g. shape, spatial location and distribution of shading information), since the model has no specific mechanisms for decomposing pictured objects into constituent parts.

## **Experiment 2**

### ***Introduction***

So far, the response profile of a given unit has been examined only when it is maximally active. However, given the distributed nature of representation in the VPM, it is pertinent to consider the additional information that may be available from studying relationships between *different* feature map units for the *same* training exemplar. For example, units may still carry some diagnosticity even when they are inactive or only partially active. In the VPM *every* SOFM unit plays a role in the representation of *every* training set exemplar, even if only at low levels of activation. Figure 7, where light and dark areas indicate high and low regions of feature map activity respectively, should illustrate this point. These contour maps fit within the same grid pattern as the images in figure 6 since they derive from the same SOFM.

*INSERT FIGURE 7 HERE*

The 10 by 10 grid structure in the 2 diagrams (figures 6 and 7) allows superimposition and, matching the lightest region on each contour map with the corresponding weight image, provides a description of the type of image which the most active unit (i.e. the whitest) responds to. Note how the activation peaks for the 2 furniture items in figure 7 (row 2, column 3 and row 7, column 9) correspond with units whose weight vectors depict unambiguous examples of lamps and clocks. Now contrast this with the 2 animals and note how the highest regions of activity correspond to units whose weight vectors depict ambiguous figures. These examples are not atypical and begin to illustrate some of the qualitative representational differences which emerge between different object categories in the VPM. When examining feature map representations for animals and furniture, a trend was observed towards the former being represented by undulating contour maps (i.e. relatively shallow peaks and troughs) and the latter being represented by much harsher contours (i.e. sharper and higher peaks). To establish whether this trend was reliable, an assessment of the *distribution of activation values* was made for each image for all 10 VPMs.

### ***Method***

Activation values (range 0-255) were recorded for all 100 SOFM units in each VPM, for every training exemplar. For diagrammatical and analytical purposes, the possible activation values (i.e. values in the range 0 to 255) were grouped into 5 ordinal bands as follows: (i) 0-49, (ii) 50-99, (iii) 100-149, (iv) 150-199, (v) 200-255. The frequencies of unit activation values falling within each band were recorded for each training exemplar and averaged across all 10 VPMs. These mean frequencies were then averaged across all exemplars *within* each superordinate.

### ***Results***

The mean frequencies for each unit activation band are plotted in figure 8 with respect to each superordinate. VPM performance was highly consistent, indicated by the low standard deviations.

*INSERT FIGURE 8 HERE*

It is evident that most images generate contour maps comprising a majority of minimally activated units. This derives largely from the learning algorithm used, in which the size of a unit's associative neighbourhood is reduced during the training cycle such that, eventually, only each winning unit's weights are updated. However, the distribution of unit activation values was not consistent across superordinate categories, with significant differences being observed within each frequency band (one-factor ANOVAs [df 3, 559] for the 0-49 band,  $F = 48.2$ ,  $P < 0.0001$ ; for the 50-99 band,  $F = 46.28$ ,  $p < 0.0001$ ; for the 100-149 band,  $F = 43.41$ ,  $p < 0.0001$ ; for the 150-199 band,  $F = 2.72$ ,  $p < 0.05$ ; for the 200-255 band,  $F = 6.19$ ,  $p < 0.0005$ ). Post-hoc pairwise comparisons (Bonferroni) revealed the following differences: 0-49 band - animals vs. all others, musical instruments vs. all others; 50-99 band - animals vs. all others; 100-149 band - animals and musical instruments vs. clothing and furniture; 200-255 band - furniture vs. all others.

### ***Discussion***

Experiment 2 confirms that category differences exist not just in the response profile of winning units but also across distributed VPM representations. Whilst some images generated unique contour maps with fairly sharp peaks and a residual surface of marginally active units (i.e. furniture and clothing), other images generated very similar contour maps with plateaux of moderate to high activity and less tendency for localised peaks (i.e. animals and musical instruments). The former tended to be represented as unique exemplars and the latter as exemplar amalgamations or abstractions. This pattern of representation may offer an alternative conceptualisation of hierarchically organised perceptual representations. For example, Sartori et al. (1993) proposed that living things have structural representations which are deeper, whereby their item-specific representations reside further down the hierarchy relative to non-living things. Representation of exemplars and categories within the VPM suggest this hypothesis might be better articulated in terms of the extent of overlap between patterns in a neural network. Within such a framework, certain categories can emerge naturally because their exemplars are more tightly clustered together. Others have little perceptual basis for categorical coherence, resulting in discernible exemplar representations. Differentiation of exemplars in the former would necessitate more intricate, and presumably time-consuming, representational interrogation whereas differentiation in the latter might require only a 'shallow' comparison between

representations. Moreover, local neurological damage (within an area of the map critical for representing visually crowded categories) may selectively impair recognition of living things and musical instruments by removing the subtle exemplar differences previously encoded within the damaged or missing units. By contrast, localised damage in other areas of the map would not have such a profound effect because fewer exemplars would be affected.

Although categorical inequalities were observed in the VPM representations, it remains untested whether these would modulate accuracy of full object recognition (i.e. object identification/ naming). For example, although we suggest that differentiation between living things requires a more intricate interrogation of perceptual representations, it does not necessarily follow these items will ultimately be more difficult to identify. Indeed, given that human object naming purportedly favours a more generic level (e.g. ‘bird’ rather than ‘sparrow’ (Rosch et al., 1976)), it is plausible that a high degree of overlap between item representations in certain categories might serve to *enhance* object identification. In experiment 3 we investigate whether categorical inequalities at the level of perceptual representation can impact upon basic level identification.

### **Experiment 3**

#### ***Introduction***

The previous experiments have demonstrated how formation of perceptual categories may be moderated by global visual similarity and how certain groups of items may gain greater categorical coherence at a pre-semantic stage of processing. Following assumptions of the cascade model (Humphreys et al., 1988), one should also expect the level of perceptual overlap in a group of items to be reflected at later stages of object recognition because competing perceptual representations should each activate their own semantic representations which should then compete with the target semantic output. To explore this idea in a connectionist framework, we used the representations which had previously emerged in the VPM as a set of input vectors for a supervised network trained to allocate each perceptual representation to a basic level name category. By measuring the learning time taken to categorise each VPM representation correctly, it should be possible to assess whether the nature of

pre-semantic representation has a bearing on actual object classification.

### ***Method***

The CM was used to simulate basic level naming. In the recurrent part of the CM one cycle was implemented as this was sufficient to elicit an unambiguous output state (cf. Hinton and Shallice, 1991). Cross-entropy<sup>5</sup> was instantiated as the measure of unit error during training although, for the purpose of data presentation, all error terms are expressed in the standard form of root mean square (RMS) error. Ten individual CM simulations were run, each using output representations of a different one of the 10 VPMs described in experiments 1 and 2. For training purposes the output values from each VPM were standardised within the range 0-1 and the 560 patterns were presented in random order across the 100 unit input layer of each CM. The subsequent hidden layer, output layer and clean-up layer comprised 15, 32 and 10 units respectively. Each of the 32 output units represented a distinct name category (i.e. target categories were locally represented), with the first 4 representing superordinates and the remaining 28 representing basic level names. Thus, for each input, 2 output units were simultaneously active at any one time. This method of classification is not intended to simulate the complex processes involved in object naming and nor is it proposed to be neurally plausible. Nonetheless, it does provide an unambiguous measure of the model's ability to separate each input pattern into a coherent categorical grouping (Allison, Ellis, Flude and Luckman, 1992; Humphreys et al., 1995; Luckman et al., 1995; Quinn and Johnson, 1997) and, moreover, ensures that the target output representation for each input is qualitatively similar, thereby eliminating confounding factors at the level of output representation. For the purposes of this investigation, the error term for each training pattern was measured across the entire output layer at intervals of 10 epochs (i.e. every 5600 pattern presentations) and training was aborted when the RMS error for *every* pattern in the training set had dropped below 0.01. For analysis, RMS errors for each training exemplar were averaged across all 10 CM simulations at each 10 epoch interval and, for graphical representation, these mean error terms were averaged across all items within each superordinate, providing a mean RMS error for each superordinate category at every 10 epochs.

---

<sup>5</sup> RMS can be ineffective for overcoming local minima when target patterns are large and sparse because units may become pinned incorrectly due to momentum. Cross-entropy overcomes this problem by increasing exponentially for completely incorrect outputs.



## **Results**

Figures 9a, 9b and 9c display the decrease in average RMS error for pattern classification in each superordinate category during early, middle and late stages of training respectively.

*INSERT FIGURE 9 HERE*

Animals and musical instruments generally produced a higher error rate than furniture and clothing. Note how a marked difference emerges between these 2 categorical groupings soon after training commences (figure 9a) and continues (figure 9b) until a late stage of training, when error rates for all 4 categories converge (figure 9c). Repeated-measures ANOVAs showed significant error differences between superordinate groups ( $F [3, 139] > 3.51, p < 0.05$ ) at *each* 10 epoch interval up until 480 epochs. Post-hoc pairwise comparisons (Bonferroni) demonstrated the following significant error differences at epoch intervals 100 through to 300 inclusive: (i) animals vs. furniture *and* clothing; (ii) musical instruments vs. furniture *and* clothing. These data show that animal and musical instrument VPM patterns are more difficult for the CM to learn.

## **Discussion**

The category dissociations observed in learning correlate broadly with those found in experiments 1 and 2. Furniture and clothing items, which generated the most distinctive contour maps, were more easily classified by the CM. In some respects this is surprising because target outputs for the CM represented basic and superordinate categories and were, therefore, *more generic* than the level of abstraction that was possible in the VPM representations. It follows that the CM must re-represent some of the more atypical VPM representations, forcing them to converge within basic and superordinate groups. Given that animal and musical instruments exhibited greater tendency towards coarse-coding in the VPM, one might expect the CM to be more adept at assigning them to their appropriate categories. The fact that this did not happen suggests that the level of categorical coherence seen across some animal (and musical instrument) VPM representations may actually be *more inclusive* than the basic level. If this is the case, the biggest problem for the CM may well have been one of *disambiguating* basic level categories due to perceptual crowding at a more global level rather than grouping together unique exemplars at the basic level. This interpretation does not accord

with the notion that all basic level categories reflect maximal within category similarity and maximal between category dissimilarity (Rosch et al., 1976). Indeed, our data suggest that, for animals and musical instruments, these parameters might be better fulfilled by a level of categorisation more *inclusive* than the basic level. Conversely, for most artefacts the optimal level of categorisation may be more *exclusive* than the basic level.<sup>6</sup> The data presented here are consistent with the primate learning study of Gaffan and Heywood (1993) whose subjects were trained to discriminate between pairs of basic level objects and animals selected from Snodgrass and Vanderwart (1980). The monkeys learned more slowly to discriminate basic level living things from each other and the authors concluded that visual discriminability must be particularly difficult within living thing taxonomies. The implication of these findings, and those of experiment 3, is that living things are perceptually crowded not just within basic level groups but also at a more global level. Gaffan and Heywood did not discuss primate discrimination of musical instruments but they did report, in another experiment utilising the same stimulus set, that musical instruments and living things suffer from similar visual processing constraints when degraded line drawings are presented to neurologically intact human subjects<sup>7</sup>.

Experiment 3 demonstrates how pre-semantic category bias may give rise to category specificity in full object recognition. These data concur with the cascade model which predicts slower recognition for structurally similar stimuli. However, whilst the pre-semantic module of cascade (i.e. the structural description system) is assumed to represent similarity by shared parts and part configurations, the VPM abstracts a less processed, more global type of similarity (e.g. spatial location, global shape and the distribution of shading information across an object's surface). Whilst such global representations may capture *some* information about shared parts and configurations, it would be wrong to imply that objects are explicitly specified by constituent parts within the VPM.

---

<sup>6</sup> However, the simulations reported here only consider global visual similarity and do not take into account functional or tactile similarity which may also play a critical role in the natural coherence of basic level groups.

<sup>7</sup> Although their comparison was only between musical instruments and tools.

## Experiment 4

### *Introduction*

In experiment 3 the CM was *slower* to learn correct classifications for animals and musical instruments. Nonetheless, it was able to learn correct classifications for all training patterns if given sufficient time (i.e. > 480 training epochs). Although increased learning time is required to differentiate between exemplars from perceptually crowded categories, it is not necessarily the case that these items will be more prone to error once learning has reached saturation (i.e. when error rates converge at a low levels for all items). To test this hypothesis in a supervised network like the CM is not straightforward because, when fully trained, it will be 100% accurate for all familiar patterns (i.e. those experienced during training) giving rise to ceiling effects. However, given that neural networks can generalise learned mappings *beyond* training items, it is viable to test them with unfamiliar stimuli. Within the context of the current study, a generalisation task using novel stimuli would serve 2 purposes. Firstly it would give some indication of whether pre-semantic categorical differences can explain error rates when testing occurs under less than optimal conditions (i.e. when an easy task is made more difficult by imposing constraints – in this case testing with unfamiliar rather than familiar items). Secondly, it would test whether the model can apply learned mappings from one set of images to another. This is important because if a model fails to generalise, it may be dismissed as an artefact.

In this experiment we explored the emergence of category effects in the full model (i.e. the VPM and CM combined) after presenting it with only 80% of the pictorial stimuli. At testing phase the model was presented with the remaining 20% and its ability to generalise (as measured by classification accuracy) was assessed in relation to the different object categories. The target item classifications in the CM were the same as in experiment 3 (i.e. superordinate and basic). Given the findings of the previous experiments, we can make some tentative predictions about how the model should perform. We know that animal patterns tend to be tightly clustered together in the VPM such that even some basic level animal categories share similar representations. This should make superordinate classification quite easy for the model although it may have difficulty in disambiguating some basic level categories. A broadly similar pattern of performance would be expected for the category of musical instruments. Clothing and furniture exemplars, on the other hand, have very distinctive VPM

representations which do not cohere naturally at either the superordinate or even basic level; although their appropriate classifications were easier for the CM to learn in experiment 3, this may have been because there was much less chance of confusing similar items. Indeed, given the low levels of visual overlap in these categories, it is possible that learned mappings are of little use in categorising *novel* exemplars. For this reason, the model is expected to generalise more effectively with visually crowded categories (i.e. animals and musical instruments).

### **Method**

*Creation of VPM Training and Test Sets:* The set of 560 images was systematically subdivided in the ratio of 4:1 to create training and test sets (448:112 images respectively). Training sets included 4 of the 5 subordinates representing each basic level category and test sets included the remaining subordinate. Five training/ test set combinations were created such that each subordinate appeared in one of the 5 different test sets. The presence of a subordinate, whether in the training or test set, was characterised by all light/ dark variations and left-right inversions (i.e. 4 instantiations of each subordinate). In this way, training/ test set construction was counterbalanced across superordinate, basic and subordinate categories such that all images appeared the same number of times.

*Training:* One randomly configured VPM simulation was run for each different training/ test set combination (i.e. 5 in total). When training was complete the 448 output vectors for each VPM were transformed within the range of real values between 0 and 1. These transformed vectors formed the training set for the CM (the CMs were identical in specification to those described in expt. 3). Six randomly configured CMs were trained with each of the 5 sets of transformed VPM output vectors (i.e. 30 unique CM simulations in total). Presentation of training patterns was randomised within each epoch and training was aborted when the RMS error for *every* pattern had dropped below 0.01.

*Testing:* The appropriate test set (i.e. the set of images that was mutually exclusive to the training set) was presented to each of the 5 fully trained VPMs and the resulting contour maps (i.e. the activation values of all 100 units of the SOFM) were recorded for each image. Activation values in each contour map were standardised within the range 0-1 and, in this way, formed the *novel* test sets for the CMs.

These test sets (each containing 112 patterns) were then presented to the appropriate CMs. Activation values were recorded across the CM output layer for each item and scored for their level of accuracy.

*Scoring:* For scoring purposes, superordinate and basic level classifications were treated independently. For a correct superordinate generalisation the appropriate CM output unit (1 of 4) had to be activated to a level of at least 0.49 (min 0, max 1) and no other competing 'superordinate' units could be more highly active than the correct unit. Furthermore, unless a competitor was activated to a value that was at least 0.1 less than the target unit, classification was scored as incorrect. In this way 'proximity' and 'gap' criteria were set for measuring model performance (see Hinton and Shallice, 1991; Plaut and Shallice, 1993). For correct basic level generalisation the appropriate basic level output unit (1 of 28) had to be activated to a level of at least 0.49. The procedure for dealing with competing basic level units (i.e. gap criteria) was identical to that for superordinate units. The type of basic level classification error for each pattern was also recorded. If a different basic level unit from within the same superordinate was activated to a higher level than the target unit, the response was classed as a 'within-category error'. This was also the case if a competing basic level unit from within the same superordinate was activated to within 0.1 of the target unit. An error subtype of 'unrelated category error' was recorded if the most active unit (provided that its activity level exceeded 0.49) was a 'basic level' unit belonging to a different superordinate. These error classifications are often used in cognitive neuropsychological assessments of visual object recognition (e.g. Done and Gale, 1997; Hodges, Salmon and Butters, 1991; 1992). All other erroneous classifications were treated as non-responses including any instances where no unit was activated above the imposed threshold value of 0.49. The number of correct and incorrect classifications was recorded for all 30 CMs, and broken down with respect to the superordinate categories. Comparisons of generalisation performance were drawn, firstly, between levels of abstraction (superordinate vs. basic) and, secondly, between superordinate categories (animals vs. musical instruments vs. clothing vs. furniture).

### ***Results***

Collapsing across all categories, generalisation accuracy far exceeded the levels predicted by chance alone (Mean  $\pm$  S.D. accuracies for superordinate and basic level generalisation were 68.7%  $\pm$  5.7% and

42.8%  $\pm$  5.1% respectively). Chance accuracy would be 25% (i.e. 1 in 4 units) for superordinate generalisation compared with only 3.57% (1 in 28 units) for basic level generalisation. With chance performance partialled out, superordinate generalisation was still superior to basic level generalisation ( $t [1, 29] = 3.279, p < 0.005$ ). Category effects in generalisation accuracy were observed at both superordinate ( $F [3, 87] = 20.99, p < 0.0001$ ) and basic levels ( $F [3, 87] = 67.62, p < 0.0001$ ). At the superordinate level, accuracy for furniture items was consistently lower compared to other categories, whilst accuracy for animals exceeded that of all other categories except musical instruments (Bonferroni corrected Scheffe F-tests,  $p < 0.008$ ). At basic level, accuracy for furniture was consistently lower than in all other categories ( $p < 0.008$ ). Figure 10 displays the mean percentages of novel stimuli that were correctly identified at both basic and superordinate levels. The scores predicted by chance occurrence are also provided for comparative purposes. The type and frequency of basic level generalisation errors are plotted in figure 11. The frequency of each error type is expressed as a percentage of all basic level generalisation errors made within each superordinate.

*INSERT FIGURES 10 AND 11 HERE*

Non-responses were the most frequent error type for each class simply because the level of activation across basic level outputs often failed to reach the imposed threshold. The proportion of non-responses varied across categories ( $F [3, 87] = 6.81, p < 0.0005$ ) with both animals and musical instruments generating significantly fewer than other categories ( $p < 0.008$ ). The proportions of within-category errors also varied across categories ( $F [3, 87] = 21.98, p < 0.0001$ ) with (i) animals generating more than all other categories *except* musical instruments and (ii) clothing generating fewer than any other categories ( $p < 0.008$ ). It is of particular interest that, in many studies of object naming, these errors would be categorised as 'semantic'. In these simulations, however, such confusions can only be visual since input representations did not encode explicit 'semantic' features.

### ***Discussion***

This experiment has demonstrated a different pattern of category deficit to that reported in experiment 3. As predicted, superordinate categories of novel animals and musical instruments were

more easily recognised compared to the other categories. This further supports the view that these categories are crowded at a global level and a prediction is that category decision in normal subjects (i.e. forced choice of superordinate labels) should be quicker for items in these classes. Basic level classification of animals and musical instruments was less accurate (relative to superordinate classification) suggesting that boundaries of some basic level *perceptual* representations within these categories may be fuzzy. Nonetheless, the model was considerably more accurate at classifying basic level animals and musical instruments relative to furniture and, considering that chance performance is only 3.6%, the model has arguably captured useful diagnostic information about object categories. We did not predict a difference in classification accuracy between clothing and furniture since these categories elicited similar patterns of learning performance in experiment 3. However, novel clothing items were recognised more accurately at both levels of abstraction compared with furniture. This pattern cannot derive from item effects because the presentation of exemplars was counterbalanced so that each item was ‘novel’ in the same number of experiments. In experiment 1, items of furniture accounted for a higher number of ‘winning units’ than items of clothing (34.5% and 28.3% respectively), suggesting greater visual redundancy in the latter. Intuitively, one might expect basic level clothing categories to carry more redundant visual information because their general shape characteristics are determined by those of basic level human body parts (e.g. trousers/ shorts - *legs*, gloves - *hand*, shoe/ boot - *feet*, t-shirt/ jacket - *torso*) which tend to be shape invariant. However, whilst the shape of some furniture items is partly constrained by human body shape, different functional specifications result in considerable variation within basic level categories (e.g. *chairs*: armchair, deckchair, dining chair, office chair) and some items of furniture have no relationship with human body shape whatsoever (e.g. lamp, clock). That basic level non-living categories can have low between-item similarity and high within-item variability has also been pointed out by Laws and Neve (1999) and Turnbull and Laws (2000). The data from this experiment (and indeed experiment 1) concur with this although the extent of visual redundancy would vary between categories (i.e. some artefact categories would have greater visual redundancy than others).

Again we would point out that no semantic features have been encoded in any simulations so far. One might argue that the weight matrices of the CM perform some kind of semantic function on grounds

that they facilitate assemblance of certain object groupings (e.g. ‘pulling together’ visually diverse items such as furniture) and disambiguation of others (e.g. enhancing the fine-grained distinctions between basic level animal categories). However, such properties derive from imposition of target output states for the model rather than explicit encoding of semantic attributes. Although the proposed model cannot be viewed as a full model of object recognition (since there is no semantic system to mediate between perceptual representations and object classifications), it does give some indication of the level of object identification that may be achievable from visual information alone.

An important, and indeed consistent, finding of the simulations so far is the similar performance profile for animal and musical instrument stimuli which accords with many cases described in the literature (e.g. Stewart, Parkin and Hunkin, 1992; Warrington and Shallice, 1984). This apparent anomaly has been difficult to explain in terms of relative deficits for living things or biological entities (Parkin and Stewart, 1993). It has been pointed out that both living things and musical instruments are relatively unfamiliar *and* visually complex categories. Whilst these assertions are undoubtedly true, they cannot account for our data. Our results suggest, rather, that similar performance characteristics for these classes arise through high levels of within-category visual overlap. We would certainly not contend that this explanation accounts for all living thing/ musical instrument category-specific impairments. Indeed, we have not considered the role of the semantic system (or systems) in object classification/ naming, so we are unable to rule out explanations at the level of semantic knowledge (e.g. Caramazza et al., 1990; Farah and McClelland, 1991; Warrington and Shallice, 1984). Nonetheless, our model does show that perceptual crowding can be a potentially important factor in at least some cases of category specific impairment. In the next experiment we test this hypothesis in a larger model which incorporates a semantic layer within the CM. Unlike the previous experiments which have investigated model performance under ‘normal’ conditions, experiment 5 focuses on the model after it had been artificially lesioned. From this we may predict the kind of errors that might be expected following damage at different stages of the visual object recognition process.

## **Experiment 5**

### ***Introduction***



In experiments 1 to 4 we have demonstrated how perceptual crowding might underlie category specific impairments. We have also demonstrated the level of object identification that is achievable for our image set *without* the use of semantic information. We have not yet considered the potential modulatory effects of semantics on representations which have evolved in the VPM. Most category specific impairments discussed in the literature have been described in terms of object naming/ recognition accuracy and, under the assumptions of contemporary models of object recognition, such impairments must reflect contributions of both perceptual and semantic representations. It is difficult to directly probe perceptual representations in human subjects without activating some semantic information, due to the non-discrete nature of visual object processing (Humphreys et al., 1988). Thus, if there is any validity in the claim that perceptual crowding can underlie some category impairments, it is important to show that visual redundancy effects are still present at later stages of object processing. In this experiment, a new variation of the model is introduced. Rather than simply mapping VPM representations onto locally represented object categories (as in expts 3 and 4), this model is intended to capture, more plausibly, the mapping process between perceptual and semantic representations. The function of the VPM is identical to before but, in this experiment, the CM is trained to assign VPM representations to a semantic representation which comprises *distributed* features and micro-features. Once trained, the model is lesioned incrementally at different processing stages (early vs. mid vs. late) and the impact on recognition performance for different object categories is measured.

### ***Method***

The outputs of the 10 VPMs trained in experiments 1 and 2 were used as training sets for the CMs in this experiment (i.e. the 560 output vectors from *each* VPM formed the set of training vectors for one distinct CM, resulting in 10 unique CM simulations). Each CM had 100 input, 50 hidden and 32 output units with each layer being fully connected to the previous. The number of hidden units was increased in this study because distributed output representations were used instead of local representations. One impact of this is a greater overlap between output patterns, rendering learning more difficult and necessitating an increase in processing resources. However, the hidden layer was still of minimum size necessary for consistent, accurate learning of all patterns. Unlike the previous

experiments there was no clean-up layer in this model. However, all units in the output layer were fully interconnected to allow mutual excitation or inhibition. Implementing a different form of recurrence in this experiment served two purposes. Firstly, initial investigations had demonstrated that a large clean-up layer was needed to cope with the new set of distributed output representations (Gale, 1997). This prolonged model training considerably and did not lead to more accurate learning than intra-layer connections. Secondly, when a large clean-up layer was implemented, it only accounted for a very small change in the output representations of most training patterns (i.e. the majority of the categorisation was done by the feed-forward part of the network) rendering this method an uneconomic use of processing units.

As with experiments 3 and 4, the output units represented superordinate and basic level category information. However, in this experiment, semantic categories were delineated by pools of micro-features rather than localised category markers. The set of possible output states comprised 28 different distributed binary patterns (i.e. one for each of the 28 basic level categories). Each distributed pattern represented both superordinate and basic level information. Representations for basic level categories sharing the same superordinate incorporated the same superordinate information (table 2). Each of the 28 possible outputs was associated with 20 distinct VPM representations (i.e. the 5 subordinates of each basic level group represented by left-right inversions at 2 levels of light/ dark). Each possible output pattern comprised 32 microfeatures (i.e. bits), of which, 16 were involved in the representation of superordinate properties and 16 in the representation of basic level properties. Each output pattern comprised an equal number of 1s and 0s and, across all 28 possible output patterns, each *unit* was set to 1 and 0 an equal number of times. Thus, no output unit was biased towards assuming a particular binary value (i.e. the probability of any given unit being active was 0.5). Moreover, each output unit was involved in the representation of superordinate and basic level properties an equal number of times, removing any division of labour between units representing general and specific information (following Tippett, McAuliffe and Farah, 1995). For the reader's clarification, 8 examples of semantic outputs are displayed in table 2.

*INSERT TABLE 2 HERE*

Such great care was taken to counterbalance all sources of output representation variation so as not to introduce any theoretical assumptions about the representation of semantics for different object categories. There was no distinction between perceptual and functional semantic units (cf. Farah and McClelland, 1991) and nor were there differing levels of inter-unit correlation for different categories (cf. Devlin et al., 1998). This is not because we believe that such assumptions are necessarily flawed. However, it is difficult to justify them on a neural basis given what little is known about the fundamental nature of semantic representation in the brain. For example, whilst it is undoubtedly true that living and non-living things tend to be *defined* more in terms of perceptual and functional attributes respectively, it does not logically follow that these attributes form the *basis* for such concepts (e.g. Keil, 1989; Komatsu, 1992; Medin, 1989; Murphy and Medin, 1985). Thus, in our model, we have avoided any assumptions about representational *inequalities* at the level of semantics and our investigation will focus purely upon the mapping process between VPM and distributed semantic representations in order to ascertain whether perceptual crowding effects are still apparent in a *neutral* semantic system.

The CMs were trained using RMS error as a measure of unit discrepancy. Training patterns were presented randomly without replacement within each epoch and training was aborted when the RMS error for *every* pattern had dropped below 0.01. When training was complete, each of the 10 CMs was systematically lesioned, by random connection removal, at 3 distinct locations, each under 3 levels of severity. The levels of lesion severity were set at 5%, 10%, 20% removal of connections and lesions were made at the following locations: (i) between input and hidden units (IH); (ii) between hidden and output units (HO); (iii) between units within the semantic output layer (OO). Each CM was tested with the full set of 560 patterns at baseline (i.e. no lesioning) and then under each lesion severity level at each of the 3 sites (i.e. 10 tests for each CM in total). Errors across output units were expressed as mean unit output discrepancies, whereby the activation value of each unit was compared with its target state for a given pattern and the modulus of the discrepancy was averaged, within each test pattern, across (i) those units involved in the representation of superordinate information and (ii) those units involved in the representation of basic level information. These mean unit output discrepancies were then averaged across the 4 instantiations of each subordinate image (i.e. the 2

left/ right inversions and 2 light/ dark variations), resulting in 140 superordinate and 140 basic level mean unit output discrepancies for each CM simulation. These values were then averaged across all 10 CM simulations for each item thereby facilitating a by-items analysis (35 items for each superordinate) with the basic level and superordinate error score for each item being synthesised from 640 individual unit measures (i.e. 16 unit output discrepancies \* 4 instantiations \* 10 unique CM simulations). These final error values were analysed within a 4-factor analysis of variance model. Level of visual crowdedness (high vs. moderate to low) was a between items factor with animals and musical instruments being pitted against furniture and clothing. Lesion severity level (5% vs. 10% vs. 20%), lesion site (IH vs. HO vs. OO) and level of semantic classification (superordinate vs. basic) were nested within-items factors.

If visual crowding effects are still observable at post-perceptual stages of visual object recognition, it must be predicted that items from visually crowded categories (i.e. animals and musical instruments) will generate higher mean unit output discrepancies after simulated lesioning. Furthermore, given the greater activation frequency of superordinate information (i.e. superordinate feature patterns were activated 7 times more often as basic level feature patterns), we would also have to predict a *relative* overall preservation of superordinate information. However, based upon the findings in experiment 3, it would be reasonable to expect an interaction between level of semantic categorisation and visual crowdedness since disambiguation of basic level categories appeared to be more difficult for items belonging to visually crowded groups. Thus, we would expect highest unit output discrepancies for units involved in representing basic level information for visually crowded categories. Naturally, one would expect the level of lesioning severity to predict global error rates but there is no a-priori basis for assuming a *qualitatively* differential impact upon visually crowded and non-crowded categories. Finally, the site of lesioning should influence propensity for error because each layer of the CM will play a different role in representing training items: at input level representation, visual similarity effects will be of a similar magnitude to those reported in the previous experiments because the CM input layer simply encodes the VPM output; the hidden layer, however, will need to re-represent the similarity relationships between training items in order to facilitate correct semantic classification at output (given that no categorical biases are encoded in the semantic output representations); finally,

the interconnections between semantic units would be expected to act as a fine-tuning mechanism for reducing RMS error. Given the nature of representational change that is likely in the model, we would predict a large visual crowding effect for early (i.e. IH) lesions but would expect this effect to be attenuated (though not lost completely) after later lesions (i.e. HO and OO).

### **Results**

The mean unit output discrepancies after lesioning are plotted in figure 12 (min 0, max 1). The 3 levels of lesion severity (5%, 10% and 20%) are plotted on the x axis for each lesion types (IH, HO and OO). Mean unit output discrepancies for items from highly crowded categories (i.e. animals and musical instruments) are indicated by *black* and *white* bars (for units involved in *superordinate* and *basic* level representation respectively). Mean unit output discrepancies for items from moderate to low crowded categories (i.e. clothing and furniture) are indicated by *dark grey* and *light grey* (for units involved in *superordinate* and *basic* level representation respectively).

As predicted, all main effects were significant. Visual crowdedness was a significant factor ( $F[1, 139] = 25.8, p < 0.0001$ ), with crowded categories generating higher discrepancies than uncrowded categories (overall means 0.155 and 0.113 respectively). Level of semantic classification was also a highly significant predictor ( $F[1,139] = 73.0, p < 0.0001$ ), with basic level information being more vulnerable to lesioning (overall mean superordinate and basic level unit output discrepancies 0.109 and 0.159 respectively). As would be expected in any neural model, the extent of lesion severity was highly predictive of classification accuracy ( $F[2, 278] = 1347, p < 0.0001$ ), with overall mean unit output discrepancies of 0.046, 0.132 and 0.224 for 5%, 10% and 20% levels of lesion severity respectively. Lesion site was also a significant factor ( $F[2, 278] = 226.8, p < 0.0001$ ), with earlier lesions always generating a higher error rate than later lesions (overall mean unit output discrepancies were 0.215, 0.114 and 0.074 for IH, HO and OO lesions respectively).

*INSERT FIGURE 12 HERE*

There was an interaction between level of semantic classification and visual crowding ( $F[1, 139] = 48.5, p < 0.0001$ ); for visually uncrowded categories, the mean overall level of unit output discrepancy was

similar for superordinate and basic level units (0.111 and 0.109 respectively); for visually crowded categories, however, the mean unit output discrepancy was much greater for units involved in the representation of basic level information (0.20 vs. 0.11). So, although visual crowding effects are still observable at the post-perceptual stages of our model, these effects only seem to be tapped by tasks which rely on basic level semantic knowledge. There was also an interaction between lesion site and visual crowdedness, whereby early lesions produced the largest effects of visual crowding. The ratio of crowded: uncrowded mean unit output discrepancies were as follows: for IH lesions 0.265 vs. 0.164; for HO lesions 0.121 vs. 0.110; for OO lesions 0.082 vs. 0.066. Post-hoc tests confirmed that the differences were significant only for IH and OO lesions ( $p \leq 0.005$ ). The fact that HO lesions did not generate visual crowding effects would suggest that the hidden units were compensating for the differential effects of visual similarity, thereby re-representing visually crowded items in a way which made them more discernible from each other and, moreover, re-representing visually diverse items in a way which made them more categorically coherent. However, given that a visual crowding effect was still observable after OO lesions, it would appear that connections between units within the semantic later played a greater role in generating correct activation states for visually crowded items, even though the effect size was relatively small. Finally, although there was no effect of visual crowding after HO lesions, there was an interaction between level of semantic categorisation, visual crowding and lesion site ( $F[2,276] = 12.4, p < 0.0001$ ) with post-hoc tests revealing a significantly higher mean unit output discrepancy for basic level information within crowded categories after IH, HO and OO lesions.

### ***Discussion***

In this experiment, the model was required to map between VPM representations and distributed semantic representations. In the former, qualitative differences between categories had evolved (see experiments 1 and 2) but, in the latter, categorical biases were deliberately avoided in order to eliminate theoretical assumptions about the organisation of semantic memory which might account for category effects. Although the semantic representations we adopted clearly embody *some* assumptions (for example, we have made the assumption that different items are represented by

equivalent levels of superordinate and basic level information), none of these would predispose any particular category of items towards a higher error rate after lesioning. Whilst it is contentious whether this reflects the true nature of representation in semantic memory, our aim was simply to investigate whether visual crowding effects can be propagated forward in a 'neutral' semantic system. Our data support the view that visual crowding can underlie category impairments even when lesioning occurs at post-perceptual levels of processing. However, as the site of lesion damage moves from perceptual towards semantic representations, the effects are reduced considerably in our model. Nonetheless, there is a consistent trend for relative loss of basic level semantic information for items belonging to visually crowded categories, suggesting that representational inequality at a pre-semantic stage of visual processing has some potential in accounting for loss of more detailed semantic information for living things and musical instruments. This explanation is not mutually exclusive to other explanations which posit between-category differences in the nature of semantic representation (e.g. Durrant-Peatfield et al., 1997; Farah and McClelland, 1991). Indeed, it is possible that the visual crowding effects demonstrated here may be further exaggerated or attenuated depending upon what assumptions one makes about the organisation of information underlying semantic categories.

## **Experiment 6**

### ***Introduction***

In the introduction, we discussed potential concerns about how a model's behaviour can be contingent upon properties of a particular training set. To recap, the subjective choice of features used to represent a range of objects within a connectionist model may give rise to interesting emergent behaviour that would not be observed if the same items were encoded across a completely different set of features. Our aim, in using images as visual input, was to reduce the risk of subjectivity in feature selection. Each image in our training set was represented over 2500 features (pixels) and each individual feature could take any value between 0 and 255, giving considerable scope for between-item variability. The value of any given feature for any given image was determined by its spatial position within that image, rather than being pre-determined by the investigators. In this way,

features in our model were simply components of a picture rather than having a direct relationships with nameable visual features (e.g. 'is large', 'has eyes', 'is brown', etc.). By adopting this approach, we aimed to reduce the level of bias that is introduced when an experimenter selects a specific set of features to represent a diverse range of objects. However, this method is not fully immune to experimenter bias because it is arguable that the exact choice of items in the training set will have just as important a bearing on the model's final behaviour as the set of features chosen to represent those items. We adopted rigorous procedures when selecting images for this study: basic level categories were chosen carefully to reflect the range of visual diversity within each superordinate category; the choice of images was given to an individual who was not a member of the investigative team; all images were required to depict objects in typical and canonical perspectives; all images were standardised for size and centralised within the input space; all images were left-right inversed to reduce standard orientation biases. Moreover, in experiment 4, we tested a variant of our model with stimuli which were not experienced during training (N.B. these stimuli were novel to the neural network but not to the original set of chosen stimuli). However, despite these control measures, it would be impossible to eliminate subjectivity in image choice simply because a set of images cannot be chosen at random from the population of all possible pictures. The question still remains as to whether the model would generate similar behaviour for an entirely different set of pictures and we address this in the final experiment by testing the model with images obtained from a new source.

### ***Method***

A new test set of images comprising 2 examples of each basic level animal and musical instrument category (28 images in total falling within the same basic level categories that were used in experiments 3, 4 and 5) was assembled using the 1995 Grolier Multimedia CD-ROM encyclopaedia. No images of furniture or clothing were available within this volume so we were not able to use these categories. This new source of stimuli was ideal for this study because all depicted items were very high quality coloured line drawings depicting each object, in a canonical view, against a plain white background. Each image was converted to 8-bit greyscale and centred within a 50 by 50 pixel grid, thereby conforming to the method used for constructing the original image set. These new images were then presented to each of the 5 VPMs used in experiment 4, which had already been trained with



448 images from the original set of items.<sup>8</sup> After training, the output vectors from each VPM were standardised (range 0-1) and presented to half of the fully-trained CMs (chosen at random) which had been paired with each VPM in experiment 4 (i.e. 15 CMs in total were tested). Classification performance was scored by the same criteria used in experiment 4. Although this method does not allow us to validate our model with a completely novel *training* set, it does offer a useful test of whether the model's performance is contingent upon properties of a *specific* training set of images chosen by one person. If this is indeed the case, we would expect a different pattern of generalisation performance to that seen in experiment 4. If, however, our model's performance is independent of the training set, we should expect a similar pattern. To test this formally, we ran a by-subjects analysis of variance on the data from the 15 CMs which were used in both this experiment and experiment 4. A repeated measures by-items analysis would not be appropriate because the actual test items differed, both in source and quantity. Test set (expt. 4 vs. expt. 6), level of classification (superordinate vs. basic level) and category (animals vs. musical instruments) were nested within-items factors in the ANOVA model.

### **Results**

Model performance is plotted in Figure 13, showing the level of consistency between the 2 image test sets (original vs. new). The small confidence intervals at each measure indicate high levels of between-subject consistency and this is despite the fact that the 15 CMs did not all experience identical training sets.

*INSERT FIGURE 13 HERE*

For analytical purposes, chance performance was partialled out (25% and 3.57% for superordinate and basic level classification respectively). There was no main effect of category ( $F[1,14] < 1$ ) but level of classification was a significant factor ( $F[1,14] = 4.71$ ,  $p < 0.05$ ; overall superordinate (SO) and basic level (BL) means adjusted for chance were 49.6 and 48.6). There was a significant main effect of test set ( $F[1,14] = 5.6$ ,  $p < 0.05$ ) although the effect size was very small (but, nonetheless, consistent) with

---

<sup>8</sup> Each of these 5 training sets had omitted a *different* subordinate from each basic level category, meaning that the composition of each training set was different but that all images were represented in 4 out of 5 sets - for further clarification, see the method section of experiment 4.

adjusted means of 50.1 and 48.2 for original and new test sets respectively. The 3-way interaction was significant ( $F[1,14] = 22.3, p < 0.0001$ ) although post-hoc tests (Bonferroni) showed significant differences between sets *only* for musical instruments at the *superordinate* level (50.6% vs. 45.2%). It is notable that the 2 significant main effects derived from very small effect sizes (approx 1% and 2% respectively for level of classification and test set). A small number of replications (15 in this case) would not usually provide sufficient power to detect as significant, such a small difference between means. However, the classification performance of CMs was highly consistent between simulations, and it is this consistency which accounts for such small differences reaching significance.

### ***Discussion***

Although there was a small quantitative difference in classification accuracy levels between the original and novel set, the overall profile of performance were very similar (figure 13). The data show that our model's behaviour must be somewhat contingent upon properties of a particular training set although we would argue that the magnitude of the difference is far too small to dismiss our model's performance as a toy domain. Whilst we would accept that our model will not generate the exact response profile reported in this paper for *every* possible set of images, we would suggest that it does capture important elements of the simulated domain and that our results demonstrate how visual crowding can be an important factor underlying at least some category specific impairments.

## **SUMMARY OF RESULTS**

In this brief section we provide the a brief summary of the key findings reported within this paper. In experiment 1, a visual processing module (VPM) based on Kohonen's (1982a; 1982b; 1988) self organising feature map was presented with standardised, digitised images of animals, musical instruments, clothing, and furniture - categories which have been extensively used in testing of patients with category specific agnosias. The VPM was able to represent a diverse set of 560 greyscale images deriving from 7 basic level categories within each of 4 superordinates. When the weight vector for each output unit in a trained VPM was used to generate an image of the type of stimulus that would generate maximal response in that particular unit, there was a strong tendency for furniture

and clothing exemplars to be depicted at an exemplar level, whereas animals and musical instruments tended to be visible in amalgams of several images, thus implying that they are less visually distinctive. In experiment 2, patterns of activation across the whole VPM surface were analysed. Different types of contour map could be identified for categories of animals or musical instruments compared with furniture and clothing. For the latter categories, coding in the VPM was more localised with the majority of output units having background levels of activity. Such localised representation resulted in little overlap in VPM patterns across basic level categories (i.e. minimal visual crowding). Categories of animals and musical instruments, by contrast, were characterised by relatively higher levels of activity in a greater number of VPM units, and there was considerable overlap of activity across basic level categories (i.e. high visual crowding). In experiment 3, the rate at which acquired patterns of activity in the VPM would lead to correct classification by a supervised Categorisation Module (CM) varied between categories. Although the CM could eventually identify the correct category for each VPM pattern with similar accuracy, animals and musical instruments required substantially greater training time, relative to furniture and clothing, to reach a state of re-representation that would permit a comparable level of identification accuracy. This data supported the view that visual crowding may underlie some processing difficulties for living things and musical instruments. In experiment 4, the fully trained model (i.e. VPM and CM combined) was tested for its ability to generalise mappings from basic and superordinate level categories it had already acquired, to *novel* exemplars. Basic level object recognition of novel stimuli was significantly less accurate for furniture compared to animals, musical instruments and clothing. Classification of superordinate category was significantly more accurate for animals and musical instruments than furniture and clothing. Recognition errors for novel images of animals and musical instruments tended to be within-category errors (i.e. members of a perceptually related basic level category) rather than recognition failures which was the case for furniture and clothing. In experiment 5, localised category output representations were replaced by distributed representations within the CM in order to more plausibly simulate the mapping process between pre-semantic and semantic representations. After lesion damage, visually crowded categories were more prone to categorisation failure, although the effect size was moderated by lesion site. Finally, experiment 6 replicated the results of experiment 4 using a completely novel test set of pictures. Although a small quantitative difference was observed,

the general pattern of performance across the two different test sets was remarkably consistent suggesting that the behaviour of our model is quite tolerant of changes within the stimulus set.

## **GENERAL DISCUSSION**

In this paper we have described a connectionist model of object processing which takes, as input, digitised greyscale images and classifies each image according to its semantic category. The model has provided an alternative approach to exploring variables assumed to operate at pre-semantic levels of processing and, more specifically, their potential role in emerging category impairments. Before discussing the implications of our findings, we firstly focus upon some strengths and limitations of the approach.

Firstly, our model does not seek to explain category specific deficits in terms of representation within the semantic system. In experiment 5 we demonstrated how, in a neutral semantic system, lesion damage at the level of semantics can give rise to a relative impairment in activating correct information for visually crowded categories. This pattern derives, however, from a variable operating at a pre-semantic stage of processing (i.e. visual crowding) rather than from variables operating within the semantic system itself. Thus, visual crowding has some potential in explaining category specific deficits for living things (and musical instruments), even when the locus of damage is post-perceptual. However, this explanation does not preclude the possibility that the same type of impairment may also derive from factors operating at a semantic level. For example, Farah and McClelland (1991) have demonstrated how a bias in the distribution of sensory and functional semantic information can underlie both living and non-living thing deficits. Our model is not mutually exclusive to such accounts but does demonstrate the potential contribution of factors which have not been addressed in some earlier models.

Secondly, our model only comfortably accounts for one pattern of deficit, namely an impairment in recognising living things. Living things (and musical instruments) were more difficult to classify (expt. 3) and more prone to the disruptive effects of lesion damage (expt. 5). However, our simulations provided no evidence that the reversed pattern (i.e. a non-living thing deficit) might be

attributable to pre-semantic factors, suggesting that impairments in recognising non-living things might be better accounted for by representational characteristics of object concepts within the *semantic* system. Although our model was less accurate at recognising *novel* pictures of non-living stimuli (excluding musical instruments – expt. 4), this pattern emerged *before* the model was lesioned and does not, therefore, simulate the emergence of a non-living thing disorder after brain damage. It does, however, support the view that non-living things may be more difficult to recognise under normal conditions owing to greater within category structural variability at the basic level (see Laws and Neve, 1999; Turnbull and Laws, 2000). It must also be noted that our model simulates living thing deficits for *pictorial* stimuli and does not specifically address those cases where such impairments are demonstrated across different input modalities. Pictures are generally assumed to activate semantic representations via stored structural representations (Humphreys et al., 1988) whereas words are thought to tap conceptual knowledge more directly (Caramazza et al., 1990; Chertkow, Bub and Caplan, 1992; Riddoch and Humphreys, 1987a). Whether the distinction is actually this clear cut or not, the fact that some neurological patients demonstrate greater difficulty in accessing semantics via one modality (e.g. Bub, Black and Hampson and Kertesz, 1988; Riddoch and Humphreys, 1987b) does suggest different semantic access procedures between different types of stimuli. Those cases where living thing deficits are demonstrated under testing with both verbal and pictorial stimuli can only be explained by visual crowding effects by assuming that verbal definitions rich in visual information tap stored structural representations. However, there appears to be no theoretical basis for speculating that an item's visual similarity to within-category associates should influence object *name* perception. Nonetheless, one can argue minimally that where a living thing deficit is restricted to pictorial stimuli (see the investigation of patient JB described by Riddoch and Humphreys, 1987b), visual crowding may play a role in the impairment.

A relative strength of this model is its ability to predict a co-occurrence of impaired performance for both living things *and* musical instruments. This pattern is not so easily accounted for by theories positing representational biases at a level of semantic processing. For example, a sensory/ functional distinction may neatly predict a living/ non-living dissociation but it is difficult to conceive how it could explain a dissociation between musical instruments and other artefact categories (e.g. tools).

Connectionist models have been criticised by some (e.g. Reeke and Sporns 1993) for being 'homuncular' (see Reeke and Edelman, 1988) whereby a human interpreter must provide suitable neural inputs and infer appropriate behaviour from a model's output states. It is claimed that many modellers utilise a set of engineering principles to build an artefact which merely behaves in a similar way to a human, yet does not capture underlying processing in the simulated domain. Moreover, connectionist models can often oversimplify a problem domain, permitting tolerable explanations yet offering little by way of testable predictions. Reeke and Sporns (1993) argue that ideal models should have 'inputs that are actual sensors and outputs that are actual effectors working on the environment' (p. 598) and also fulfil strict psychological or biological plausibility criteria. Whilst our model falls a considerable way short of fulfilling these very stringent criteria<sup>9</sup>, we have made an attempt to address some general criticisms of neural models. Given the degree of pre-processing applied to our pictorial stimuli, our input data cannot really be described as sensory in nature. Nonetheless, our images constitute a level of representation that is closer to actual visual data than an inferred list of nameable attributes. For example, a greyscale image provides information about shape, spatial location and the distribution of shading across the object's surface. This type of information does not, on its own, specify a depicted object in terms of its parts and, therefore, has some parallels with bottom-up visual data. By contrast, an inferred list of visual features implies that a certain amount of processing has already been carried out to extract those features which are relevant in the object recognition process. Given the lack of conclusive evidence about which type of visual features are important, this approach may sometimes be misleading. By utilising pictorial stimuli, we have not completely avoided the homunculus problem because there is still a degree of experimenter subjectivity in stimulus selection. However, we have provided data in experiment 6 which suggests that our model's behaviour is not an artefact of a specific training set.

The issue of subjectivity in image choice is an important consideration: there is always a possibility that one could select a completely different set of images depicting items from within the same

---

<sup>9</sup> It is of course questionable whether Reeke and Sporn's criteria for modelling are reasonable in all domains. For example, it is quite possible that the introduction of unnecessary complexity may actually undermine a model's ability to test the adequacy of key theoretical assumptions.

taxonomies, which would *not* give rise to the same pattern of behaviour observed in these experiments<sup>10</sup>. The question then arises as to whether our stimuli provide a *reasonable* reflection of our chosen basic level and superordinate categories and whether any systematic biases exist which might contribute towards the behaviour of the model. Within any given set of images there will always be idiosyncracies that are difficult to control. For example, our 5 examples of snakes vary from a simple vertical squiggle to a complex set of convolutions. The complexity of the coils reflect movement and are accidental to the basic simple shape, yet the model has no way of knowing this and must represent all variations. Conversely, there is less shape variation within the images of frogs because posture is very similar in all pictures (for example, there were no images which depicted frogs in mid-leap). Such variations would be difficult to control in pictorial stimuli and are, largely, a by-product of the way in which animals are depicted in reference media. For example, frogs are typically photographed in a stationary, seated position because this is probably the position in which they are most commonly visible (and most easily photographed). To control for variations in the shape of living things which derive from movement would be very difficult and the fact that some of our chosen basic level living thing categories reflect shape/ movement interactions is purely accidental. Nonetheless, it could be argued that these variations constitute a confounding factor since they are present in only one of 4 categories. We would, however, argue that they do not undermine but, rather, strengthen the case for visual crowding being a potentially important factor in category specific impairments. For example, although some of our basic level animal categories may have been rendered less visually coherent by movement/ shape interaction, the category of living things, as a whole, was still visually crowded relative to furniture and clothing. Moreover, living things generated a very similar pattern of model performance to musical instruments which, by contrast, are not prone to shape/ movement interactions: indeed, even though accidental shape variations might hypothetically render basic level animal classification more difficult (compared to musical instruments), our data suggests otherwise, since there was no difference in the level of basic level classification accuracy for *novel* animals and musical instruments in both experiments 4 and 6.

Another important issue relating to stimulus characteristics is whether the *actual* categories of items

---

<sup>10</sup> For example, it is likely that line drawings would give quite different results because they have little textural or shading information.

utilised in these experiments would elicit recognition difficulties in genuine category specific patients. The CM in our model confused many different basic level categories within each superordinate (e.g. snake/ frog; bird/ mouse; violin/ guitar; flute/ saxophone), whereas most visual agnosics, unless severely impaired, would confuse only those basic level objects which shared similar shapes and configurations of features (e.g. flute/ saxophone; violin/ guitar) rather than items which have very different shapes (e.g. snake/ frog; bird/ mouse). This issue relates particularly to our representation of living things because the basic level categories within this group *all* appear to be visually dissimilar to each other (cf. musical instruments, for example, where some items like guitar and violin have similar shapes and configurations of features). All our basic level categories were specifically selected to reflect the range of visual diversity within each superordinate group so, for example, the basic level animal categories we chose represented several different genera (e.g. reptiles, amphibians, mammals, birds, etc.). If, instead, we had selected all our living thing stimuli from within the *same* genera (e.g. frog, toad, newt, salamander), it would be arguable that we were deliberately constraining similarity and that this selection of items may have generated a pattern of performance atypical of the whole category of animals. To represent a broad range of different animals in which several genera were represented by more than one basic level category would require a much larger training set and this would make it difficult to counterbalance the number of basic level categories within each of the 3 other superordinates. In practice, our model's 'visual' experience is highly limited and, owing to control measures, cannot reflect overall category size. In the real world, the category of animals probably includes many more basic level items than any other category. This, combined with the fact that living things tend to be visually crowded, visually complex, relatively unfamiliar *and* of low word frequency, may well mean that greater learning effort is required to make accurate within-category discriminations for living things.<sup>11</sup> Within our simulations, however, only visual crowdedness (and possibly visual complexity) could influence category learning because both familiarity and category size were counterbalanced (and word frequency did not apply in this model). Considering category size alone, the model had no greater expertise with any one of the 4 categories. Were we to increase the number of exemplars in each superordinate (in proportion with overall category size) such that

---

<sup>11</sup> Differentiating items from different genera would be relatively easy and would not require the same level of expertise that is needed within a field like ornithology, for example. By contrast, within-genera discriminations would tend to be more difficult for the non-expert and this is why they probably elicit a higher error rate in visual agnosic patients



living things accounted for more than a quarter of the training set, the model might become more expert in classifying animals and may, perhaps, make fewer ‘coarse-grained’ visual discrimination errors (e.g. frog/snake) and more ‘fine-grained’ discrimination errors (e.g. spider/beetle). Nonetheless, even if this were the case, the basis of any confusions would still be high visual crowding and we would predict that an increase in category size would only increase the size of the crowding effect seen within these simulations. In short, we suggest that if a crowding effect emerges in a category that is deliberately set up to include a visually diverse range of items, it should be accounted for in models seeking to explain category specific impairments.

We have discussed possible criticisms of the stimuli presented to our model and have argued that these do not undermine the principal finding of this study. However, in addition stimulus characteristics, connectionist models may also be evaluated by the type of architecture used to implement a particular theory. In a departure from some other simulations of category specific agnosia, we utilised an unsupervised module (the VPM) to model the formation of perceptual categories. Whilst unsupervised neural networks are not immune from experimenter effects, they are potentially useful in modelling perceptual processes because they develop representations without a-priori assumptions about the information that is extracted from the environment (Luckman et al., 1995; Schynns, 1991).<sup>12</sup> Thus, the VPM representations which evolved in our simulations did so in response to the topology of visual stimuli rather than being pre-ordained categorical representations. A full evaluation of the psychological and biological plausibility of our model is beyond the scope of this article. Nonetheless, it is worth pointing out some broad similarities with neural representation in the human brain. For example, there appear to be separate pathways in visual object recognition, most notably the ventral visual pathway and a separate route which is critically dependent on the posterior parietal and frontal lobes (Gainotti et al., 1995; Logothetis and Sheinberg, 1996). The inferotemporal cortex in the ventral visual pathway is considered to have a role for storing the ‘central representations’ of objects (Logothetis and Sheinberg, 1996; Mishkin, Malamut and Bachevalier, 1984). Coding in the inferotemporal cortical system is characterised by neither sparse, localised

---

<sup>12</sup> Although it should be pointed out here that whilst unsupervised networks make weaker assumptions about the environment, they embody stronger assumptions about the *form* of representation that will develop (cf. 3-layer back-propagation networks).

representations (i.e. single neurones signalling an object) nor highly distributed representations (Fotheringham and Young, 1997). Instead, neighbourhoods of neurones appear to be tuned to classes of stimuli whereby each cell may respond to a range of stimuli varying along a particular dimension. Physiological selectivity of stimuli to which inferotemporal cortex cells respond (e.g. facial stimuli) is correlated with physical similarity derived from multidimensional scaling of the stimuli. Thus, in some respects, the VPM in our model appears to represent the visual properties of objects in a manner similar to that of the inferotemporal cortical system by reducing dimensionality of the external world and representing objects within multidimensional spaces. Moreover, the overall clustering behaviour of the 10 *individual* VPM simulations in experiments 1 and 2 was similar suggesting that it is the configurations of neuronal firings, rather than the activation values of single cells, that generate meaningful representation over time. So, whilst our model is relatively simple and does not capture anything like the full complexity of human visual processing, it has a certain degree of biological and psychological plausibility.

It is well documented that apparent living thing deficits may emerge through poor control of variables known to affect naming performance in normal and neurological subjects (e.g. Funnell and Sheridan, 1992; Stewart et al., 1992). The artefact explanation has been examined under several paradigms, one of which is statistical modelling. Here, multiple regression is used to see whether partial correlation of the categorical factor (living vs. non-living) is still a significant predictor of recognition performance after complexity, familiarity, frequency and within-category similarity have been taken into consideration (e.g. Farah et al., 1991; Kurbat, 1997). Kurbat and Farah (1998) demonstrated that, when measurement error was minimised or modelled using simulated data (e.g. Monte-Carlo simulations), living things still produced higher error rates in patients and slower reaction times in normal subjects. Thus, the categorical factor (living vs. non-living) was still a significant predictor of recognition accuracy. However, this may be a premature conclusion. A number of studies have suggested that visual similarity is a key variable influencing the organisation of the structural description system as well as subsequent access to an amodal semantic store (Forde et al., 1997; Humphreys, et al., 1988; Humphreys et al. 1995; Lloyd-Jones and Humphreys, 1997; Riddoch and Humphreys, 1987b; Sartori and Job, 1988). Moreover, from the simulation data presented in this paper, it appears that a self

organising system can, over time, develop different ways of representing living and non-living things based upon their visual similarity. These qualitatively different types of representation generate different performance characteristics in learning (expt. 3), recognition of novel items (expts. 4 and 6) and post-lesion classification (expt. 5). Thus visually crowded categories may not simply require finer visual differentiation for stored knowledge to be retrieved (Forde et al., 1997; Humphreys et al., 1988) but, during learning, may actually give rise to different types of representation: when learning was complete within our model the VPM was organised, to a certain extent, along categorical lines. This questions the validity of the debate as to whether category specific deficits result from categorical organisation of the structural description system or greater susceptibility of visually crowded categories because, in a self organising system, visual similarity can lead to some degree of categorical organisation. Since measures of visual similarity which have been used previously (e.g. Farah et al., 1991; Humphreys et al., 1995) are less precise than the competitive learning algorithm used in our VPM, this may explain why some statistical models of patient performance require the categorical factor as an additional predictor variable.

The cascade model of Humphreys et al. (1988) holds that competition between similar items at one level of visual processing will propagate forward to subsequent levels. Thus, increased competition between structural representations, especially if degraded, will be reflected in the semantic (and phonological) system by within-category recognition errors (e.g. 'cow' for deer). Superordinate categorisation tasks, by contrast, will induce few, if any, errors (Humphreys et al. 1988; Sartori and Job, 1988; Sartori et al. 1993; Silveri and Gainotti 1988). Like the simulation of Humphreys et al. (1995), the findings in experiment 5 demonstrate how lesioning within the 'semantic' system can result in an impairment for living things even though the category bias is pre-semantic. Although the hidden and output layer of the CM 'separated out' the visually crowded exemplars to a certain degree, the effects of visual crowding were still evident within representations at these later processing stages. For example, lesions at all 3 sites (i.e. IH, HO and OO) impaired access to basic level semantic information for all items but the effect was more pronounced for those belonging to visually crowded categories. However, activation of superordinate features was, under mild HO and OO lesions, slightly worse (relative to basic level features) for visually uncrowded categories. Thus a low degree of lesioning (i.e.

5%) in the HO and OO stages of our model accords to some extent with the incongruous finding reported in patient RC (Moss et al., 1998) who was significantly better at sorting living things (93% correct) than non-living things (63% correct) into superordinate categories despite a typical category specific agnosia for living things. Unlike the simulation of Humphreys et al (1995), the hidden layer in our CM performed somewhat different functions for visually crowded and uncrowded categories evidenced by a dramatic reduction in the visual crowding effect observed after HO lesions (cf. IH lesions). On one hand, the CM appeared to be re-representing items from visually crowded categories by exaggerating subtle inter-item differences (i.e. *separation*) whereas, on the other, it re-represented items from visually uncrowded categories by reducing the amount of discrepancy between visually dissimilar items within the same category (i.e. *convergence*). The dual function of the CM hidden layer has been borne out in cluster analyses of the representations at each stage of a fully trained CM (Gale, 1997). Thus, semantic proximity can exert a strong modulatory effect upon visual similarity.

As we have already pointed out, a relative strength of our model is its ability to account for a co-occurrence of recognition problems for both living things *and* musical instruments. That such an apparently bizarre pattern of patient deficit should be so consistently reported has been the focus of much interest (e.g. Gainotti et al., 1995; Parkin and Stewart, 1993). Meanwhile, purported category specific deficits for non-living things tend to be restricted to small manipulable objects rather than large, outdoor objects (e.g. Warrington and McCarthy, 1987) and there is frequently a co-occurrence with impaired knowledge for body parts. The explanation offered by Gainotti et al. (1995) is that, for the former, patients fail to retrieve visual-perceptual features defining the different members of each category. For the latter, however, the critical knowledge is mainly somato-sensory, and depends upon motor representations. Thus, non-living thing deficits would arise through difficulties in accessing somato-sensory, rather than visual-perceptual, representations. If this is the case, the influence of perceptual crowding should hold true for somato-sensory relatedness as much as it does for visual relatedness in our simulations. Thus it is possible that if the somato-sensory representation system is also self-organising, the converse pattern of representation to that reported here will occur (i.e. greater coarse-coding for non-living things and little representational overlap for living things).

## REFERENCES

Allinson, N.M., Ellis, A.W., Flude, B.M., and Luckman, A. (1992) A connectionist model of familiar face recognition. Paper presented at Digest of IEEE Colloq. on *Machine Storage and Recognition of Faces*, January, 1992: London, U.K.

Arguin, M., Bub, D.N., and Dudek, G. (1996) Shape integration for visual object recognition and its implication in category-specific visual agnosia. *Visual Cognition*, 3, 221-275.

Basso, A., Capitani, E. and Laiacona, M. (1988) Progressive language impairment without dementia: a case with isolated category specific semantic defect. *Journal of Neurology, Neurosurgery, and Psychiatry*, 51, 1201-1207.

Battig, W.F., and Montague, W.E. (1969) Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 80, 1-146.

Baum, E.B., and Haussler, D. (1989) What size net gives valid generalisation? *Neural Computation*, 1, 1-46.

Behl-Chadha, G. (1996) Basic-level and superordinate-like categorical representations in early infancy. *Cognition*, 60, 105-141.

Behl-Chadha, G., Eimas, P.D., and Quinn, P.C. (1995) Perceptually-driven superordinate categorisation by young infants. Paper presented at the biennial meeting of *The Society for Research in Child Development*, Indianapolis, IN.

Bub, D.N., Black, S., Hampson, E., and Kertesz, A. (1988) Semantic encoding of pictures and words: some neuropsychological observations. *Cognitive Neuropsychology*, 5(1), 27-66.

Caramazza, A (1998) The interpretation of semantic category-specific deficits: what do they reveal about the organization of conceptual knowledge in the brain ? *Neurocase*, 4, 265-272.

Caramazza, A., Hillis, A., Leek, E.C., and Miozzo, M. (1994) The organization of lexical knowledge in the brain: evidence from category- and modality-specific deficits. In L.A. Hirschfeld and S.A. Gelman (Eds.) *Mapping the Mind*. Cambridge NY: Cambridge University Press.

Caramazza, A., Hillis, A.E., Rapp, B.C., and Romani, C. (1990) The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology*, 7, 161-190.

Caramazza A, and Shelton J.R. (1998) Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *Journal of Cognitive Neuroscience*, 10, 1-34.

Chertkow, H., Bub, D., and Caplan, D. (1992) Constraining theories of semantic memory processing: evidence from dementia. *Cognitive Neuropsychology*, 9(4), 327-365.

Collins, A.M., and Loftus, E.F. (1975) A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.

Collins, A.M., and Quillian, M.R. (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behaviour*, 8, 240-247.

Coltheart, M, Inglis, L, Cupples, L, Michie, P, Bates, A, Budd, B (1998) A semantic subsystem of visual attributes. *Neurocase*, 4, 353-370.

Cosmides, L., and Tooby, J. (1994) Origins of domain specificity: the evolution of functional organization. In L.A. Hirschfeld and S.A. Gelman (Eds.) *Mapping the Mind*. Cambridge NY: Cambridge University Press.

Damasio, A.R. (1990) Category-related recognition defects as a clue to the neural substrates of knowledge. *Trends in Neuroscience*, 13, 95-98.

Devlin, J.T., Gonnerman, L.M., Andersen, E.S., and Seidenberg, M.S. (1998) Category-specific semantic deficits in focal and widespread brain damage: a computational account. *Journal of Cognitive neuroscience*, 10, 77-94.

Dixon, M., Bub, D.N., and Arguin, M. (1997) The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive neuropsychology*, 14 (8), 1085-1130.

Done, D.J. and Gale, T.M. (1997) Attribute verification in dementia of Alzheimer's Type: Evidence for the preservation of distributed concept knowledge. *Cognitive Neuropsychology*, 14, 547-571.

Durrant-Peatfield, M.R., Tyler, L.K., Moss, H.E., and Levy, J.P. (1997) The distinctiveness of form and function in category structure.

Edelman, S. (1998) Representation is representation of similarities. *Behavioural and Brain Sciences* 21, 449-498.

Eimas, P.D., and Quinn, P.C. (1994) Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, 65, 903-917.

Eimas, P.D., Quinn, P.C., and Cowan, P. (1994) Development of exclusivity in perceptually based categories of young infants. *Journal of Experimental Child Psychology*, 58, 418-431.

Farah, M.J., and McClelland, J.L. (1991) A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, 120, 339-357.

Farah, M.J., McMullen, P.A., and Meyer, M.M. (1991) Can recognition of living things be selectively impaired? *Neuropsychologia*, 29, 185-193.

Farah, M.J., Meyer, M.M., and McMullen, P.A. (1996) The living/ non-living dissociation is not an artefact: Giving an a-priori implausible hypothesis a strong test. *Cognitive Neuropsychology*, 13, 137-154.

Forde, E.M.E., Francis, D., Riddoch, M.J., Rumiati, R.I., and Humphreys, G.W. (1997) On the links between visual knowledge and naming: A single case study of a patient with a category-specific impairment for living things. *Cognitive Neuropsychology*, 14 (3), 403-458.

Fotheringham, D.K., and Young, M.P. (1997) Neural coding schemes for sensory representation: theoretical proposals and empirical evidence. In M.D. Rugg (Ed.) *Cognitive Neuroscience*. Psychology Press.

Fuchs, A., and Haken, H. (1988) Pattern recognition and associative memory as dynamical processes in a synergetic system. *Biological Cybernetics*, 60, 17-107.

Funnell, E., and Sheridan, J. (1992) Categories of knowledge? Unfamiliar aspects of living and nonliving things. *Cognitive Neuropsychology*, 9, 135-153.

Gaffan, D., and Heywood, C.A. (1993) A spurious category-specific visual agnosia for living things in normal human and non human primates. *Journal of Cognitive Neuroscience*, 5, 118-128.

Gainotti, G., Silveri M.C., Daniele, A., Giustolisi, L (1995) Neuroanatomical correlates of category-specific semantic disorders: a critical survey. *Memory*, 3, 247-264.

Gale, T.M. (1997) *Perceptual and semantic information in human object recognition: A neuropsychological and connectionist study*. Ph.D. Thesis, University of Hertfordshire, U.K.



Gonnerman, L.M., Andersen, E.S., Devlin, J.T., Kempler, D., and Seidenberg, M.S. (1997). Double dissociation of semantic categories in Alzheimer's disease. *Brain and Language*, 57, 254-279.

Grolier (1995) *The 1995 Grolier Multimedia Encyclopedia*, Version 7.0. On Line Computer Systems Inc.

Hillis, A.E., and Caramazza, A. (1991) Category-specific naming and comprehension impairment: A double dissociation. *Brain*, 114, 2081-2094.

Hinton, G.E., (1989) Connectionist learning procedures. *Artificial Intelligence*, 40, 185-234.

Hinton, G.E., and Shallice, T. (1991) Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, 98, 74-95.

Hinton, G.E., McClelland, J.L., and Rumelhart, D.E. (1986) Distributed representations. In D.E. Rumelhart, J.L. McClelland, and the PDP research group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1*. Cambridge, MA: MIT Press.

Hodges, J.R., Salmon, D.P., and Butters, N. (1991) The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain*, 114, 1547-1558.

Hodges, J.R., Salmon, D.P., and Butters, N. (1992) Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia*, 4, 301-314.

Humphreys, G.W., and Forde, E.M.E. (in press). Hierarchies, similarity and interactivity in object recognition: on the multiplicity of 'category-specific' deficits in neuropsychological populations. *Behavioural and Brain Sciences*.

Humphreys, G.W., Lamote, C., and Lloyd-Jones, T.J. (1995) An interactive activation approach to object processing: Effects of structural similarity, name frequency, and task in normality and pathology. *Memory*, 3, 535-586.

Humphreys, G.W., Riddoch, M.J., and Quinlan, P.T. (1988) Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67-103.

Keil, F. (1989) *Concepts, Kinds and Cognitive Development*. Cambridge, MA: MIT Press.

Kohonen, T. (1982a) Self-organised formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.

Kohonen, T (1982b) Clustering, taxonomy and topological maps of patterns. in M. Land (Ed.), *Proceedings of the Sixth International Conference on Pattern Recognition*. Silver Spring, MD: IEEE Computer Society Press.

Kohonen, T (1988) *Self-Organisation and Associative Memory*. Berlin: Springer-Verlag.

Komatsu, L.K. Recent views of conceptual structure. *Psychological Bulletin*, 112,3, 500-526.

Kurbat, M.A. (1997) Can the recognition of living things really be impaired? *Neuropsychologia*, 35, 813-827.

Kurbat, M.A., and Farah, M.J. (1998) Is the category-specific deficit for living things spurious? *Journal of Cognitive Neuroscience*, 10, 355-361.

Laiacona, M., Barbarotto, R., and Capitani, E. (1993) Perceptual and associative knowledge in category specific impairment of semantic memory: a study of two cases. *Cortex*, 29, 727-740.

Laiacona, M., Capitani, E., and Barbarotto, R. (1997) Semantic category dissociations: a longitudinal study of two cases. *Cortex* 33(3), 441-461.

Laws, K.R., and Neve. C. (1999) A 'normal' category-specific advantage for naming living things. *Neuropsychologia*, 37, 1263-1269.

Lloyd-Jones, T.J., and Humphreys, G.W. (1997) Perceptual differentiation as a source of category effects in object processing: Evidence from naming and object decision. *Memory and Cognition*, 25(1), 18-35.

Logothetis, N.K., and Sheinberg, D.L. (1996) Visual Object Recognition *Annual Review of Neuroscience*, 19, 577-621.

Luckman, A.J., Allinson, N.M., Ellis, A.W., and Flude, B.M. (1995) Familiar face recognition: A comparative study of a connectionist model and human performance. *Neurocomputing*, 7, 3-27.

McClelland, J.L., and Rumelhart, D.E. (1981) An interactive activation model of context effects in letter perception, 1. An account of basic findings. *Psychological Review*, 88, 375-407.

McClelland, J.L., and Rumelhart, D.E. (1988) *Explorations in parallel distributed processing*. Cambridge, MA: Bradford Books.

McRae, K., de Sa, V.R., and Seidenberg, M.S. (1997) On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126, (2), 99-130.

Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

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

Mauri, A., Daum, I., Sartori, G., Riesch, G., and Birbaumer, N. (1994) Category-specific semantic impairment in Alzheimer's Disease and temporal lobe dysfunction: A comparative study. *Journal of Clinical and Experimental Neuropsychology*, 16, 689-701.

Medin, D.L. (1989) Concepts and conceptual structure. *American Psychologist*, 44(12), 1469-1481.

Mervis, C.B., and Greco, C. (1984) Parts and early conceptual development: Comment on Tversky and Hemenway. *Journal of Experimental Psychology: General*, 113(2), 194-197.

Mishkin, M., Malamut, B., and Bachevalier, J. (1984) Memories and habits: Two neural systems. In G. Lynch, J.L. McGaugh and N.M. Weinberger (Eds.), *Neurobiology of Learning and Memory*, New York: The Guildford Press.

Moss, H.E., Tyler, L.K., Durrant-Peatfield, M., and Bunn, E.M. (1998) 'Two Eyes of a see-through': impaired and intact semantic knowledge in a case of selective deficit for living things. *Neurocase*, 4, 291-310.

Murphy, G.L., and Medin, D.L. (1985) The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316.

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

Parkin, A.J., and Stewart, F. (1993) Category-specific impairments? No. A critique of Sartori et al. *Quarterly Journal of experimental Psychology*, 46A, 505-509.

Perry, C. (1999) Testing a computational account of category-specific deficits. *Journal of Cognitive Neuroscience*, 11, 312-320.

Plaut, D.C., and Shallice, T. (1993) Deep Dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, 10, 377-500.

Plunkett, K., and Elman, J.L. (1997) *Exercises in Rethinking Innateness: A Handbook of Connectionist Simulations*. Cambridge, MA: MIT Press.

Quinn, P.C., and Eimas, P.D. (1986) On categorisation in early infancy. *Merill-Palmer Quarterly*, 32, 331-363.

Quinn, P.C., and Johnson, M.H. (1997) The emergence of perceptual category representations in young infants: A connectionist analysis. *Journal of Experimental Child Psychology*, 66, 236-263.

Reeke, G.N.Jr., and Edelman, G.M. (1988) Real Brains and artificial intelligence. *Daedalus Proc. Am. Acad. Arts Sci.*, 117, 143-173.

Reeke, G.N.Jr., and Sporns, O. (1993) Behaviourally based modelling and computational approaches to neuroscience. *Annual Review of Neuroscience*, 16, 597-623.

Riddoch, M.J., and Humphreys, G.W. (1987a) Picture Naming. In G.W. Humphreys and M.J. Riddoch (Eds.), *Visual Object Processing: A cognitive Neuropsychological Approach*. London: Lawrence Erlbaum Associates Ltd.

Riddoch, M.J. and Humphreys, G.W. (1987b) Visual object processing in optic aphasia: a case of semantic access agnosia. *Cognitive Neuropsychology*, 4, 131-185.

Rosch, E. (1973) Natural categories. *Cognitive Psychology*, 4, 328-350.

Rosch, E. (1975) Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.

Rosch, E., and Mervis, C.B. (1975) Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.

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

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986a) Learning internal representations by error propagation. In Rumelhart, McClelland, and the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1*. Cambridge, MA: MIT Press.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986b) Learning representations by back-propagating errors. *Nature*, 323, 533-536.

Sacchett, C., and Humphreys, G.W. (1992) Calling a squirrel a squirrel but a canoe a wigwam: A category-specific deficit for artefactual objects and body parts. *Cognitive Neuropsychology*, 9, 73-86.

Sartori, G., and Job, R. (1988) The oyster with four legs: A neuropsychological study on the interaction between vision and semantic information. *Cognitive Neuropsychology*, 5, 105-132.

Sartori, G., Job, R., and Coltheart, M. (1993) The neuropsychology of visual semantics. In D.E. Meyer and S. Kornblum (Eds.), *Attention and Performance. XIV*. Cambridge, MA: MIT Press.

Sartori, G., Job, R., Miozzo, M., Zago, S., and Marchiori, G. (1993) Category-specific form-knowledge deficit in a patient with herpes simplex encephalitis virus. *Journal of Clinical and Experimental Neuropsychology*, *15*, 280-299.

Schynns, P.G. (1991) A modular neural network model of concept acquisition. *Cognitive Science*, *15*, 461-508.

Sheridan, J., and Humphreys, G.W. (1993) A verbal-semantic category-specific recognition impairment. *Cognitive Neuropsychology*, *10*, 143-184.

Silveri, M.C. and Gainotti G (1988) Interaction between vision and language in category-specific impairment. *Cognitive Neuropsychology*, *5*, 677-709.

Sirigu, A., Duhamel, J.-R., and Poncet, M. (1991) The role of sensorimotor experience in object recognition. *Brain*, *114*, 2555-2573.

Small, S.L., Hart, J., Nguyen, T., and Gordon, B. (1995) Distributed representations of semantic knowledge in the brain. *Brain*, *118*, 441-453.

Snodgrass, J.G., and Vanderwart, M. (1980) A standardised set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, *6*, 174-215.

Stewart, F., Parkin, A.J., and Hunkin, N.M. (1992) Naming impairments following recovery from Herpes Simplex Encephalitis: Category-specific? *Quarterly Journal of Experimental Psychology*, *44A*, 261-284.

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

Tippett, L.J., McAuliffe, S., and Farah, M.J. (1995) Preservation of categorical knowledge in Alzheimer's disease: A computational account. *Memory*, 3, 519-533.

Tranel, D., Logan, C.G., Frank, R.J., and Damasio A.R. (1997) Explaining category-related effects in the retrieval of conceptual and lexical knowledge for concrete entities: operationalization and analysis of factors. *Neuropsychologia*, 35, 1329-1339.

Turnbull, O.H., and Laws, K.R. (2000) Loss of stored knowledge of object structure: implications for 'category-specific' deficits. *Cognitive Neuropsychology*, 17, 365-389.

Tversky, B., and Hemenway, K. (1984) Objects, parts and categories. *Journal of Experimental Psychology: General*, 113, 635-657.

Tyler, L.K., Moss, H.E., Durrant-Peatfield, M.R. and Levy, J.P. (2000) Conceptual structure and the structure of concepts: a distributed account of category specific deficits. *Brain and Language*, 75, 195-213.

Waite, J., and Vincent, J.M. (1992) A probabilistic framework for neural network facial feature location. *BT Technological Journal*, 10, 20-28.

Warrington, E.K., and McCarthy, R.A. (1983) Category-specific access dysphasia. *Brain*, 106, 859-878.

Warrington, E.K., and McCarthy, R.A. (1987) Categories of knowledge: Further fractionations and an attempted integration. *Brain*, 110, 1273-1296.

Warrington, E.K., and Shallice, T. (1984) Category-specific semantic impairments. *Brain*, 107, 829-853.

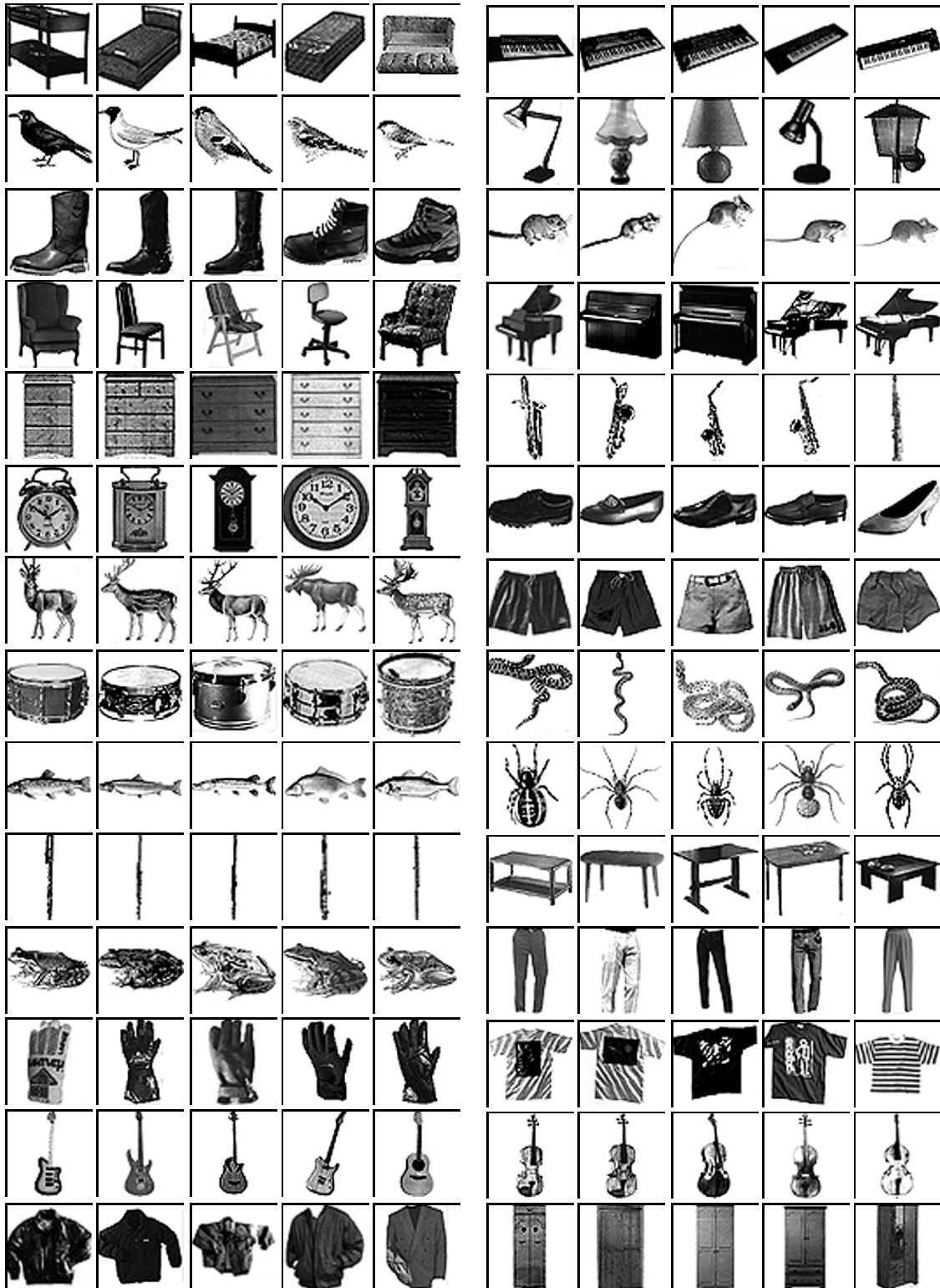
Yamane, S., Kaji, S., Kawano, K. (1988) What facial features activate face neurons in the inferotemporal cortex of the monkey? *Exp. Brain Res.*, 139, 209-214.





**APPENDIX A: The Image Set**

All images are displayed in alphabetical order of their basic level name (starting in column 1 and continuing in column 2). The basic level categories are: bed, bird, boot, chair, chest, clock, deer, drum, fish, flute, frog, glove, guitar, jacket, keyboard, lamp, mouse, piano, saxophone, shoe, shorts, snake, spider, table, trousers, t-shirt, violin, wardrobe. Five subordinates appear for each basic level group but, in the interests of brevity, the light-dark variations and left-right inversions do not appear.



**APPENDIX B:** Formal description of VPM training procedures1. *Description of VPM module.*

The Self Organising Feature Map (SOFM) module consists of  $n$  input neurons fully connected to  $M$  output neurons in a 2-dimensional grid of  $R$  by  $C$  rows.

2. *Algorithm used for training the SOFM*

(i) Initialise all the connections between the input layer and the SOFM layer with random values in the range of 0 to 1.

(ii) The connections that emanate from the input units to the  $i$ th output unit form the weight vector  $w_i = (w_{i0}, w_{i1}, \dots, w_{ij}, \dots, w_{in})$

(iii) For each epoch,  $t = 1, 2, \dots, T$ ,

(iv) For each input pattern,

$$A^k = (a_0^k, a_1^k, a_j^k, \dots, a_n^k) \quad k = 0 \dots K-1$$

For each unit in the SOFM output map  $I = 1 \dots M$

Calculate the euclidean distance between each unit weight vectors and the input pattern using

$$d_i = \sqrt{((a_1^k - w_{i1})^2 + \dots + (a_j^k - w_{ij})^2 + (a_n^k - w_{in})^2)} \quad (1)$$

Store the euclidean distances for each output unit in the SOFM output map  $x_i$

(v) Select the winning unit  $g = \min(d_i)$

(vi) Calculate the row  $r$ , and column  $c$  position of  $g$

(vii) Determine the neighbourhood size  $n$  from

$$n(t) = \text{INT}(n_0 (1 - t/T)), \quad (2)$$

with  $n_0$  being the initial neighbourhood size such that the neighbourhood includes at least half the map and  $T$  the maximum number of epochs.

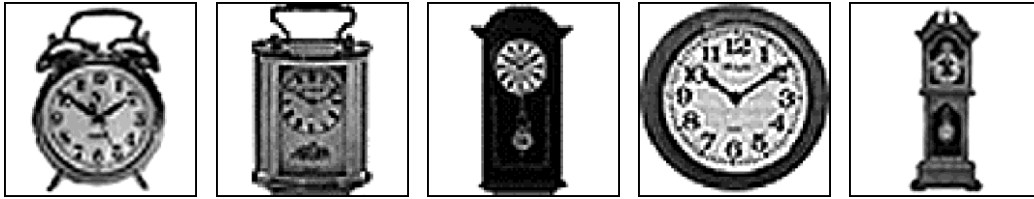
(viii) Create a list of  $g$  plus the units in the neighbourhood of  $g$ ,  $(r-n, r+n, c-n, c+n)$

(ix) For each unit  $h$  in the *neighbourhood* list move the units weight vector  $w_g$  closer to  $A^k$  using the equation:

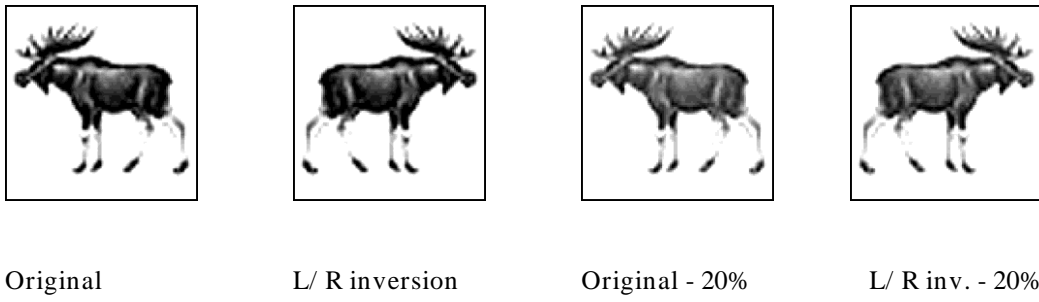
$$\Delta w_{hj} = \beta(t)[a_j^k - w_{hj}] \quad (3)$$

where  $j = (0, 2 \dots N-1)$  and  $\beta(t) = \beta_0 [1 - t/T]$ , with  $\beta_0$  being the initial training rate.

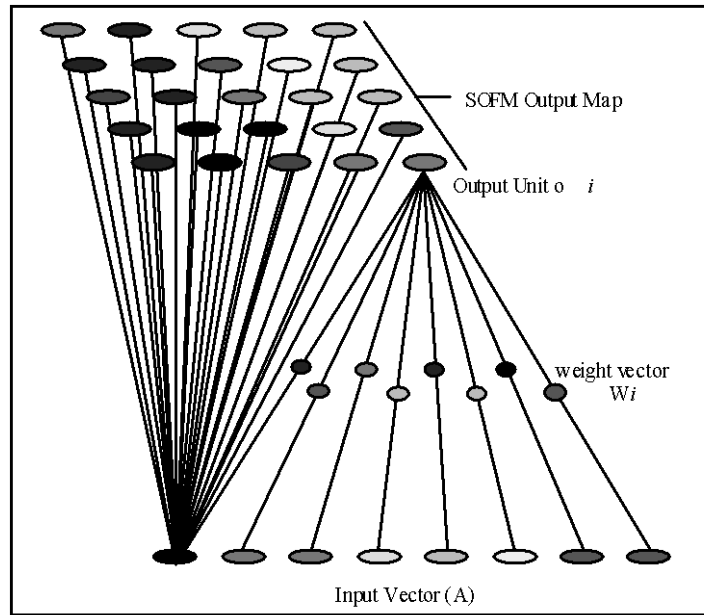
**FIGURES AND TABLES**



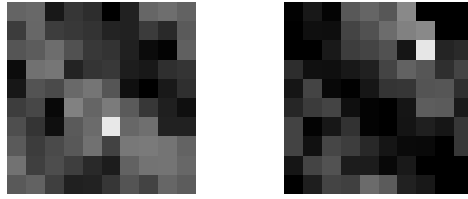
*Figure 1* Examples of a set of greyscale images presented to our model. Each image fits within a 50 by 50 pixel grid such that the principal dimension comes within one pixel of the grid border. The above examples depict the 5 subordinate images of the basic level category ‘clock’ which, in turn, is one of the 7 basic level categories representing furniture.



*Figure 2* Examples of left-right inversions and light/ dark variations for one subordinate image



*Figure 3* A schematic representation of the VPM - a self organising feature map (SOFM) connected to an input vector (A) by weight vector ( $W_i$ ). This diagram shows (i) how each SOFM output unit connects to every input unit and (ii) how each output unit will have its own n-dimensional weight vector, where n is the number of values in the input vector. Greyscale information depicts variability in both unit activation and weight values.



*Figure 4* Greyscale depictions of self organising feature map representations for 2 different training patterns. Light squares indicate areas of high activity (peaks) and dark squares indicate areas of low activity (troughs)

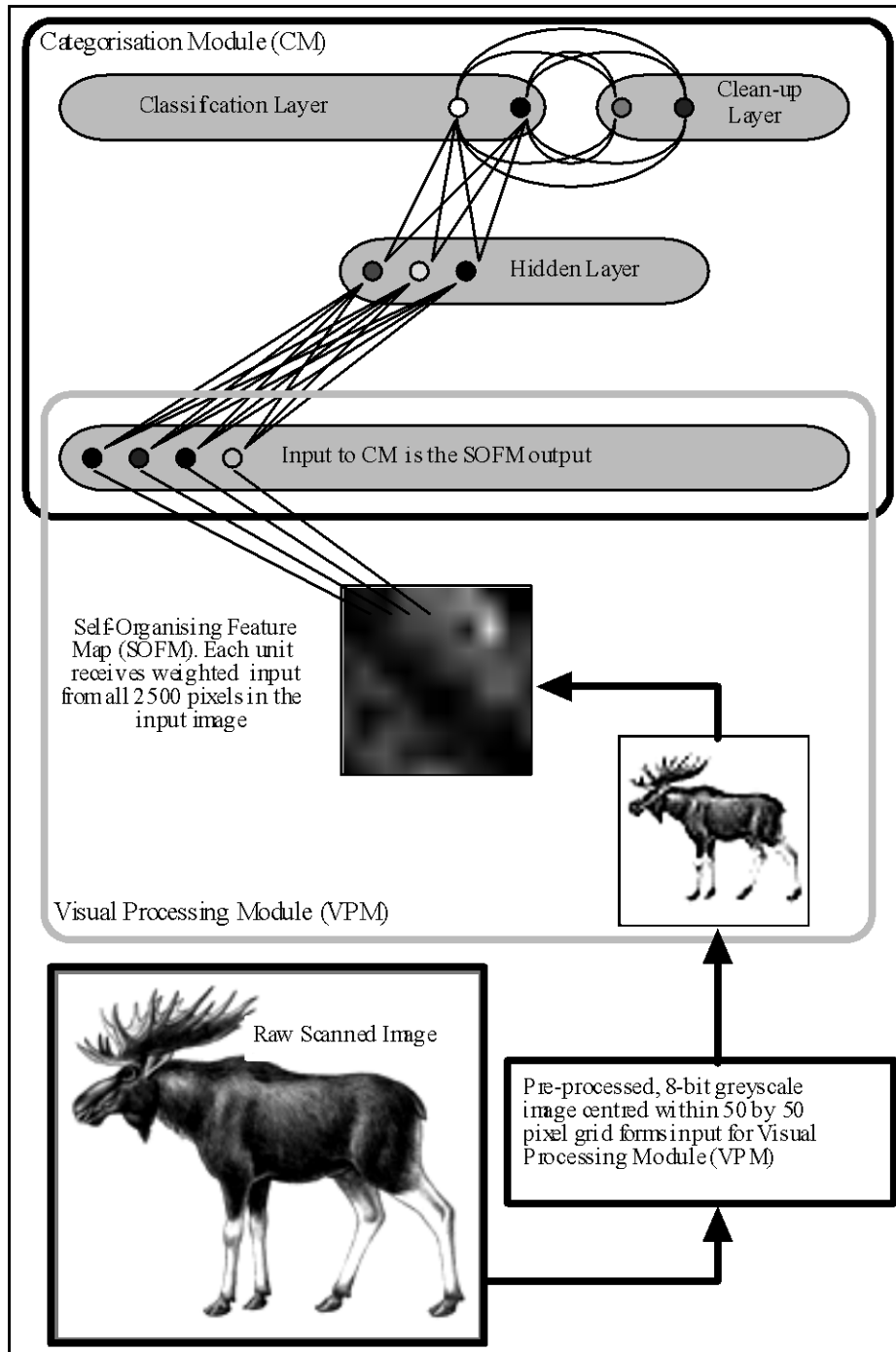


Figure 5 Schematic representation of the full model showing the interaction between VPM and CM

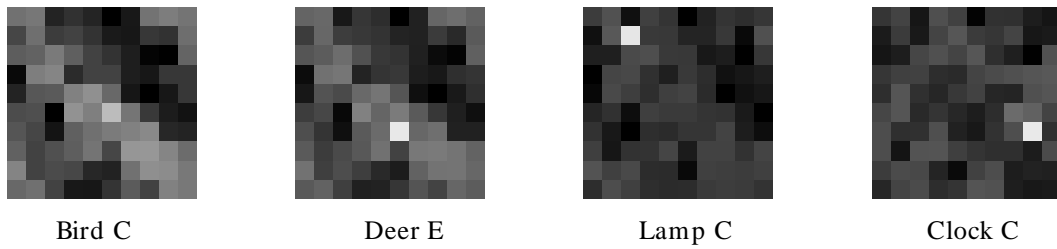




*Figure 6* A representation of the weight vectors for one SOFM. The images are mapped over the surface of a fully-trained 10 by 10 unit feature map, whereby each image represents a plot of the weight vector connecting the corresponding unit to the input array. Thus the first image in the top left hand corner is a plot of the weight matrix for the unit in the top left hand corner of SOFM. The competitive learning algorithm dictates that the weight vectors for a winning unit and its neighbouring units are moved closer in Euclidean distance to any training pattern for which they are highly activated (Kohonen, 1982a 1988; Schynns, 1991; Luckman et al., 1995). Thus, each image depicts the stimulus or configuration of stimuli to which the corresponding unit is maximally responsive. It should be kept in mind that the surface is wrapped around on itself such that patterns on opposite sides and corners are actually neighbouring each other.

<b>VPM</b>	<i>Animals</i>	<i>Furniture</i>	<i>M/Insts</i>	<i>Clothing</i>
1	20	34	18	28
2	19	35	18	28
3	20	36	16	28
4	20	36	16	28
5	19	33	18	30
6	19	34	17	30
7	20	34	18	28
8	18	36	19	27
9	20	33	20	27
10	19	34	18	29
<b>Mean</b>	19.4 ( $\pm$ 0.7)	34.5 ( $\pm$ 1.2)	17.8 ( $\pm$ 1.2)	28.3 ( $\pm$ 1.1)

*Table 1.* A breakdown of ‘winning units’ for each superordinate category within each of the 10 VPMs



*Figure 7* Examples of self organising feature map output contour maps for 4 training patterns (see appendix A for the actual images presented). Light areas indicate regions of high output activation, whereas dark areas indicate regions of low unit response. Note how the furniture training patterns generate sharper peaked contour maps than the animal training patterns.

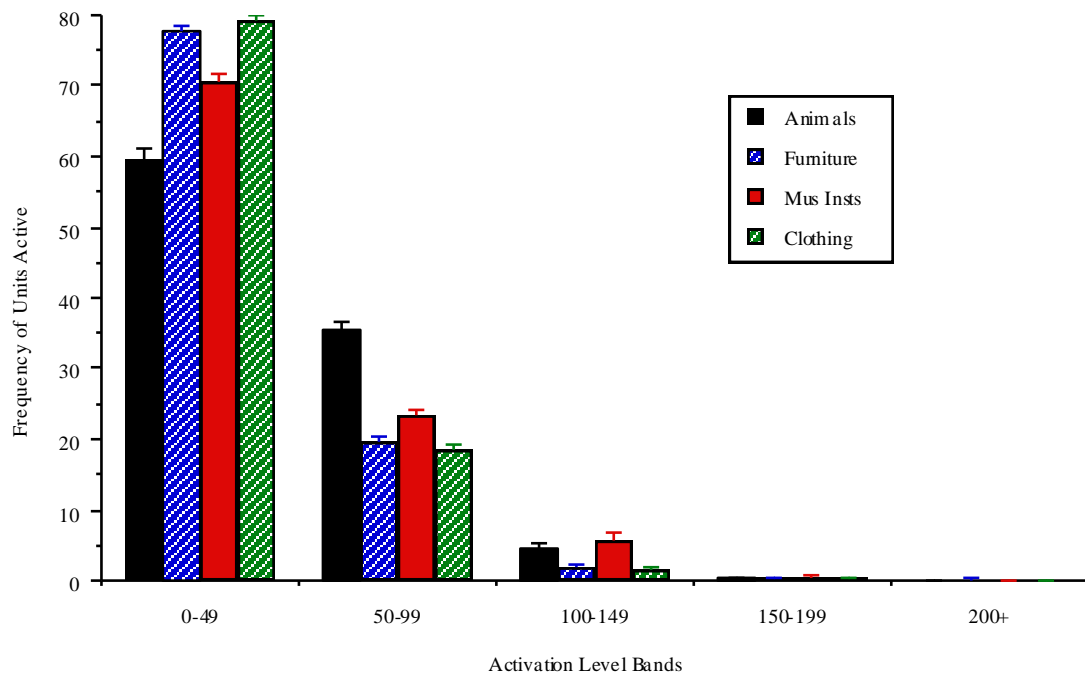


Figure 8 Frequency plot of activation values across a 100 unit Self Organising Feature Map for training patterns from 4 different semantic categories. Error bars indicate the standard deviations.

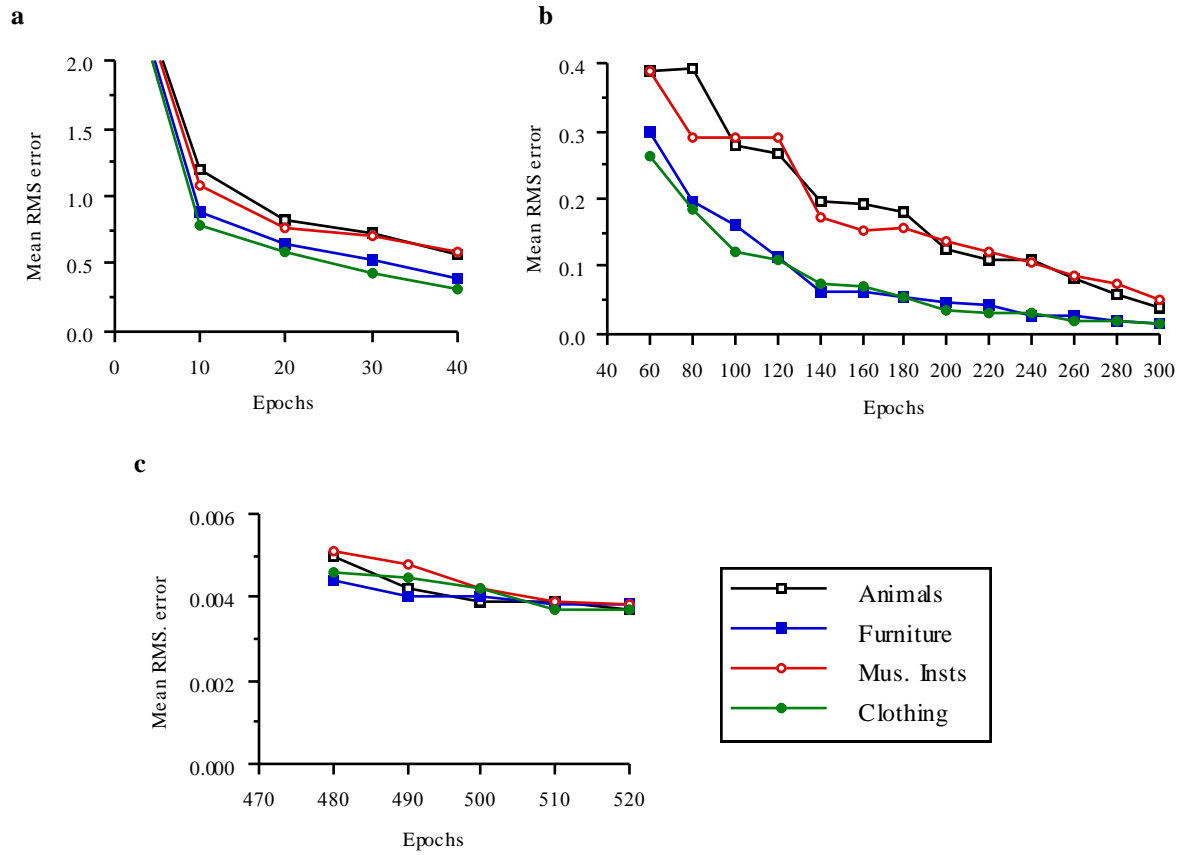


Figure 9 (a, b and c) Categorisation accuracy over time for training exemplars (i.e. VPM representations) from each of the 4 superordinate categories. Graphs a, b and c depict early, middle and late stages of training respectively.

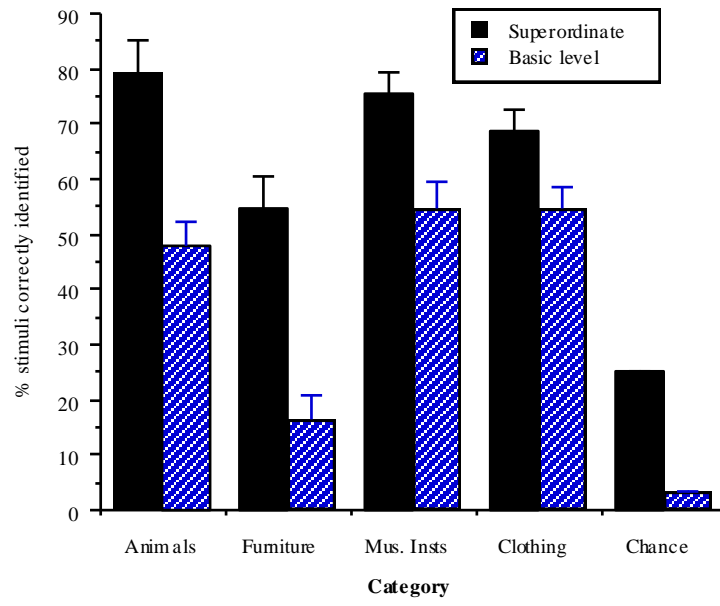
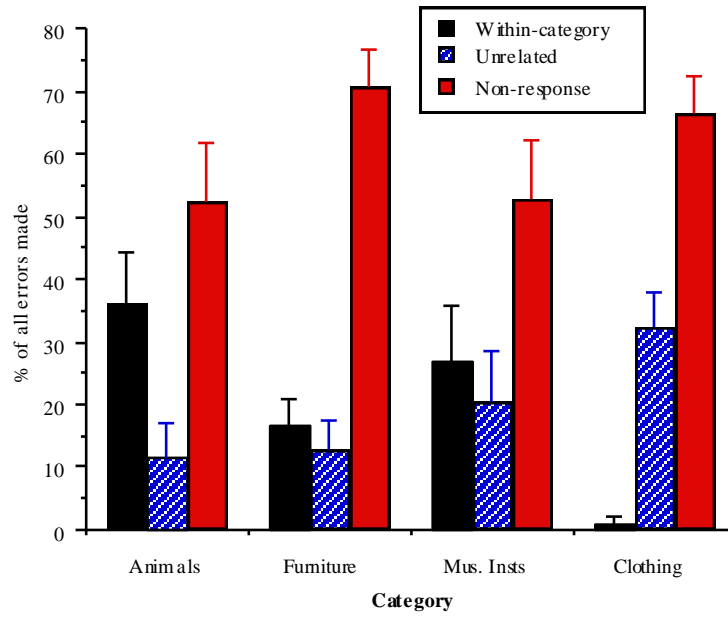


Figure 10 The percentage of novel stimuli identified correctly at the superordinate and basic level. Error bars indicate standard errors of the mean.



*Figure 11* The distribution of generalisation error types for each object class. Error bars display the standard errors of the means

<i>Exemplar</i>	<b>Binary Semantic Feature Vector</b>	<b>Feature Order</b>
Animal 1	<b>10101010 10101010</b> 11111111 00000000	<b>SO SO</b> BL BL
Animal 2	<b>10101010 10101010</b> 00001111 00001111	<b>SO SO</b> BL BL
Furniture 1	00000000 11111111 <b>01000111 10011100</b>	BL BL <b>SO SO</b>
Furniture 2	11001100 11001100 <b>01000111 10011100</b>	BL BL <b>SO SO</b>
Musical 1	00111100 <b>01010101 10111000</b> 00111100	BL <b>SO SO</b> BL
Musical 2	01100110 <b>01010101 10111000</b> 01100110	BL <b>SO SO</b> BL
Clothing 1	<b>01010101</b> 11011000 11001001 <b>01100011</b>	<b>SO</b> BL BL <b>SO</b>
Clothing 2	<b>01010101</b> 10101111 00001100 <b>01100011</b>	<b>SO</b> BL BL <b>SO</b>

*Table 2* Example semantic output vectors used in experiment 6. It can be seen that the group of 32 output units divide into 4 banks of 8 and the superordinate feature banks are denoted by bold typeface. In the example shown here, superordinate properties for animals are represented across the first 2 banks, whereas for furniture they are represented across the last 2. For musical instruments, they are represented across the second and third banks and for clothing, across the first and fourth banks (in the last column, the letters 'SO' and 'BL' represent superordinate and basic level properties respectively). Not only do these representations remove any architectural distinction between superordinate and basic level information, but they are perfectly counterbalanced across *all* training patterns such that no unit is biased towards assuming a particular value.



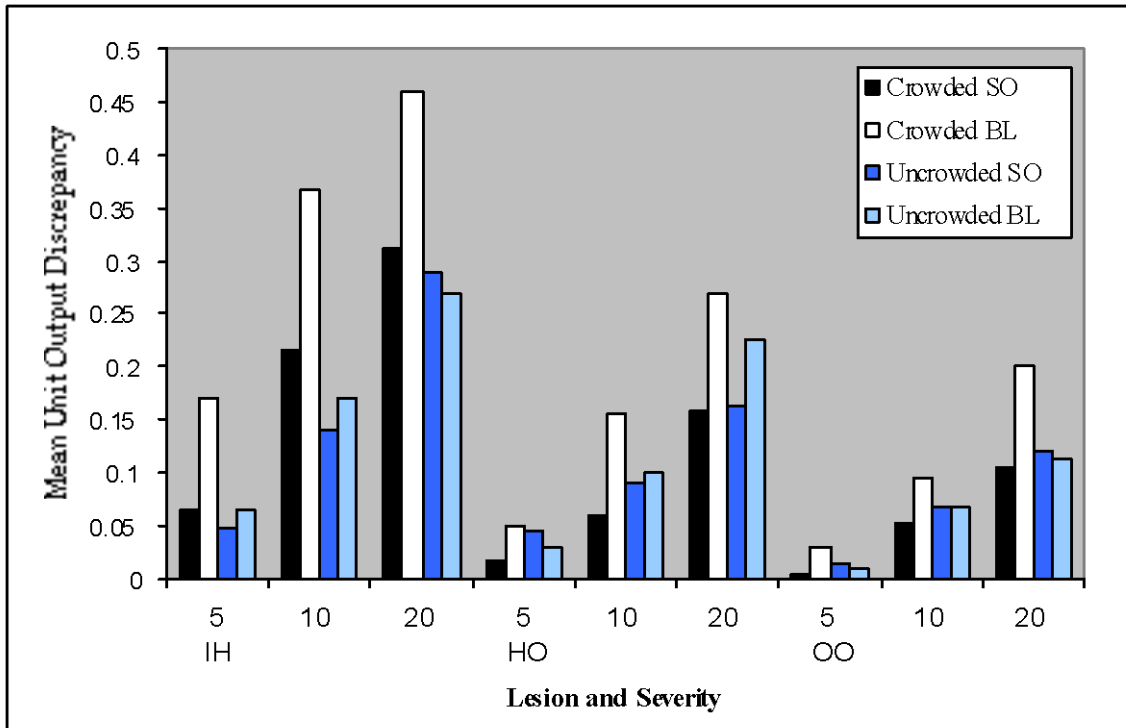


Figure 12 Mean Unit Output Discrepancies for units involved in the representation of superordinate (SO) and basic level (BL) information, for both Crowded and Uncrowded categories. Data are plotted to show the effects of incremental lesioning at each possible lesion site.

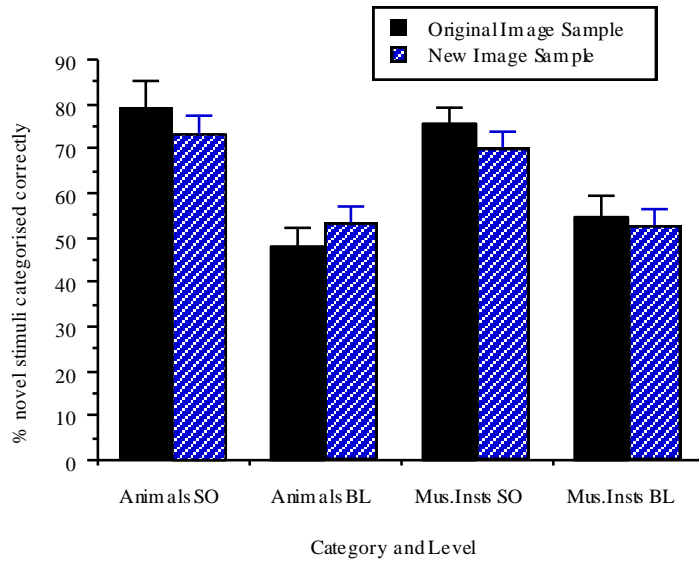


Figure 13 Consistency of model performance over 2 different image samples. Error bars indicate 95% confidence intervals