Max–Planck–Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 079

Categorical Perception of Familiar Objects

Fiona N. Newell[1] & Heinrich H. Bülthoff[2]

March 2000

Fiona Newell is now at:
Department of Psychology
Áras an Phiarsaigh
University of Dublin
Trinity College, Dublin 2 Ireland
Email: fnewell@tcd.ie

[1] AG Bülthoff, E–mail: fiona.newell@tuebingen.mpg.de
[2] AG Bülthoff, E–mail: heinrich.buelthoff@tuebingen.mpg.de

Categorical Perception of Familiar Objects

Fiona N. Newell & Heinrich H. Bülthoff

Abstract. We investigated whether familiar, 3-D objects are categorically perceived in the same way that other visual stimuli (e.g., colour and faces) are perceived as categorical. A continuum of shape change between fifteen pairs of objects was created and the images along the continuum were used as stimuli. In Experiment 1 participants were first required to discriminate pairs of images of objects that lay along the shape continuum. Then participants were asked to classify each morph-image into one of two pre-specified shapes. We found evidence for categorical perception in some but not all of our object pairs. In Experiment 2 we lloked in more detail at specific object pairs by increasing the difficulty of the discrimination task and found that more object pairs were categorically perceived. In Experiment 3 similarity ratings for each object pair were collected. These similarity scores correlated with the degree of perceptual categorisation found for the object pairs. Our findings suggest that familiar objects are perceived categorically and that categorical perception is closely tied to inter-object perceptual similarity.

1 Introduction

It is well documented that visual stimuli varying on a monotonic scale are often not perceived as gradually changing. Instead, the elements along this continuum are often perceived as belonging to discrete categories. For example, observers rarely report perceiving continuity of colour change along the wavelength but report "shifts" in colour categories from, say, red to orange to yellow etc. Facial identity and facial expressions are similarly perceived as discrete categories. These and other findings have suggested to researchers that the brain somehow categorises perceptually similar stimuli into qualitatively different categories to allow for more efficient processing of the perceptual world (see Harnad, 1987 for a review).

The world is filled with a rich variety of shapes, both man-made and living. Remarkably, by perceiving the similarities and the dissimilarities between objects, we can create classes of objects to effectively reduce the overwhelming number of entities in the world to more manageable proportions (Rosch, 1975; Rosch, Mervis, Gray, Johnson, and Boyes-Braem, 1976; Tversky and Hemenway, 1984). The number of possibilities of different object shapes in the world is endless yet each shape can be categorised into separate groups, even if the shapes are unfamiliar (Murphy, 1991). These findings suggest that there exists a specialised perceptual system that can group particular objects into different object classes based on their shape.

Recent research has found that recognition performance to objects is related to the degree of perceptual similarity between the items in a task, such that recognition is faster and more efficient when inter-object similarity is low and less efficient when inter-object similarity is high (Edelman, 1995a; Newell, 1997). For example, if the task involves recognising objects from different classes then object recognition can occur by building a description of the object based on its component parts (Biederman, 1987; Biederman and Gerhardstein, 1993; Marr, 1982, Marr and Nishihara, 1978). Thus recognition could be achieved when a generic description of the object's structure is built from the image and subsequently matches a memory representation of that object's description. Part-based descriptions of objects are often sufficient to discriminate between different object classes since these descriptions are mostly unique across classes (Biederman, 1987; Rosch et al, 1976). It is argued that much of object recognition occurs at the most general category level of abstraction, that is, the basic level, and is a fast and efficient process (Biederman, 1987; Jolicoeur, Gluck and Kosslyn, 1984; Rosch et al., 1976).

Yet basic-level recognition is not the only recognition task for the visual system and most recognition tasks involve discriminating between items within the same class (Bülthoff, Edelman and Tarr, 1995; Tarr and Bülthoff, 1995), and such items often share the same part-based descriptions (Tversky and Hemenway, 1984, Rosch et al, 1976). For example, such a task might be to recognise a Ford Escort from a Volkswagon Golf. The visual recognition system also needs to be able to discriminate between different exemplars within a category. When items within a task are highly similar, then recognition is less efficient, with performance often dependent on changes in the image characteristics such as viewpoint (Bülthoff and Edelman, 1991) or illumination (Tarr, Kersten and Bülthoff, 1998). Moreover, incidental changes in viewpoint produce a higher cost on recognition performance if the objects in the task are highly similar, and less pronounced if the objects are not very similar (Newell, 1997). In sum, research into object recognition also suggests that the recognition system is tuned to the specific shape characteristics of categories of objects and that representational object space is structured into classes of perceptually similar objects sharing a basic set of features (Edelman, 1995b).

Given the evidence that the object recognition system can distinguish between different object classes and objects within a class the question arises as to how such a system can work. Although many mechanisms can be proposed we will consider two, more obvious, mechanisms: First, the perceptual system may be tuned to a particular combination of features that represent a particular class of objects. In this case we would predict a qualitative difference in the way similar objects actually look to the perceiver. On the other hand, continuous changes between similar objects may be perceived directly and it is only the later semantic or verbal systems that are tuned to categorise these shapes. If this were the case then we can predict no effect of category membership at the perceptual level.

In the following experiments we investigated whether objects are indeed classified at the perceptual level using a paradigm often used in psychophysics called categorical perception (CP). The hallmarks of CP are usually twofold. First, the probability of identifying an object should not vary linearly along a continuum of shape change but should change relatively abruptly at the subjective category boundary. Second, pairs of shapes differing by the same physical amount should be more discriminable if they straddle this category boundary than if they lie within one category.

In the past, categorical perception within the visual domain has been demonstrated using continuous stimuli such as colours (Bornstein and Korda, 1984); facial identity (Beale and Keil, 1995), and facial expressions (Calder, Young, Perrett, Etcoff and Rowland, 1996; Etcoff and Magee, 1992). Calder et al. created photo-realistic sequences of morphed images between different expressions of a face. Subjects were first required to discriminate between pairs of images along each expression continuum that differed by equal physical increments between two different expressions. They were then required to label each morphed image as one of two pre-specified facial expressions. Calder et al. reported finding sharp, step-like functions categorising images into two different facial expressions. Moreover, they found an increased sensitivity to discriminating images that lay along the category boundary. Thus their study provided evidence that higher-order stimuli such as facial expressions are categorical at the perceptual level.

In the same way that facial expressions and facial identities are reported to promote categorical perception, we investigated whether objects are also categorical at the perceptual level. Previously, categorical effects have been demonstrated using stimuli that are highly similar, such as colours and facial expressions. However, inter-item similarity between objects can vary depending on whether the task involves objects which cross category boundaries or not. Consequently, there is no reason to believe that basic-level objects, for example, are perceived as qualitatively different at the perceptual level since each object category has a distinct shape and part-based description from the next object category (e.g., bottle versus car). These quantitative differences may, therefore, suffice for discriminating between objects (Biederman and Gerhardstein, 1993). However, smooth shape changes are likely to occur within an object category. For example, it would be possible to create a new set of bottles by averaging between two different bottles. Since bottles occur in all shapes and sizes in the real world then it might be important for the perceptual system to 'tag' similar bottle shapes together for identification purposes. This could occur through learning to identify the relevant dimensions on which the objects differ (Goldstone, 1994). This 'tagging' process would apply when discriminating between wine bottles and coke bottles, for example, but may be unnecessary when discriminating between a wine bottle and church bell. This is because the intrinsic quantitative differences between a wine bottle and church bell may suffice for efficient identification whereas the perceptual system may need to specify qualitative differences between a wine bottle and coke bottle in order for these shapes to be discriminated, thus resulting in effects of categorical perception.

Very little research has been conducted on the perceptual categorisation of shape. However, a recent investigation found evidence that local shape changes are categorically perceived (Mamassian, Kersten and Knill, 1996). In their study, Mamassian et al. asked participants to classify a specified local shape on an object surface as either hyperbolic or elliptical. Each local shape was a sample of a continuum of shape change along a surface of a 'croissant'-like object. They found that participants could reliably partition the local shapes into either hyperbolic or elliptical surfaces. A sharp step-like function relating number of responses to the continuum of local shape change marked the subjective category boundary that partitioned the two surface descriptions. The authors argued that this ability to correctly locate the partition between the classes of curved surfaces may be a property of the visual system that is used to, say, partition objects into their component parts. Their study also provided evidence (albeit indirectly, since discrimination sensitivity at the category boundary was not measured) that shape changes between similar object properties, such as from elliptical to hyperbolic surfaces, can be categorically perceived.

In Experiment 1 we investigated whether familiar objects were perceived categorically. Two main groups of objects were used as stimuli; objects from within the same basic-level category and objects from different basic-level categories. Categorical perception was tested between all possible pairs of objects within each group. We found evidence of categorical perception for some, but not all, of the object pairs. In Experiment 2 we increased the difficulty of the discrimination task used in Experiment 1 because of a concern that the discrimination task was not sensitive enough to show effects of categorical perception which may be present for some object pairs. Finally, in Experiment 3 we collected similarity ratings for the pairs of objects used as stimuli in our experiments in order to test for a possible correlation between pair similarity and categorical perception.

2 Experiment 1

In the following experiment participants performed two tasks; a discrimination task followed by an identification task. In the discrimination task participants had to discriminate between pairs of images that lay within the morphed continuum between two objects. These pairs of images were separated by equal physical increments along the shape continuum between two objects. Prior to the identification task participants were first presented with two shapes of objects and asked to memorise them. They were then shown single-image presentations of object shapes and were asked to decide which of two memorised objects each shape was more like. The discrimination (XAB) task assessed observers' ability to discriminate between object images that lay along a shape continuum and the identification task determined how observers classify these same object images.

2.1 Method

Participants

Forty-five undergraduate students from the Eberhard-Karls University of Tübingen, Germany participated in the following experiment for pay. Twenty-four of the participants were female. The participants' ages ranged from 18 to 28 years old. All participants had normal or corrected-to-normal vision.

Stimuli and Materials

The stimuli consisted of eleven morphed images from each of 15 pairs of objects. The objects were two different exemplars from 5 different basic-level categories. The categories, including the two exemplars, were as follows; BOTTLE (wine bottle, Coke bottle); LAMP (bedside lamp and desk lamp); GLASS (wine glass and beer glass); VASE (urn and single-stem vase); and BELL (church bell and hand bell). These objects were paired as five pairs of within-class objects and ten pairs of between-class objects. The between-class pairings constituted all combinations of category objects using one exemplar object from each class (e.g., wine bottle and hand bell).

The stimuli were generated using SoftImage|3*D* 3.7 software on a Silicon Graphics, Indigo2 workstation. This software package allows modelling and rendering of 3D objects. The ten basic object shapes were designed in the following manner: The objects were drawn as solids of revolution by defining the occluding contour of each object and rotating this contour around the object's elongated axis. All objects were designed with approximately the same aspect ratio of 2:1:1, with respect to the length of the elongated axis, the width and the depth of the object. The main elongated axis was the same length for all objects. The occluding contour was specified by a number of co-ordinates. All objects were created from the same basic number of co-ordinate points but the position (i.e., radial distance from the elongated axis and the position on the Y axis) of these co-ordinates was different for each object shape. The objects were designed using the same number of co-ordinates in order to allow for correspondence between the objects during shape interpolation. The interpolation routine was a SoftImage 3D, morphing algorithm applied on a pre-specified pair of 3D object shapes. This algorithm measures the distance between each of the corresponding co-ordinate positions on the objects and simulates an animation procedure by gradually moving each co-ordinate in the first object to the position of the corresponding co-ordinates in the second object. The number of steps taken to transform the shape of object 1 into the shape of object two was pre-specified. We specified 11 steps, therefore, 11 images were taken as output from the morphing procedure for each object pair. These 11 images were evenly spaced samples of the morphing routine from object 1 to object 2 along the shape continuum. See Figure 1 for an illustration of the morphing routine.

The objects were rendered under a perspective projection from their 3D models using the Silicon Graphics Indigo2 workstation with the SoftImage software. A shaded, 256 grey levels image was rendered by Lambertian shading by assuming a point light source at infinity, 45° up and 15° left from the line of sight and another light source (with the same intensity) on the viewpoint. The objects were presented in

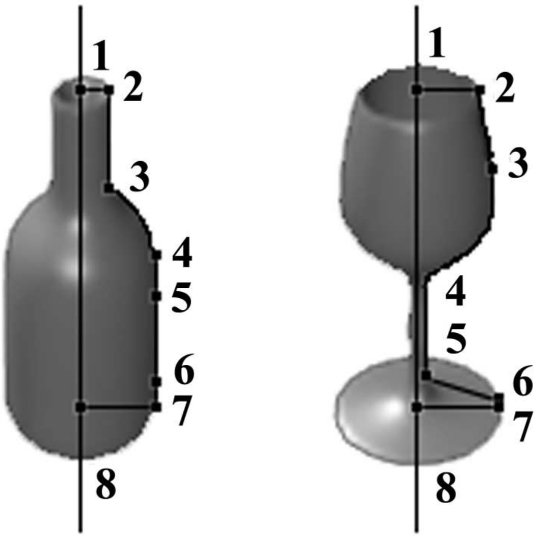

Figure 1: A schematic illustration of the nature of the morphing routine used to create the image morphs between two original object images. An object was drawn by rotating a contour (spline) around the main axis of elongation. In the figure above, the spline of a bottle and a glass is shown superimposed on the object shapes. The morphing procedure can be imagined as follows: The spline of the glass is redrawn from the spline of the bottle by hand. Then, the distance between each of the corresponding spline nodes (i.e. co-ordinates) between the objects is calculated. In our illustration the corresponding co-ordinates are marked by a number. Once the distance is calculated the position of each co-ordinate on the bottle is moved, linearly, to the position of the corresponding co-ordinate of the glass. The number of steps taken between the original position of the two co-ordinates is prespecified. A new 3-D object shape is generated at each step (see Figure 2).

a canonical view[1] against a black background. All images of the objects were 256 by 256 pixels in size.

The object images were transferred to a Macintosh Quadra computer and the stimuli were presented on a 256 colour, 16-in Macintosh monitor. The object images subtended, on average, a vertical visual angle of 5° and an horizontal visual angle of 3°. The experiment was run using PsyScope 1.04 presentation package for the Macintosh. Subjects used a button-box for responding.

Design

The experiment was divided into two separate tasks; an XAB discrimination task and an identification task. Subjects were divided into 5 groups of 9. Each group was tested on 3 pairs of objects, from the set of fifteen pairs, in both tasks. Two of these pairs were of objects from across different classes (e.g., lamp to bottle) and one was a within-class object pair (e.g., desk lamp to bedside lamp). All subjects assigned to a group saw the same 3 object pairings in both the discrimination and identification tasks. The images from each object-pair were blocked in both tasks and the order of presentation of the object pairs was counter-balanced across the subjects in each group. The order of the trials within each block was randomised across participants.

The discrimination task was based on an XAB design in which an image of the first object (stimulus X) was presented initially in a trial followed by the second and third images of objects (stimuli A and B) presented simultaneously, left and right of fixation. Stimuli A and B were always physically different from each other and stimulus X was identical to either stimulus A or B. Stimuli A and B differed by 2 steps (e.g., images 1 and 3) along the object-pair shape continuum. The order of the stimuli was counter-balanced which resulted in four orderings of any two stimuli (AAB, ABA, BAB, and BBA). In the identification task participants were presented with single images of the object-pairs from along the shape continuum.

There were $(3 * 9 * 4)$ 108 experimental trials in the discrimination task (object pairs, shape pairs, and XAB counter-balancing) and $(3 * 11)$ 33 experimental trials in the identification task (object-pairs and morphed images). We repeated the blocks of trials in the identification task four more times. These repeated trials were not included in the experimental analysis but were included as filler trials for the subject. The analysis included the first presentation of each trial. The total number of trials in the experiment was, therefore, 108 (discrimination task) and 165 (identification

---

[1] The term canonical view refers to the best and most familiar view of objects. See Palmer, Rosch and Chase (1981), Blanz, Tarr and Bülthoff (1999) for a further description.

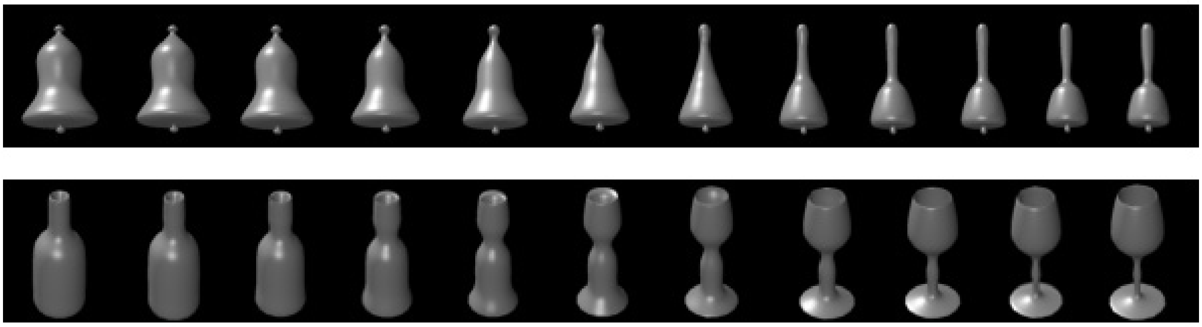

Figure 2: Illustration of two of the morphed objects in the experiment. On the top row a church bell is morphed into a hand bell (left to right). On the bottom row, a bottle is morphed into a glass (left to right). Objects are shown in the canonical view.

task).

Procedure

Participants were seated approximately 57cm away from the computer monitor. Each participant was required to perform first the discrimination task followed by the identification task with a self-timed break between the two tasks. The identification task determined the subjective category boundary for each object pair, hence it was conducted after the discrimination task in order to avoid any biasing during the discrimination task.

A fixation cross preceded the object stimuli in all tasks for 250ms. In the discrimination task, there were three object stimuli shown in any one trial. The first object image (X) was shown for 100ms in the centre of the screen, followed by a mask for 1 second. The next pair of stimuli (A and B) remained on the screen until the subject made a response. Each of the A and B stimuli were displayed 3cm to the left and right of the centre point of the screen. An inter-trial interval of 500ms followed the subject's response. In order to acquaint subjects with the XAB procedure in the discrimination task, the experiment began with a random selection of 20 practice trials. There were three blocks to the discrimination task and participants received a self-timed break between blocks.

In the identification task each trial began with a 250ms fixation cross. An object image then appeared and remained on the screen until the subject responded. An inter-trial interval of 500ms followed each subject's response. In the identification task the images from each object pair were presented in different blocks with a self-timed break between blocks. At the beginning of each block subjects saw two shapes of objects (which corresponded to the object shapes at the extreme end of each object-pair continuum). The subjects were instructed that each shape was associated to either the left or right button on the button box. (For example, the coke bottle shape belonged to the left button and the wine bottle shape to the right button.) Subjects were then instructed to decide as fast and as accurately as possible which of two object shapes each of the presented images looked more like. This procedure was repeated for each object-pair. For each participant, the object images in the identification task were always the same as those shown in the discrimination task. Subjects took approximately one hour to complete the experiment.

2.2 Results

The mean number of correct responses made to the identification tasks and the XAB tasks are shown in Figure 3. The subjective category boundary was determined by the identification performance. We determined the category boundary as the point at which the identification function crosses the 50% correct response level. We then conducted a one-way ANOVA between the discrimination performance to pairs of images that lay at either end of the shape continua (e.g., the average performance to image pairs 1-3 and 9-11) and the discrimination performance at the image pairs that straddled the category boundary. We used category position (2 levels: within and between) as the main factor. A main effect of category position was found, $[F(1,120) = 252.8,\ p < 0.001]$. Post-hoc, Newman-Keuls analyses revealed that the effect of category position was significant for all object pairs at $p < 0.05$ level of significance. The position of the category boundary and the F ratios are shown for each object pair in Table 1.

Finding a significant difference between the number of correct responses made to the images on the end of each object-pair continuum and those that straddle the category boundary is not, in itself, sufficient evidence that the object pairs are categorically perceived. For example, the discrimination may not just be best at the point of the subjective category boundary but may be equally good for other image pairs along the continuum (see objects Lamp-Bell in Figure 3 for an illustration of this point). In a second test for categorical perception, participants' discrimination performance can be predicted from the identification data. The method of deriving the predicted performance from the identification performance is widely used in the categorical perception literature (Calder et al., 1996; Liberman, Harris, Hoffman and Griffith, 1957). The performance predicts the likely outcome in the discrimination task provided the objects were categorically perceived. This predicted performance can then be correlated with participants' actual observed performance on the discrimination task. If the two are correlated for an object pair then we can say that those objects were perceived categorically.

For each object pair, participants' predicted performance for the discrimination task was derived from the procedure outlined in Appendix A (see Calder et al., 1996). The predicted performance and the observed performance are shown for each object pair in Figure 3. These predicted data were then correlated with the observed discrimination data. The results of these correlations are shown in Table 1. The results indicate that some, but not all, of the object pairs showed significant correlations between the predicted and observed discrimination performances. A significant correlation was taken as evidence for categorical perception of the object pair. The object pairs that showed correlations between the observed and predicted functions included the following: bottles, glasses, bells, lamps, vases, bottle-vase, vase-bell, glass-vase and bottle-glass. There was no correlation found between the following pairs of objects: lamp-vase, lamp-bell, bottle-bell, bottle-lamp, lamp-glass, glass-bell.

2.3 Discussion

We found evidence that familiar objects can be perceived categorically. However, not all of our object pairings were categorically perceived by the participants. First, all of the object pairs that belonged within the same semantic categories were perceived as categorical. Second, four out of ten of the between-class object pairs were perceived as categorical.

We originally argued that categorical perception may only arise to stimuli that are continuous in the real world. For example, the entire spectrum of colour in a rainbow is experienced as distinct colour bands. Similarly, changes in facial expression can be witnessed in any dynamic face. Moreover, it might even be plausible to consider that facial identity may be continuous, e.g., within families etc. Consequently we hypothesised that only images from object pairs that are possibly found in the real world would be perceived as categorical. In fact our results did not conform to this idea. Some object pairs from different classes were also perceived as categorical, therefore, basic-level class membership was not a necessary condition for categorical perception.

In the following experiments we explore possible reasons why we found that some ob-

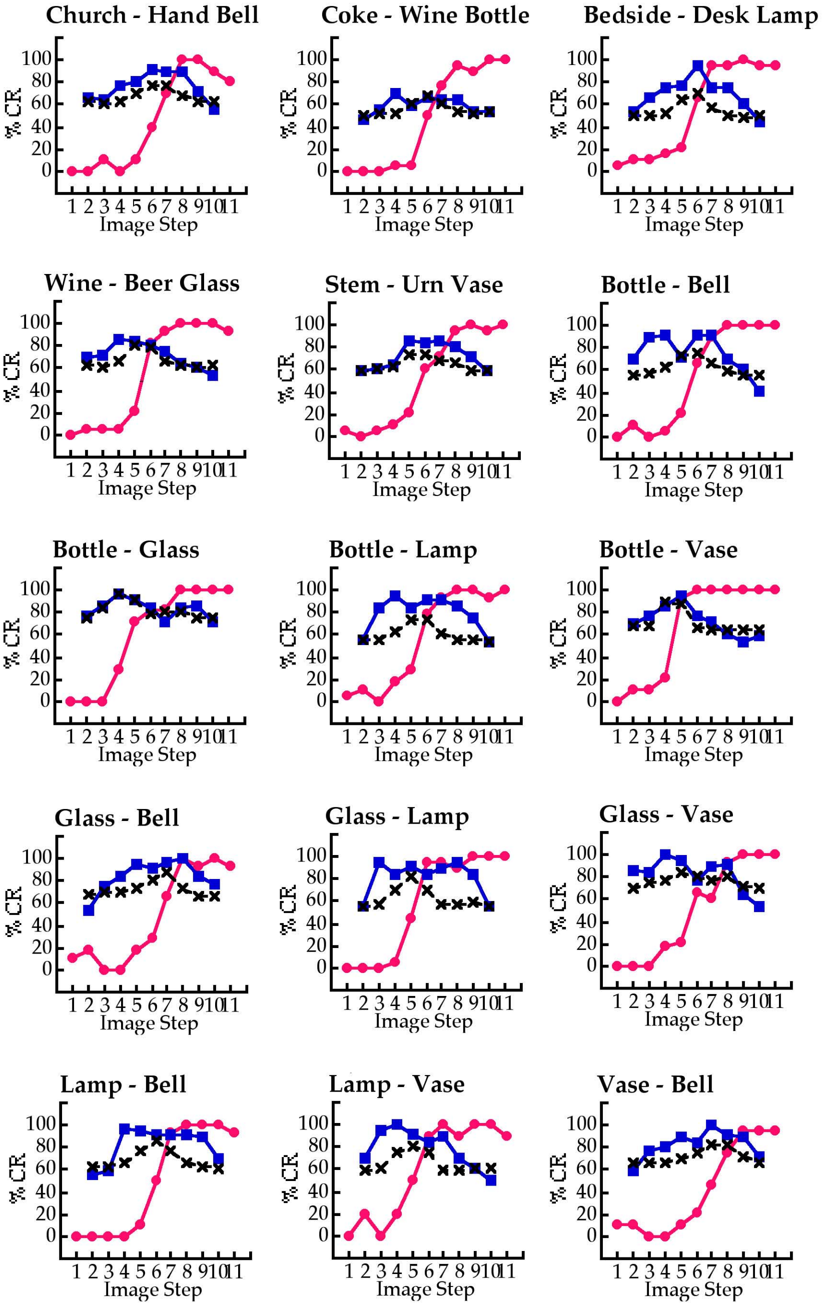

Figure 3: Plots showing identification and discrimination data for each of the 15 object pairs shown in Experiment 1. An explanation of the functions in each plot is as follows:
—●— The identification data show the mean percentage with which the first object image of the morph pair was correctly identified (e.g., church bell, coke bottle, bedside lamp etc.).
—■— The discrimination data show the mean percent correct (%CR) at discriminating the two object images at either side of each data point.
– – × – – The predicted discrimination data for each object pair. See Appendix A for a description of how the predicted curve is derived.

ject pairs were perceived categorically whereas others were not. In order to avoid possible ceiling effects in Experiment 2 we increased the difficulty of the discrimination task to those object pairs that were not perceived as categorical in Experiment 1. Finally, in Ex-

CLASS	OBJECT PAIRS	ANOVA		CORRELATION
		boundary	F-ratio	
Within	church bell - hand bell	6-8	F(1,120)=19.9 **	r=0.99 **
Within	coke bottle - wine bottle	5-7	F(1,120)=5.9 *	r=0.99 **
Within	bedside lamp - desk lamp	5-7	F(1,120)=44.7 **	r=0.75 *
Within	wine glass - beer glass	5-7	F(1,120)=10.5 **	r=0.99 **
Within	single-stem vase - urn	5-7	F(1,120)=13.3 **	r=0.86 **
Between	bottle - bell	5-7	F(1,120)=5.9 *	r=0.50
Between	bottle - glass	4-6	F(1,120)=5.9 *	r=0.79 *
Between	bottle -lamp	5-7	F(1,120)=18.1 **	r=0.53
Between	bottle - vase	4-6	F(1,120)=19.9 *	r=0.81 **
Between	glass - bell	6-8	F(1,120)=21.7 **	r=0.57
Between	glass - lamp	4-6	F(1,120)=21.7 **	r=0.36
Between	glass - vase	4-6	F(1,120)=13.3 **	r=0.86 **
Between	lamp - bell	5-7	F(1,120)=18.1 **	r=0.56
Between	lamp - vase	4-6	F(1,120)=21.7 **	r=0.54
Between	vase - bell	6-8	F(1,120)=25.6 **	r=0.76 *

Table 1: Table showing results of the one-way ANOVAs and correlations to each of the object pairs shown in Experiment 1. Objects pairs from the same class (Within) and object pairs from different classes (Between) are indicated in the first column (CLASS). The names of the object pairs are shown under OBJECT PAIRS. The F-ratio per object pair is the result of comparing the mean number of correct responses made to the object images at the extreme ends of each shape continuum to the number of correct responses made to the image pairs that straddle the category boundary, shown under ANOVA. The image numbers which straddle the category boundary are indicated for each object pair (boundary) and can also be seen in Figure 3. The table also shows the correlation between the observed discrimination performance and predicted discrimination performance for each object pair (CORRELATION).

periment 3 we investigated whether categorical perception was related to the visual similarity between the objects.

3 Experiment 2

In the following experiment participants were presented with a subset of the object pairs seen in Experiment 1. These object pairs included the pairs of objects where little or no evidence for categorical perception was found in Experiment 1. One of the possible reasons why no significant correlation was found between the observed discrimination data and the predicted data could be that the discrimination performance reflected a ceiling effect. The mean discrimination performance for the 9 object pairs that were perceived categorically was 74.2% as opposed to a mean discrimination performance of 80.1% for the object pairs that were not perceived categorically. The difference between these discrimination performances was almost significant [unpaired, two-tailed t-test: $t(13) = -1.728$, $p < 0.10$]. Although an error rate of approximately 20% is not generally considered as ceiling performance it was necessary to increase the difficulty of the discrimination task in order to rule out any possible ceiling effects which might have occurred to images that lay along the shape continuum. The plot of the glass-lamp pair of objects in Figure 3 illustrates this point: Although discrimination of pairs of images that lay at either extreme end of the continuum was low (mean 57%) the discrimination of pairs of images that lay between the extreme points was much better. In fact, the discrimination function is almost a straight line between the end pairs of images. We reasoned that an increase in the difficulty of the discrimination task may allow for effects of categorical perception to emerge because for some object pairs the discrimination task was not sensitive enough to show any effects of categorical perception.

The object pairs included in the following experiment were, therefore; bottle-bell, bottle-lamp, glass-bell, glass-lamp, lamp-bell, and lamp-vase. As already argued, one of the reasons why categorical perception might not have been observed for such object pairs was because the step size between the images along the shape continuum was sufficiently large to allow for easy discrimination. This possibility needed to be investigated. In the following experiment we decreased the step size between the images to one in the discrimination task (in Experiment 1 participants had to discriminate between images that lay two steps away on the shape continuum). We predicted that increasing the difficulty of the discrimination task would promote categorical perception.

3.1 Method

Participants

Sixteen students from the Eberhardt-Karls University of Tübingen participated in the following experiment for pay. Eleven of the participants were female. The participants' ages ranged from 21 to 30 years of age. All participants had normal or corrected-to-normal vision.

Materials and Apparatus

See Experiment 1

Stimuli

A subset of the pairs of objects used in Experiment 1 was used in the following experiment as stimuli. These object pairs included the following; bottle-bell, bottle-lamp, glass-bell, glass-lamp, lamp-bell, and lamp-vase. Shape continua (morphs) were created between each of these objects. There were 11 image steps in each shape continuum, including the original objects at each extreme (See Figure 2 for an example of the images generated).

Design

The discrimination task in this experiment was rendered more difficult than in Experiment 1 by using pairs of images that were closer together along the morph sequence (e.g. subjects discriminated between images 1-2 rather than 1-3 as in Experiment 1). In all other ways the experiment was similar to that in Experiment 1. As in Experiment 1 there were two parts to the following experiment. The first part, the discrimination task, was based on a two-way, mixed design with object pairs and image step as factors. Each participant was tested on three of the 6 object pairs. The three object pairs were randomly chosen for each participant. Images of the object pairs were blocked in each task. The order of the object pairs was counterbalanced and the order of the trials in each block was randomised across subjects.

Procedure

The general procedure was the same as that outlined in Experiment 1. All participants were presented with the identification task after the discrimination task was completed. Participants were allowed a self-timed break between blocks in each task. The experiment took approximately 1 hour for each participant to complete.

3.2 Results

The mean number of correct responses to the discrimination task for the object pairs seen in this experiment was decreased from 80.1% in Experiment 1 to 66.9% in the present experiment. This decrease was significant [unpaired t-test, $t(10) = 5.894$, $p < 0.001$]. We also found that the overall discrimination performance to the six object pairs in this experiment was not significantly different to the overall performance shown to the 9 object pairs which yielded effects of categorical perception in Experiment 1 [unpaired t-test, $t(13) = 2.038$, n.s.]. We were, therefore, assured that we had achieved the same level of difficulty in the discrimination task in this experiment as in Experiment 1.

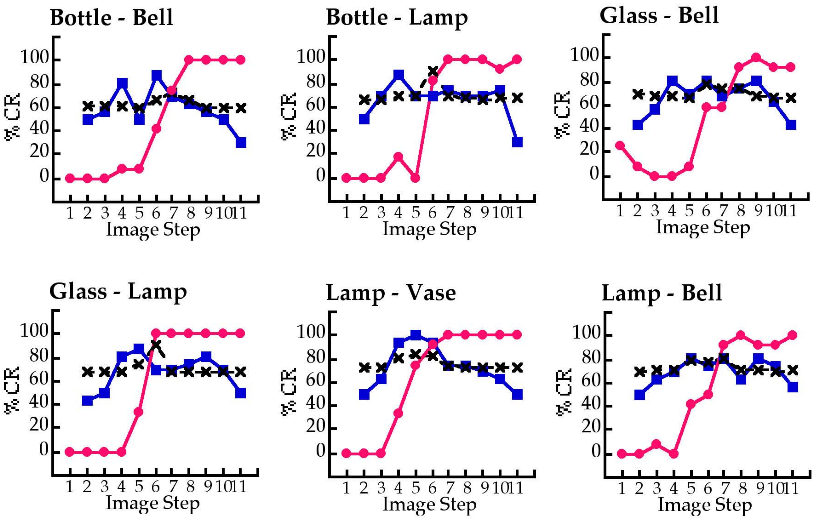

Figure 4: Plots showing identification and discrimination data for each of the 6 object pairs shown in Experiment 3. An explanation of the functions in each plot is as follows:
—●—The identification data show the mean percentage with which the first object image of the morph pair was correctly identified (e.g., church bell, coke bottle, bedside lamp etc.).
—■— The discrimination data show the mean percent correct (%CR) at discriminating the two object images at either side of each data point.
– – × – – The predicted discrimination data for each object pair (see Appendix A for the procedure).

The mean number of correct responses made to the identification tasks and the XAB tasks are shown in Figure 4. The subjective category boundary is determined by the identification performance. We calculated the category boundary as the point at which the identification function crosses the 50% correct response level. We then conducted a two-way ANOVA between the average discrimination performance to the image pairs at the extreme ends of each shape continua and the performance to images straddling the category boundary. We used objects as a between subjects factor and category position (2 levels: within and between) as the within subjects factor. An effect of object pairs was found, $[F(5, 42) = 2.447,\ p = 0.049]$. (A post-hoc, Newman-Keuls analysis revealed that the average discrimination performance to the lamp-vase pair was significantly lower than to the bottle-lamp and bottle-bell pairs.) A main effect of category position was found, $[F(1, 42) = 60.842,\ p < 0.001]$. Further Newman-Keuls analyses revealed that the effect of category position was significant for all object pairs at $p < 0.05$ level of significance. The position of the category boundary and the F ratios are shown for each object pair in Table 2.

We again employed the procedure used in Experiment 1 (see Appendix A) to predict the observers performance on the discrimination task from their performance on the identification task for each object pair. Both the predicted and the observed performance are shown in Figure 4. For each object pair the predicted performance was then correlated with the participants observed performance. The results of these correlations are shown in Table 2. We found that two of the six object pairs showed significant correlations between the observed and predicted data. These object pairs were lamp-bell and lamp-vase. There was no significant correlation between these functions for the following object pairs: bottle-bell, bottle-lamp, glass-bell, and lamp-glass.

3.3 Discussion

Although the difficulty of the discrimination task was increased in Experiment 2 we did not achieve perceptual categorisation for all

CLASS	OBJECT PAIRS	ANOVA		CORRELATION
		boundary	F-ratio	
Between	bottle-bell	5-7	F(1,42)=6.39 *	r=0.565
Between	bottle-lamp	5-7	F(1,42)=6.39 *	r=0.138
Between	glass-bell	5-7	F(1,42)=11.37 **	r=0.404
Between	glass-lamp	4-6	F(1,42)=13.34 **	r=0.183
Between	lamp-bell	5-7	F(1,42)=6.39 *	r=0.632 *
Between	lamp-vase	4-6	F(1,42)=20.21 **	r=0.895 **

Table 2: Table showing results of the one-way ANOVAs and correlations to each of the object pairs shown in Experiment 3. The F-ratio per object pair is the result of comparing the mean number of correct responses made to the object images at the extreme ends of each shape continuum to the number of correct responses made to the image pairs that straddle the category boundary. The category boundary is indicated for each object pair and can also be seen in Figure 4. The table also shows the correlation between the observed discrimination performance and predicted discrimination performance for each object pair.

object pairs. Here we found that only two of the six object pairs were perceived categorically: lamp-bell and lamp-vase. The other object pairs (bottle-bell, bottle-lamp, glass-bell, glass-lamp) were not found to be perceived categorically with an increase in difficulty in the discrimination task. It is noted that for these particular object pairs the overall discrimination performance was slightly lower than the performance to these same pairs in Experiment 1 which indicates that the lack of categorical perception found is not due to ceiling effects. However, the question arises as to why some object pairs were not found to be perceived categorically. Furthermore, we could ask why some object pairs are categorically perceived when the discrimination task is at a certain level of difficulty whereas others are not. In Experiment 3 we explore the idea that categorical perception may be related to the level of inter-object similarity between the objects.

4 Experiment 3

The results from the previous experiment indicated that some object pairs were perceived categorically by the observers whereas there was no evidence for categorical perception for other objects. One point of note from these results was that the objects that were categorically perceived were objects that appeared to be visually similar to each other. For example, objects from within the same basic-level category tend to appear similar to each other rather than objects from different categories (Rosch et al., 1976). In our results object-pairs from the same basic level categories were all perceived categorically whereas only some object-pairs from different categories were perceived categorically. These data beg the question whether visual similarity can effect the perceptual categorisation of objects. To this end we asked a number of participants to rate each object pair in Experiment 1 in terms of how visually similar the objects were to each other. We then correlated the mean rated similarity scores with the amount of categorical perception found for each object pair (i.e. the size of correlation between the observed and predicted functions for the discrimination data).

4.1 Method

Participants. Thirty students from the Eberhard-Karls University of Tübingen participated in this study. Fourteen of these participants were female. The participants' ages ranged from 18 to 28 years old. All participants had normal or corrected-to-normal vision. The students had not participated in any of the previous experiments.

Stimuli

Each of the 15 pairs of objects located at the extreme of each of the shape continua described above was organised into a 5x3 object-pair matrix. The images of the objects were greyscale, shaded images (see Experiment 1 for a description of the image rendering). The object matrix was then printed out for each participant using a colour Laserprinter with (600) dpi resolution. The order of the object pairs on each printout was randomised across all participants.

Procedure

Subjects were presented with a printout of a matrix of object pairs. They were instructed (in German) to study each pair of objects individually and to rate each pair of objects according to how similar the objects were to each other. Subjects were instructed to base their judgements of similarity on the shape of the objects only and to ignore other properties such as the object's name or function. A rating scale was demonstrated to the participants. This scale ranged from 1 (very dissimilar) to 7 (very similar). Participants were encouraged to use all of the scale in their judgements.

4.2 Results and Discussion

Participants' mean similarity ratings for each of the object pairs are shown in Figure 5. Kendal's coefficient of concordance was calculated across the participants ratings (0.406) which indicated a significant concordance across participants' rating scores, Chi^2 (N=30, df=14) = 170.46, $p < 0.01$.

The similarity ratings for each object pair were then correlated with the degree of correlation found between the predicted and observed discrimination data in Experiment 1. We found that the similarity ratings were highly correlated with the amount of categorical perception in each object pair, r=0.743, $p < 0.01$. The results suggested that inter-object perceptual similarity has a role to play in whether the objects are categorically perceived or not.

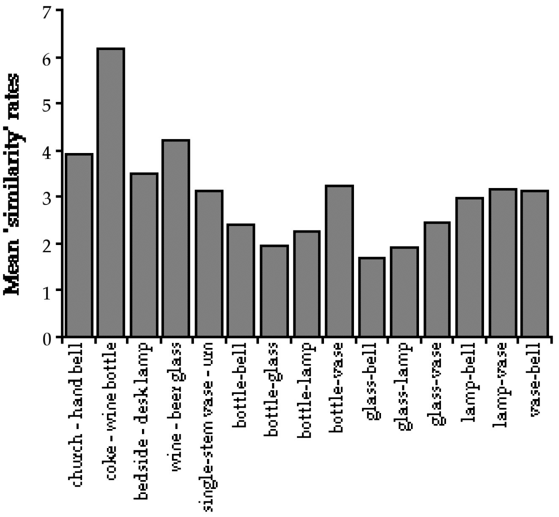

Figure 5: Plot showing participants mean rated scores their judged visual similarity between the objects in each object pair. Subjects used a scale from 7 (very similar) to 1 (very dissimilar). To allow for easier comparison the object pairs are arranged in the same order as the object plots shown in Figure 3.

5 General Discussion

In the experiments reported above we investigated whether objects are categorically perceived. Specifically we asked whether the shape of an object has a psychological salience in representational space that is not completely determined by its physical difference from other shapes. In other words, each object belonging to a single category may be perceived as qualitatively different from another object such that a continuum of shape change between two objects would result in a sharp discontinuity of categorisation performance, reflecting the category boundary between the two objects. Alternatively, the shape description of an object from a single category may not be qualitatively different from an object within the same category and in this case we would observe identification performance varying as a smooth continuous function between the two shapes. Our findings suggest that some familiar objects are indeed perceived categorically.

In Experiment 1, fifteen object pairs were tested, five of which were pairs from within the same category and ten of which were pairs of objects from different, basic-level categories. We found that for the within-category pairs the objects were perceived as categorical. Furthermore, some of the between-category object pairs were also categorically perceived. We found no evidence for categorical perception in six of the ten between-category object pairs. In Experiment 2 we increased the level of difficulty of the discrimination task for those six object pairs where no categorical perception was previously found. Our data showed that the change in task difficulty promoted categorical perception in two of the six object pairs. Finally, in Experiment 3 we found a correlation between rated similarity judgements and the degree of categorical perception found for each object pair.

The first question that needs to be addressed was why some object pairs were categorically perceived whereas others were not. We initially hypothesised that only objects which belonged to the same semantic class would be perceived as categorical. We reasoned that categorical perception is most likely to be found for objects that can exist as continuous shapes within a class in the real world because categorical perception may be a procedure for differentiating between similar objects. Furthermore, a continuum of shape change between objects from different classes is unlikely to occur in the real world. However, we found no evidence for this prediction since objects from across different basic-level classes were also perceived as categorical.

We found some evidence that categorical perception was dependent on the difficulty of the discrimination task. In Experiment 1 the discrimination task was found not to be sensitive enough to allow for effects of categorical perception to emerge in some of the object pairs. More objects were perceived as categorical in Experiment 2 where we increased the difficulty of the discrimination task. Of course the possibility remains that the level of difficulty in Experiment 2 was also not sensitive enough to allow for effects of categorical perception to emerge in all other objects. However, this is unlikely since the average discrimination performance was not at ceiling (on average 67% correct responses), suggesting that if any effects of categorical perception were present then the level of discrimination was not the factor obscuring such effects.

A further possibility why some object pairs were not perceived as categorical is that the morphing procedure itself may have introduced sufficient noise so that any effects of categorical perception that might be present were obscured. For example, morphing between two objects from different categories may not only produce a continuum of shape change between the two objects but other categories of objects may, as a consequence, emerge along this continuum. Consider morphing between a cube and a sphere. Somewhere along this morph sequence an octagon may emerge. This octagon may constitute a distinct perceptual category for the observer. Consequently, for some object pairs the task may produce effects that reflect the perceptual categorisation of three (or more) objects. The data will not show such categorisation because the participant was required to make binary decisions in both the identification and discrimination tasks. The result could, therefore, be an increase in the noise in the data such that categorical perception is not found. Clearly further research is required to investigate the effects of emergent categories on perceptual categorisation.

An increase in the familiarity of the items in a task can also emphasise physical differences between categories thus producing effects of categorical perception that did not exist prior to training. Goldstone (1994) found evidence for an increase in distinctiveness along a dimension which was relevant to the categorisation task. Conversely, he found acquired equivalence for items that belonged within the same category, i.e. that varied along a category-irrelevant dimension. Similarly, Livingston, Andrews and Harnad (1998) recently reported that learning produced within category compression effects (i.e. that inter-item similarity increases when categories are learned), and cross-category expansion effects. However, their effects were observed only when the categories were defined along many dimensions at once and not when the categories were defined along a single dimension.

Several mechanisms have been proposed to account for effects of learning on categorical perception (see Goldstone, 1998 for a review). Categorisation can be acquired by selectively attending to the relevant dimension along which the categories differ (Nosofsky, 1986) or to certain features by which they differ (Livingston and Andrews, 1995). Or the system may develop special detectors early on which looks for features to categorise the items. This process is referred to as imprinting and an example of such a process might be the face recognition system (Perrett et al, 1984). Finally, a process called differentiation may occur which effectively separates items that are psychologically equivalent and once they are separated, fine discriminations can be made between items that were originally indistinguishable. Goldstone offers the example of wine tasters who can distinguish between the upper and lower halves of a bottle of Madeira by taste to illustrate differentiation.

We might argue that our findings were due to effects of differentiation rather than other effects, such as selective attention, or imprinting. For example, the reason why some objects and not others were categorically perceived is that categorical perception is related to the degree of perceptual similarity between the objects such that objects that are highly similar are more likely to be perceived as categorical. Therefore, objects that were once confusable or indistinguishable because they are highly similar have become separated to allow for categorisation. We found that subjective similarity ratings were correlated with the degree of perceptual categorisation in the objects. The results of Experiment 2 also indirectly support this argument in the following manner: By changing the difficulty of discrimination in Experiment 2 it could be argued that we were effectively increasing the similarity between the objects. Objects are either intrinsically similar, such as objects from within the same category (Rosch, 1975), or can be made more similar by changing the difficulty of the discrimination task. For example, images 1 and 2 are more similar to each other on the shape continuum than images 1 and 3 and consequently the discrimination task is more difficult between images 1 and 2. Changing the level of discrimination in a task does not increase the similarity between the original end objects in the shape continuum. However, if these two end objects are quite dissimilar then discriminating between image 1 and 2 on the shape continuum may be equivalent to discriminating between images 1 and 3 on a shape continuum between two more similar end objects. Therefore, the distances between two images in the discrimination task are equivalent across object pairs in representational similarity space.

If categorical perception occurs when inter-object similarity is high, we could ask why the recognition of similar objects is less efficient than the recognition of objects from different basic-levels. For example, the recognition of objects from within the same class is often found to be dependent on viewpoint (Bülthoff and Edelman, 1992; Bülthoff, Edelman and Tarr, 1995) whereas the recognition of objects from different basic-levels categories is not (Biederman and Gerhardstein, 1993; Newell, 1997). If categorical perception is a tool for qualifying differences between similar objects then it might be argued that discriminating between category members might be just as good as discriminating between objects from different categories. There may be two reasons why discriminating between objects within a class is relatively less efficient: First, the specific objects within a class may be less familiar than the generic category they belong to. There is much evidence that an increase in object familiarity improves recognition performance (Edelman and Bülthoff, 1992; Newell and Findlay, 1997). This may be because subjects learn to create the appropriate object categories for the task thus optimising discrimination performance (Nosofsky, 1984). Beale and Keil (1995) reported finding categorical perception for familiar faces but not for unfamiliar faces, suggesting that perceptual categories may be learned. Second, objects from different categories are often differentiated by the structure of their parts, which is unique to each object class (Biederman, 1987). Objects from the same category, on the other hand, are often differentiated by small metric differences between the parts and the detection of such small differences may require extra visual processing in order to reliably discriminate between the objects. The performance differences found in our study may, therefore, reflect the amount of processing required to complete each task.

5.1 Implications for Representational Object Space

Recent models of object recognition have suggested that objects are represented as multiple views (Bülthoff and Edelman, 1992) and that the inter-object similarity determines the location of the object in representational space (Edelman, 1995b and c; Valentine, 1991). Edelman suggests that objects which are highly similar lie in closer proximity in representational space than objects that are less similar. To avoid confusion between objects which closely resemble each other, the perceptual system must construct category boundaries between similar object classes. Consequently all different types of round orange objects with bumpy surfaces will be categorised as the fruit orange whereas a basketball will not. Our findings propose that the differences between similar objects are qualified at the perceptual level.

If, as has been argued, objects are located in representational space based on their inter-item similarity then the visual system is faced with the problem of maintaining the uniqueness of an object despite changes with certain transformations such as changes in viewpoint, size or illumination. In other words the visual system must categorise different instances of an object as belonging to that one object. At the same time the differences between objects must be maintained. We suggest that such a process occurs at the perceptual level. If so, then effects of categorical perception should generalise across different instances of objects such as different viewpoints or different sizes. A more thorough investigation of the specifics of the structure of representational object space is a topic currently under investigation.

Acknowledgments

We thank Isabelle Bülthoff for her help with SoftImage|$3D$ and Nicole Köhnen for help in running Experiment 3. The research reported in this article was supported by the Durham University Research Foundation, Human Frontiers Science Program and The Max-Planck Society.

Correspondence concerning this article should be sent to Fiona Newell, Department of Psychology, Áras an Phiarsaigh, University of Dublin, Trinity College, Dublin 2, Ireland. E-mail: fnewell@tcd.ie.

References

Beale, J.M. and Keil, F.C. (1995). Categorical effects in the perception of faces. Cognition, **57**, 217-239.

Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. Psychological Review, **94,(2)**, 115-147.

Biederman, I. & P. C. Gerhardstein (1993). Recognising Depth-Rotated Objects: Evidence for 3D Viewpoint Invariance. Journal of Experimental Psychology, Human Perception and Performance, **19, (6)**, 1162-1182.

Blanz, V., Tarr, M.J., Bülthoff (1999) What object attributes determine canonical views? Perception, **28**, 575-599.

Bornstein, M.H. and Korda, N.O., (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. Psychological Research, **46**, 207-222.

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

Bülthoff, H.H., Edelman, S. and Tarr, M.J. (1995). How are three-dimensional objects represented in the brain? Cerebral Cortex, **5**, 247-260.

Calder, A.J., Young, A.W., Perrett, D.I., Etcoff, N.L., Rowland, D. (1996). Categorical perception of morphed facial expressions. Visual Cognition, **3(2)**, 81-117.

Edelman, S. (1995a). Class similarity and view-point invariance in the recognition of 3D objects. Biological Cybernetics, **72**, 207-220.

Edelman, S. (1995b). Representation, similarity and the chorus of prototypes. Minds and Machines, **5**, 45-68.

Edelman, S. (1995c). Representation of similarity in Three-Dimensional Object Discrimination. Neural Computation, **7**, 408-423.

Edelman, S. and Bülthoff, H.H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vision Research, **32**, 2385-2400.

Etcoff, N.L. and Magee, J.J. (1992). Categorical perception of facial expressions. Cognition, **44**, 227-240.

Farah, M.J. (1990). Visual agnosia: Disorders of object recognition and what they tell us about normal vision. The MIT Press, Cambridge, Massachusetts, USA.

Goldstone, R.L. (1994). Influences of categorisation on perceptual discrimination. Journal of Experimental Psychology: General, **123(2)**, 178-200.

Goldstone, R.L. (1998). Perceptual Learning. Annual Review of Psychology, **49**, 585-612.

Harnad, S. (Ed.) (1987). Categorical perception: The groundwork of cognition. Cambridge University Press, Cambridge, UK.

Joilcoeur, P., Gluck, M.A. and Kosslyn, S.M. (1984). Pictures and names: Making the connection. Cognitive Psychology, **16**, 243-275.

Liberman, A.M., Harris, K.S., Hoffman, H.S. and Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, **54(5)**, 358-368.

Mamassian, P., Kersten, D. and Knill, D.C. (1996). Categorical local-shape perception. Perception, **25**, 95-107.

Marr, D. (1982). Vision. Freeman, San Francisco, USA.

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, B, **200**, 269-294.

McCarthy, R.A. and Warrington, E.K. (1986). Visual associative agnosia: A clinical-anatomical study of a single case. Journal of Neurology, Neurosurgery, and Psychiatry, **49**, 1233-1240.

Murphy, G.L. (1991). Parts in object concepts: Experiments with artificial categories. Memory and Cognition, **19(5)**, 423-438.

Newell, F.N. (1998). Stimulus context and view dependence in object recognition. Perception, **27**, 47-68.

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

Nosofsky, R.M. (1984). Choice, similarity and the context theory of classification. Journal of Experimental Psychology: Learning, Memory and Cognition, **10**, 104-114.

Palmer, S. E., Rosch, E. and Chase, P. (1981). Canonical perspective and the perception of objects. In Long, J. and Baddeley, A.D. Attention and Performance, LEA. Hillsdale, N.J.

Perrett, D.I., Rolls, E.T. and Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. Experimental Brain Research, **47**, 329-342.

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

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.

Sergent, J., Ohta, S., MacDonald, B. (1992). Functional neuroanatomy of face and object processing. Brain, **115**, 15-36.

Tanaka, K. (1992). Inferotemporal cortex and higher visual functions. Current Biology, **2**, 502-505.

Tarr, M.J. and Bülthoff, H.H. (1995): Is human object recognition better described by geon-structural-descriptions or by multiple-views? Comment on Biederman and Gerhardstein (1993). Journal of Experimental Psychology: Human Perception and Performance, **21(6)**, 1494-1505.

Tarr, M.J., Kersten, D. and Bülthoff, H.H. (1998). Why the visual recognition system might encode the effects of illumination. Vision Research, **28**, 2259-2275.

Tversky, B. and Hemenway, K. (1984). Objects, parts and categories. Journal of Experimental Psychology, General, **113(2)**, 169-191.

Valentine, T. (1991). A Unified account of the effects of distinctiveness, Inversion, and Race in Face Recognition. The Quarterly Journal of Experimental Psychology, **43A(2)**, 161-204.

Appendix A

The following procedure was used to calculate participants predicted discrimination performance from their identification performance: This procedure is adapted from that used by Calder et al. (1996) and Liberman, Harris, Hoffman, and Griffith, (1957). First, the observers ability to discriminate between images that differ by a constant physical amount was determined by calculating the mean discrimination performance to the pairs of images that lay at either end of the object shape continuum (e.g., discriminating between image pairs 1,3 and image pairs 9,11). These images were assigned to their appropriate categories approximately 100% of the time therefore any effects of categorical perception would be minimised. Second, in order to predict the effects of categorical perception on the discrimination of image pairs that lay along the entire continuum the following calculation was conducted: We calculated the difference between the number of times each image in the pair was assigned to a particular object shape in the identification task. These figures were then added together. We then multiplied the identification difference by a constant of 0.3. This had the effect of aligning the predicted function with the observed discrimination function and also brought the respective range of variabilities more comparable. The predicted and observed functions are shown in Figures 3 and 4. The constant had no effect on the correlation between the observed and predicted curves but did help to show the fit between the two curves. In summary, the function used to calculate the predicted discrimination performance for each object pair was the sum of the mean discrimination performance to the pairs of images on either end of the shape continuum and 0.3 of the identification difference for each pair of images tested.