VISUAL COGNITION, 2004, *11* (1), 29–69

Face-space-R: Towards a unified account of face recognition

Michael B. Lewis
Cardiff University, UK

Versions of the ''face space'' are considered and built upon to develop an explicitly defined model of face recognition based on stimulus generalization that is similar to models of animal learning. This face-space-R model is implemented using realistic numbers of known faces. The model is able to account for distinctiveness, caricature, and race effects. It also predicts which faces will be falsely recognized and accounts for mirror effects. The application of the model to face learning and development is considered, as well as the effects of brief presentation. By varying parameters of the model, it is possible to match its performance to that of humans, leading to an estimate of the dimensionality of face space (of between 15 and 22 dimensions for same-race faces).

The ''face space'' has been an important concept for the investigation of how we recognize faces (see Valentine, 2001 for a review). One influential formulation of the face space was described by Valentine (1991), although such a space had been implicit in many earlier accounts of face processing (Brennan, 1985; Goldstein & Chance, 1980; Reed, 1972; Rhodes, 1988; Valentine & Bruce, 1986). A face space is a mechanism or construct used to generate an internal representation of perceived and stored faces. As such, a face space must be able to encode features of faces so that they can be uniquely discriminated and associated with a particular identity where appropriate. As Craw (1995) demonstrates, such face spaces are an essential part of any system that can recognize faces, be it human or

Please address all correspondence to: M. Lewis, School of Psychology, Cardiff University, PO Box 901, Cardiff, CF10 3YG, UK. E-mail: LewisMB@Cardiff.ac.uk

I would like to thank Hadyn Ellis, Robert Johnston, Tim Valentine, Mike Burton, and two anonymous reviewers for their useful comments on earlier drafts. This research was supported by EPSRC grant GR/R37777.

http://www.tandf.co.uk/journals/pp/13506285.html DOI:10.1080/13506280344000194

automatic.[1] It is not necessarily the case, of course, that face spaces employed by computer systems have the same properties as human face spaces; however, they are similar in the fact that they are each populated by stored vectors against which perceived probe-face vectors can be compared and evaluated. Establishing that a probe face has a particular identity involves an evaluation of the similarity between the vector of the probe face and stored exemplar vector representations of ''known'' faces. How this similarity is calculated and how a response is made defines the properties of the face space.

There are many types of face spaces. These differ in their formulation, often in quite subtle ways, to account for a wide range of the empirical evidence from the face recognition literature. The main versions of face space are reviewed below and their strengths and weaknesses are considered. A further revision to the face-space concept is developed and referred to as ''face-space-revised'' (or face-space-R). Here, it is illustrated how this new model is a development from earlier versions of face space, and how it has many things in common with them. The main difference between face-space-R and the earlier incarnations is that ease of recognition is explicitly defined and based on *all* known exemplars, rather than a limited subset. This ease-of-recognition calculation is derived from a stimulus–discrimination model taken from the animal-learning literature. In order to evaluate face-space-R as a model of face recognition, a series of computational implementations were performed. It is described how these lead to an estimate for the number of dimensions that a typical face space may possess and to explanations for aspects of face recognition, including distinctiveness effects, caricature effects and race effects. Finally, the model is evaluated with regard to false positive recognition of faces (i.e., the belief that someone is familiar when they are not—an occurrence that can have disastrous results when it occurs in the forensic setting).

THE EXEMPLAR-BASED FACE SPACE

In Valentine's (1991) unified account of distinctiveness, inversion, and race, an exemplar-based face space was proposed that could account for many documented effects from the face-recognition literature. These effects included: The faster and more accurate recognition of distinctive compared with typical faces (Light, Kayra-Stuart, & Hollander, 1979; Valentine & Bruce, 1986), along with

[1] Craw (1995) discusses the face space in terms of a manifold. A manifold is a vector space in its most general form without any assumptions. Any representation can be described as a projection onto a manifold. It is the properties of the manifold that define the face space. Even a system that stores and retrieves individual templates for each face can still be considered to be represented by a face manifold albeit, a non-Euclidean manifold.

the faster face categorization for typical compared with distinctive faces (Valentine & Bruce, 1986 using a task where an image had to be judged as a face or a jumbled faces). It also accounted for the influence of inversion and race on distinctiveness, and why faces from other races are harder to recognize (Goldstein & Chance, 1980).

Valentine's (1991) face space had two important assumptions, which will be carried into the current face-space-R model. The first assumption is that the similarity between representations in the face space is based on a Euclidean metric (where similarity is an analogue of distance in the real world). While this is an important assumption, it has little empirical backing; nevertheless, it has remained relatively unchallenged (although see Craw, 1995). This lack of any challenge is probably because non-Euclidean similarity metrics are difficult to envisage and their consequences are uncertain. In addition, it can be stated that a non-Euclidean space will behave like a Euclidean space over small distances. As, in general, one is attempting to distinguish between relatively similar exemplars then even a non-Euclidean face space may behave like a Euclidean face space in the majority of cases.

The second assumption of Valentine's face-space model is that, if one samples faces from a homogeneous population (i.e., faces from a single race), then the vector representations of these faces will describe a normal distribution along any dimension. This assumption is borne out by the fact that faces are biological constructs and the majority of biological entities vary along normal distributions. The central-limit theorem can account for this.[2] Further, measurement of facial features has not revealed that they deviate from a normal distribution (e.g., Burton, Bruce, & Dench, 1994; Shepherd, Ellis, & Davis, 1977). A consequence of this multivariate normal distribution of faces is that the face space will have densely packed exemplars close to the average face (i.e., typical faces) and sparsely spaced exemplars away from the average face (i.e., distinctive faces). More densely packed exemplars make a face categorization decision easy, whereas they make learning a new face in that region or recognising a known face more difficult. Burton and Vokey (1998) have illustrated that care should be taken when considering multidimensional spaces using a two-dimensional representation or terms such as densely packed. At higher dimensions, while it is the case that the centre of the space will be most densely packed this does not mean that most faces will occur at the centre. The actual distribution of faces in terms of their distance from the centre of the face space is considered in more detail below.

[2] The central limit theorem states that if one samples over any type of random distribution then the samples will tend towards a normal distribution as the sample size increases. With regard to facial dimensions, these may be made up of a variety of elements, some that will vary normally (e.g., nose length), and some that will not vary normally (e.g., eye colour). As a facial dimension will combine many of these elements, the distribution of the dimension will be normal.

While these two assumptions of the face space will be kept in the current development, further assumptions by Valentine (1991) will be dropped (or at least developed—described later). Valentine stated that the ease of recognition of a face would be determined by: (1) by the error of encoding; (2) the similarity of the probe to the most similar exemplar; and (3) the similarity of the probe to the second most similar exemplar. While this is similar to the current development, it would be a simplification of it. Nevertheless, Valentine was able to account for many of the observed distinctiveness and race effects using this account. The framework was also used to make predictions regarding race effects that were subsequently confirmed (see Valentine & Endo, 1992).

The exemplar-based face space, just like all other face spaces considered here including the face-space-R, is agnostic to the nature of facial dimensions. It does not matter what aspects of a face we are encoding whether it is specific distances between features or more holistic configurations. Research into the nature of dimensions has continued somewhat in parallel to the theories of face space. While it is useful to remain agnostic, one might think of dimensions as being configural (as described by Tanaka & Farah, 1993) or second-order relations (as described by Rhodes, 1988). The appearance of these dimensions may be similar in form to the eigenfaces generated by principal-component analysis of faces (e.g., Turk & Pentland, 1991). These like holistic images of variable transparency that can be combined to generate a variety of facial image.

THE NORM-BASED FACE SPACE

The face-space framework described above was just one of two that Valentine (1991) described. It was, however, the preferred version as described in later work (e.g., Valentine & Endo, 1992). The alternative face space was one based upon the notion of a prototypical face or face schema. This norm-based face space was a product of the legacy of schema theory (e.g., Goldstein & Chance, 1980; Vernon, 1955). A form of this model was described (as the prototype hypothesis) by Valentine and Bruce (1986) but it was developed and described in more detail by Valentine (1991). This face space is similar to that described above but differs in that the similarity metric is assumed to be based upon a prototypical norm face. A deviation from this norm is encoded—but not as a simple-point vector as in the exemplar-based face space. The exact similarity metric within such a face space has not been defined explicitly but proponents of this version of face space make some statements regarding its properties (e.g., Rhodes, Brennan, & Carey, 1987; Rhodes & Tremewan, 1994; Stevenage, 1995). From these accounts, one can determine that the proposed metric has the direction of the vector from the norm face as being more important than the magnitude of the vector—it may, therefore, employ some version of the vector dot product in the calculation of similarity.[3]

In spite of the lack of a clear similarity metric, the norm-based face space was popular throughout the 1990s. Its popularity was due to the fact that it could account for the caricature advantage, whereas the exemplar-based account was perceived to be unable to account for such effects. The caricature advantage is the finding (by Rhodes et al., 1987 and others) that a face whose features have been slightly exaggerated in a direction away from an average face is recognized better than the original (or veridical) face. The norm-based explanation for this effect was derived from the fact that the caricature's representation would have the same direction (i.e., angle away from the norm face) as the veridical but its magnitude would be larger (Rhodes, 1996).

This version of the face space also made predictions regarding so-called ''lateral caricatures''. These were images transformed like caricatures but in a direction orthogonal to the direction of caricature. Norm-based models predict that these would be harder to recognize than equivalent anticaricatures (transformed towards the norm face) because lateral caricatures alter the direction of the vector but anticaricatures do not. Early reports of experiments suggested that these predictions were confirmed (see Carey, 1992 or Rhodes & Tremewan, 1994). More rigorous experiments, however, revealed that these predictions were contradicted (Lewis & Johnston, 1998; Rhodes, Carey, Byatt, & Proffitt, 1998).

One problem with the norm-based face space is that, in several studies, the caricature advantage is limited—there is a recognition advantage only with small exaggerations. An exaggeration beyond 16% for line-drawn faces results in worse recognition (Rhodes et al., 1987). A similarity metric based on the direction of the vector cannot easily explain this finding. A second problem is that the norm-based face space cannot handle the effect of distinctiveness in other-race faces (Valentine & Endo, 1992) without postulating multiple norms (while this has been considered feasible the implications of multiple norms would mean that the model would behave in a way indiscriminable from the exemplar models). As a consequence, there has been little interest in norm-based face spaces following the publication of exemplar-based versions of the face space, which can account for caricatures, two of which are discussed below.

[3] The dot product of two vectors is a measure of their similarity. If two vectors have the same magnitude then the size of their dot product will increase as the angle between then gets smaller. This, however, does not provide a good similarity metric on its own because the size of the dot product also increases as the magnitude of one of the vectors increases. A similarity metric based purely on the dot product would not obey the rule that similarity is greatest when the two vectors are identical.

THE ABSOLUTE CODING FACE SPACE

The ‘‘absolute coding’’ face space is a version of the exemplar-based face space used by Byatt and Rhodes (1998) to account for caricature effects. One effect of caricaturing is that it moves the vector representation of a face into a region of the face space of lower exemplar density. It is asserted that it is this lower exemplar density that makes the caricatured image easier to recognize than the veridical (Byatt & Rhodes). Also, a trade off between the advantage of being in an area of lower density and the disadvantage of being further from the stored exemplar places a limit on the degree of caricature that will improve recognition.

The absolute coding account is based on Valentine’s (1991) exemplar-based version with the modification that the ease-of-recognition calculation is not based on just the two nearest exemplars but instead on all exemplars within a certain range. The size of this range and how the exemplars act together to make recognition harder, however, remains undefined. The absolute-coding face space is again consistent with the model proposed here but face-space-R explicitly states how *all* exemplars combine to determine ease of recognition.

THE VORONOI FACE-SPACE

Like the absolute coding face space, the Voronoi[4] face space was proposed to account for the caricature advantage. Lewis and Johnston (1998, 1999a) proposed that the face space was tessellated (or irregularly tiled) into identity regions. These are regions where a probe representation would lead to recall of a particular identity. These regions are defined by the Voronoi cells surrounding a veridical, known exemplar’s representation. Voronoi cells are based on a geometric nearest-neighbour algorithm (see Lewis & Johnston, 1999a for details).

A normal distribution of exemplar faces will create Voronoi cells that have their centres at a caricatured representation. The caricatures, therefore, will be better recognized because their representations are more central within the identity region than the veridical representations are. This version of the face space can account for both the limited caricature advantage and why one gets a smaller caricature advantage for photographic faces than line-drawn faces (4.4% photographic caricature for best likeness; Benson & Perrett, 1991). This prediction comes from a more accurate probe-face representation obtained from a photographic than a line-drawn face. The size and shape of the Voronoi cells will have less influence on an accurate representation than an inaccurate representation. The model also explains why the caricature effect is larger for

[4] A Voronoi diagram is a structure that divides a Euclidean space into discrete ‘‘cells’’ based on discrete ‘‘sites’’ in the space. Each cell is defined by the set of locations that have a particular site as its closest site. The resulting diagram, therefore, is a tessellation of the space into convex polygons each surrounding a single site (see Fortune, 1992 for more details).

typical faces than distinctive faces (see Benson & Perrett, 1994). This prediction was derived from a implementation of a five-dimensional Voronoi model. Lee and Perrett (1997) have found, that for brief presentations (as little as 67 ms), the caricature advantage could still be found for 40%-exaggerated caricatures. The Voronoi model may be able to account for this brief-exposure effect, which has not previously been considered—indeed, Lee and Perrett (2000) suggest that their result lends support to the Voronoi model. It will be shown below, however, that the face-space-R model currently being proposed actually predicts the caricature effects observed under brief exposures.

One problem with the Voronoi face space is that the identity regions (i.e., regions of the face space that return a particular identity) fill its whole volume. As a consequence all probe faces produce some identity, which will obviously lead to false positive recognition of any unfamiliar face.

The face-space-R model developed here is similar to the Voronoi face-space account in that it involves comparisons between exemplars. Indeed, if the face space is sparse enough then the current model produces Voronoi cells. The calculation of identity cells, however, is somewhat different in the current model and this produces a face space that does not suffer from the same problems.

THE CATEGORISATION FACE SPACE

While not explicitly designed as a version of face space, Nosofsky's (1986) Generalised Context Model (GCM) has been applied to face-processing tasks in a manner similar to a face-space formulation (e.g., Busey & Tunnicliff, 1999; Nosofsky, 1991; Zaki & Nosofsky, 2001). This categorization model can be turned into a model for recognizing individuals if one considers that every individual face is its own category. Recognition is therefore the categorization of an image to a particular identity. Indeed, Valentine and Ferrara (1991) described how the GCM could be applied to face recognition. The GCM is a model of categorization that is primarily based on a summed-similarity rule. Briefly, the probability of a response r being made is the sum of the similarity of the probe to exemplars of r divided by the sum of the similarity of the probe to all exemplars. The GCM model produces a probability of response for every category and these, by definition, add up to one. Hence the GCM model will always produce an identity to a particular probe item. Like the Voronoi model, the GCM model cannot produce an ''unfamiliar'' response. Another problem with the GCM is that, for highly distinctive items, caricaturing will always lead to improved recognition. Like the norm-based face space, it is unable to account for a limited caricature effect. Similar problems exist with the SimSample model proposed by Busey and Tunnicliff (1999). This model is largely based on the GCM but is applied to the forced-choice recognition of faces.

The task of recognizing a face and producing a unique identity is quite different to the tasks for which the GCM and similar models were designed. In their application to face processing, the models are asked, what is the probability that a probe face comes from the set of stored faces? In face recognition in the real world, mere familiarity is not particularly useful. The main purpose of a face space is to associate a probe face with a unique identity (thereby allowing an appropriate response). This is a different task to the forced-choice tasks employed by Busey and Tunnicliff (1999), and Zaki and Nosofsky (2001). In their respective experiments, where in each case participants learnt a set of new faces, they found that the response to a morph face (from the new faces) was greater than that to the parent faces of that morph. The morph faces were generated by locating anchor points on two faces and producing a new face, which had the shape and texture pattern of a mix of each parent face. This technique of morphing immediately demonstrates how different the task is from face recognition. If one takes a morph of two famous (and easily recognized) people, then one can immediately see that the morph will not be easier to recognize (i.e., identify) than the parent faces—indeed, it will probably be unrecognizable (see Figure 1).

Lee, Byatt, and Rhodes (2000) applied the GCM to explaining caricature effects. They attempted to use the model to account for actual performance on the recognition of caricatures, veridicals, and anticaricatures of 16 very famous faces. While the model is very good at accounting for the recorded data, adopting the implemented model would lead to some worrying generalizations. First, some of the faces would produce increased recognition no matter how far it was caricatured. This is obviously wrong because the face would stop looking even face-like after a point. Second, no matter what face is used as a probe, the sum of the probabilities of recognizing one of the 16 celebrities will be one. Every face, therefore, will produce positive recognition. Obviously, this would not work as a general model for face recognition.

It may be possible to revise the GCM account in order to address issues of recognition rather than forced-choice categorization. Indeed, the introduction of the response-criterion term (k) by Shin and Nosofsky (1992) offers a method for modelling tasks that include the option to say that an item is not from any previous category. It has been shown, however, that such models are very similar to the type of model proposed here (Wills, Reimers, Stewart, Suret, & McLaren, 2000).

A NEW VERSION OF THE FACE SPACE

Here, a new (or at least updated) version of face space is presented and tested. This version is different from the versions just described. It has the advantage over some of the earlier versions in that it is explicit in its formulation and it may be preferable to others because it can account for a wider range of data. This version is an exemplar-based account but one that considers the effects of the proximity (or similarity) of *all* other exemplars. The exemplar-based models

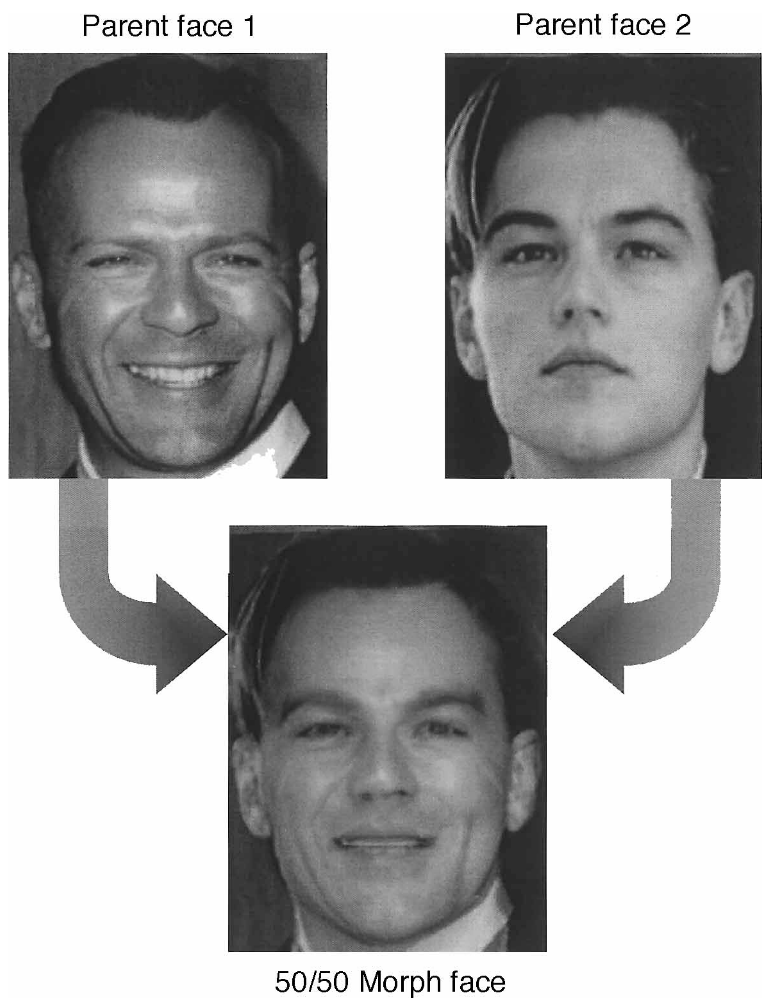


Figure 1. When presented with just the morph face, 18 out of 20 judges (psychology undergraduates) failed to recognize the face as someone familiar. All 20 judges subsequently recognized and named the two parent faces.

described previously used only a limited subset of faces to calculate how well a face is recognized.

The proposed face-space-R model makes use of ideas borrowed from models of discrimination of stimuli that date back to ideas of choice (i.e., response selection) developed by Thurstone (1927). Case V of the law of comparative judgement states that a term corresponding to the probability that A is perceived as different from B is proportional to the *difference* of the activation of A and the activation of B to some stimulus.[5] Further, work conducted by Spence (1936, 1937—based on observations by Pavlov) also suggests that *activations are subtracted* to predict responses. Blough (1975) provides a clearer formulation of such subtractive-discrimination models. These models were intended to account for the size of a response that an animal makes to a stimulus after varying degrees of positive or negative learning. The predicted size of a response can be reduced to the schedule of reinforced trials (which produce an excitation gradient) and nonreinforced trials (which produce an inhibition gradient). Responses also generalize to other, similar, stimuli according to a similarity gradient. Spence originally worked with an hemi-elliptical similarity gradient but work that is more recent employs a Gaussian function (Blough). The actual size of a resultant response will be a sum of the excitatory activation minus the inhibitory activation.

This Thurstonian (or Spencian) idea of choice based on activation dispersal and competition has been recently updated to account for a range of categorization data for humans, animals, and artificial neural networks (Jones, Wills, & McLaren, 1998; Wills & McLaren, 1997, 1998; Wills et al., 2000). These studies have shown that, in a range of categorization tasks, a winner-takes-all algorithm (e.g., those based on Spencian concepts) outperforms the ratio rule (e.g., those employed in the GCM). Although derived from a mutually inhibitory neural network (as in Burton, Bruce, & Johnston's, 1990 model of face recognition), the model can be stated as the final response is proportional to the initial activation of the most activated unit minus the activation of all other alternative units.

The concepts developed for Spencian discrimination learning can be generalized and applied to face recognition. In face recognition terms, the discrimination of a face can be interpreted as its recognition. The response to a stimulus probe face will be recognition of a face as that particular person; so, instead of having a single response function (as in many animal studies—e.g., key pressing), a discrimination face model will have a response function to every recognizable identity. Such discrimination learning has already been applied to face recognition in order to account for the peak-shift effect[6] found using morphed faces (Lewis & Johnston, 1999b). The face-space-R model,

[5] Conversely, the term corresponding to the *difference* of the activation of A and the activation of B to some stimulus is proportional to the probability that A is perceived as different from B. This suggests that the probability of A producing a response is the difference between the activation of A and the activation of B to some stimulus. The important element of Case V is that it is difference between the activations that lead to categorization or discrimination.

[6] Peak shift is the larger response to a version of a stimulus that is changed from the original reinforced stimulus than to the reinforced stimulus when that change makes the stimulus less like some other stimulus that has been negatively reinforced (see Purtle, 1973 or Thomas, Mood, Morrison, & Wiertelak, 1991 for reviews).

however, requires generalization of Spence's framework from a single discrimination to multiple discriminations (i.e., discrimination of every known face). Further, we know that faces vary on many dimensions and so the current model must also generalize Spence's ideas (based on single dimensions such as tone wavelength) to a multidimensional space.

The face-space-R model

The proposed model is a Spencian development of the face-space concept and, as such, makes many of the same assumptions of previous face-space models. The important common assumptions are the Euclidean nature of similarity and the normal distribution of exemplars. We assume that an *N*-dimensional Euclidean face space is populated by *n* known exemplars (or recognizable identities—more on this later) represented by the vectors f_i.

The concept of discrimination gradients (i.e., activation of identity is based on similarity between probe on representation of known face) replaces some original aspects of the face-space framework as described by Valentine (1991). Originally, it was proposed that the ease of recognition be defined by the error of encoding. This can be seen as being equivalent to the degree of spread of the discrimination gradients over the face space. The degree of this error will decrease to a minimum as a face is attended to for longer or under better conditions. In addition, Valentine stated that ease of recognition would be influenced by: The similarity of the probe and the nearest exemplar; and the similarity of the probe and the next nearest neighbour. In the face-space-R model, these terms are replaced by: The activation of the most similar exemplar; and the activation of all other exemplars.

The probe face (x) is assumed to produce, within the face space, excitation of all known identities. The degree of excitation will be dependent on a Gaussian function of similarity of the probe to the exemplar. The activation of any particular exemplar representation (act_i) is defined as:

$$act_i(x) = e^{-\|xf_i\|^2/2\sigma^2} \tag{1}$$

where σ is the degree of spread of the discrimination gradients, which will also be referred to as the error. This equation refers to the situation where the subject has had sufficient time to attend to a clearly present good quality view of a face. If the face is presented briefly, blurred or at an unusual angle the degree of spread would be wider (the σ value would be larger). This equation means that the activation of a particular exemplar is determined by distance between the probe and the exemplar and the function of this distance/activation relationship is a Gaussian curve. The activation of all the identities will be positive to a greater or lesser degree. Competition is required, therefore, to determine which identity, if any, is discriminated and retrieved for that particular probe.

The competition for the retrieval of identity is assumed to operate by subtracting the activation of all others from each exemplar (cf. discrimination learning, as proposed by Spence, 1936, 1937). This competition operates in a similar manner to the operation of the face-recognition units (and other layers) of Burton and colleagues' (1990) Interactive Activation and Competition (IAC) model of face recognition. The response of each known identity ($r_i(x)$) is determined by the equation:

$$r_i(x) = act_i(x) - \Sigma_{\forall j \neq \mathrm{i}}\{act_j(x)\} \tag{2}$$

This equation means that the response of a particular face exemplar is determined by that exemplar's activation minus the activation of all other exemplars. This calculation can be interpreted as an explicit formulation of the absolute coding model. The first part of the expression is dependent on the distance from probe to target, whereas the second part is a measure of exemplar density at the probe. Such a procedure leads to at most one identity producing a response that is greater than zero.[7] If a positive response is found then the probe is identified and the ease of this identification is determined by the size of the response. If, however, no exemplar produces a positive response then the probe is interpreted as an unfamiliar face. Adding many more faces to the face space will, of course, have the effect of reducing the responses to probes, but, as long as the dimensionality is high enough and the error term is lower enough the negative contribution of these faces will remain small enough for accurate probe faces to be correctly responded to and recognised.

With low numbers of dimensions, it is possible to visualize that the face-space-R produces identity regions similar to those of the Voronoi model (regions where a probe results in a common identity). Figure 2 shows the activations, responses, and identity regions for probes in a simple one-dimensional face space. A number of important features can be seen even in this simple demonstration: In particular, distinctiveness effects and the caricature advantage.

How the face-space-R behaves with many exemplars in high-dimensional space is not easy to imagine. It may behave differently to how a low dimensional equivalent system would operate. A one-dimensional version of face-space-R that could recognize a complete set of faces would also report any novel faces as being familiar. Whether this is also true in high dimensional spaces is an

[7] Equation 2 cannot produce more than one positive response for any pattern of activations. This is because if a particular identity has a positive response then its activation must be greater than the sum of all other identities. The next largest identity cannot produce a positive response because it would have to be larger than the sum of all other activations but we already know that it is smaller than the largest activation. In this way, equation 2 guarantees that only one positive response will be produced for any probe face.

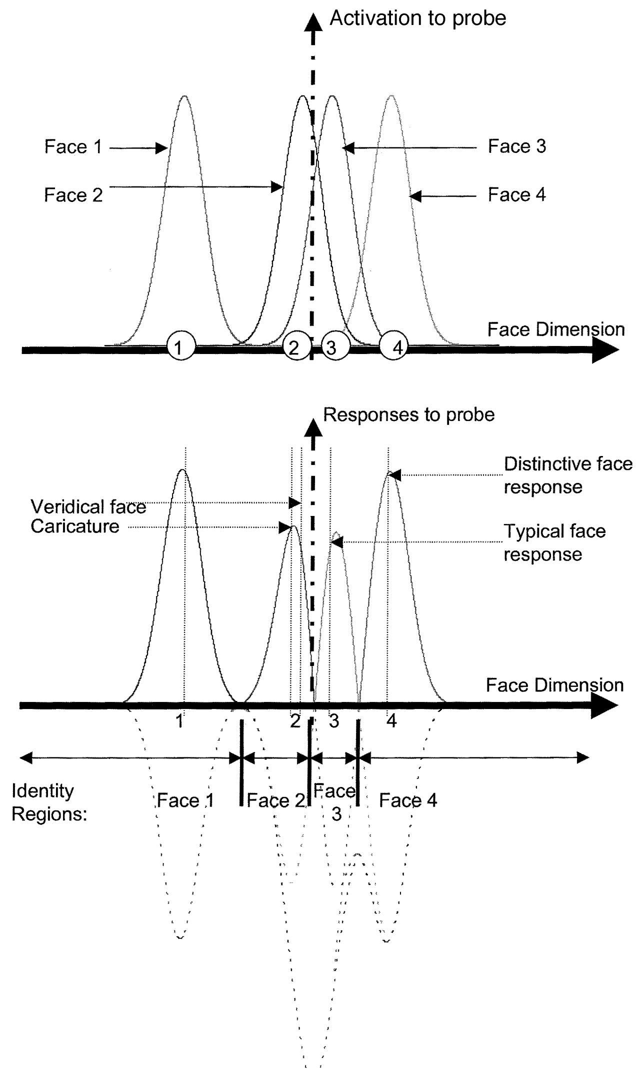


Figure 2. The upper graph shows the activation patterns produced by a probe (as it moves along the face dimension) with Gaussian dispersion for four stored exemplars in a one-dimensional space. The lower graph shows the pattern of responses for the four identities to a probe moving across the dimension (the dotted lines show the negative responses when the faces are not being recognized). Four identity regions are generated where a positive response for a particular identity is produced. The typical faces show lower response when the probe is on them compared to the distinctive faces when the probe is on them. In addition, the caricature produces a larger response than a veridical face.

important question. In order to answer this and many other questions it was necessary to investigate this form of face space computationally using a realistic number of exemplars and a realistic number of dimensions.

Familiarity effects and learning new faces

Unlike previous formulations of face space such as Valentine's (1991) exemplar-based face space, all faces do not necessarily have to have equal representations in term of their strength within the face-space-R model. This point is important if we are to model the fact that people that have been seen more times are recognized faster than less well-known people (e.g., see Lewis, 1999). Also, a varying degree of familiarity is important if the learning of new faces is going to be incorporated into the model. In order to understand the nature of the possible variations in face representations it is necessary to appreciate what each exemplar represents.

Face spaces are populated by vector representations of faces. Exactly what constitutes an exemplar depends on the particular theorist's point of view. It is possible that a new exemplar is placed within the model each time a new face is encountered even if they are just passed on the street. This leads to many exemplars that cannot be recognized because we cannot effectively recall every face we see. This type of theorizing was considered by Valentine (2001) to explain how the Voronoi face space could overcome the problem of false positive recognition described above. An alternative interpretation of the exemplars is that representations only exist for faces that can be explicitly recognised by the individual as a unique person. This situation is more in line with Valentine's (1991) original description. The former interpretation has the problem that making a face-familiarity decision would take place at some stage subsequent to face-space processing and so should not be effected by typicality, whereas Valentine and Bruce (1986) have demonstrated that it is. This second interpretation requires some system by which an unfamiliar face moves from not having an exemplar to having a complete exemplar in one step.

The face-space-R model falls between the two situations described above. All faces encountered are encoded within the model but only some are encoded to a degree where recognition is possible. Also, the move from unfamiliar to recognizable can proceed through a series of steps and so a single presentation of a face may not be sufficient for that person to be recognizable. In order to incorporate learning and familiarity into the model it is necessary to introduce a strength parameter (λ_i) for each face. This parameter ranges from 0 for an exemplar representation never before encountered, to 1 for a highly familiar face. This term acts to modify the calculation of the activation of each possible exemplar according to the following modification of equation 2:

$$r_i(x) = act_i(x)\,.\,\lambda_i - \Sigma_{\forall j \neq \mathrm{i}}\{act_j(x)\,.\,\lambda_j\} \tag{3}$$

If λ_i is small then it is unlikely that $r_i(x)$ will be positive and so recognition will not occur. This situation models performance when an individual is seen briefly and not recognized on a subsequent occasion. As λ_i gradually increases through further contact with that person's face then the $r_i(x)$ value will increase until it reaches a positive number and recognition occurs. This situation represents where a person is a new acquaintance and is recognized. Further exposure to the face will further increase λ_i and so also increase the $r_i(x)$ value. This would represent a highly familiar person who would be recognized more quickly than an acquaintance. The actual λ_i value required for recognition to occur will be dependent on the number and proximity of surrounding exemplars (i.e., the local density). Typical faces will require a larger λ_i value than distinctive faces, so making them harder to learn. It is possible that even a maximum value of λ_i may not be sufficient for some new faces to be recognized and in such situations new dimensions will need to be added to the face space for recognition to occur.

The large number of people who are seen but who do not become acquaintances (i.e., people passed on the street) will have little influence on the responses of well-known exemplars. This is because the λ_i's for these items will be close to zero.

In the simulations that follow, a simplification is made to the model in that only two values of λ_i are considered. All items that are not recognizable are given a value of zero and so can be ignored. All items that are recognizable are given a value of one and so represent equal familiarity between faces.

The computational model

The principles of the face-space-R model were implemented using randomly generated normally distributed data. The model has three free parameters. The first parameter is the number of faces that the model knows and can recognize. There is no evidence for how many faces an average person can recognize and so a rough (and debatable) estimate was used. This was taken to be either 1000 or 5000 and was used throughout the computational work (most work was conducted using 1000 exemplars but where the outcome of the modelling work was affected by the number of exemplars both 1000 and 5000 were used and the results were contrasted).

The second parameter was the number of dimensions. There has been some debate as to how many dimensions we use to identify faces. Shephard and colleagues (1977) employed a method of multidimensional scaling on the free clustering of a set of 100 faces. Analysis revealed that the pattern of clustering could be accounted for using just three dimensions that could be interpreted as hair, face shape, and age. While this technique reveals what are important factors in indicating similarities between faces, it does not identify all the dimensions of recognition. Shepherd and colleagues also employed a factor analysis on the subjective ratings of facial features. The 32 most important features of 100 faces were rated and subjected to a factor analysis, just 10 or 11 factors were able to account for 60% of the total variance. Both these methods offer solutions to the question of how many dimensions we use to encode faces. More recent work by Burton, Bruce, and Hancock (1999) have employed up to seventy (albeit dichotomous) dimensions within their complete model of face recognition. Estimates of the dimensionality of face space, therefore, lie between 3 and 70. The parameter of the dimensionality in face-space-R model was varied throughout in order to generate an estimate for its ideal value (and possibly real value if the model is to be believed).

The third parameter was the spread of the discrimination gradient or error (σ). This is a measure in distance units, which are arbitrary but are related to the spread of the exemplars. The error, therefore, is expressed as a ratio of the Gaussian used to generate the exemplars. An error of 1 would imply that the spread of the excitation is the same as the spread of the exemplars. This third parameter was also allowed to vary.

Finding the dimensionality of face space

One consequence of this model is that it is possible to generate an estimate of the number of dimensions typically employed in a human's face space. All that is required is to know the value of some property of face processing that is dependent on dimensionality. One such property comes from the research into caricature effects.

Caricatures of faces can be systematically generated by exaggerating the features of a face away from a calculated average. Rhodes and colleagues (1987) found that such caricaturing produces images that are easier to recognize than the original. By allowing participants to vary the degree of caricature, Benson and Perrett (1991) identified the degree of exaggeration that produces best likeness of that person. As mentioned earlier, on average, this best likeness occurred at an exaggeration of 4.4% away from the prototypical face for photographic quality faces. For line-drawn faces this value increased to 16% exaggeration (Rhodes et al., 1987).

We can use the value of 4.4% exaggeration for best likeness to constrain the parameters of the model. If the model operates in a similar manner to human face spaces then an equivalent caricature advantage should be found. The size of the caricature advantage in the model, however, will be influenced by the number of dimensions employed and the error. By varying the two free parameters (i.e., error and number of dimensions), it is possible to generate a locus of values (i.e., value combinations for the two parameters) that lead to a caricature advantage of approximately 4.4%.

At this point, we have one expression (the relationship between the parameters and the caricature effect) and two unknowns (dimensionality and σ—error). It is not possible, therefore, to obtain a single estimate of the dimensionality. With two further assumptions, however, it is possible. These assumptions refer to the number of dimensions required to recognise a face in optimal conditions such as from a front-view photograph (as employed by Benson & Perrett, 1991). The first assumption will be called the *necessity assumption*. It claims that the dimensionality of the face space will not be less than that required to recognize all known exemplars. That is, all probe faces placed at exact representations of stored exemplars must produce positive identification. If this is not the case then some of the stored faces will not be recognized by the model. It can be considered that if the human system were unable to recognize some faces then further dimensions would be added to the face space in the manner of the uniform-developmental model proposed by Johnston and Ellis (1995). This assumption would seem to be an imperative for the face space to work. The second assumption will be called the *sufficiency assumption*. This states that the number of dimensions of the face space is not more than that required to recognize all known exemplars. That is, none of the dimensions are redundant and could be removed without violating the necessity assumption. This second assumption makes sense from a developmental point of view but is not an imperative in the way the necessity assumption is.

The model can be employed in order to find a set of parameters that satisfy both the empirical finding from the caricature experiment and the two assumptions described above. Finding these ideal parameters involved calculations of the caricature of best likeness under a range of conditions.

The implementation

A set of 1000 (or 5000 for the second implementation) normally distributed (along n dimensions) exemplars was generated. For each of these exemplars, a caricature for best likeness was found. This was found by considering probe faces along a line connecting the average face to the target face and beyond (see Figure 3). For these probe faces, the response of the target identity was found. The probe face with the largest response was deemed the best likeness. The percentage caricature for best likeness was calculated as

$$\text{Percentage caricature for best likeness} = \left\{ \frac{\| Of_{maximum} \|}{\| Of_{verifical} \|} - 1 \right\} \times 100 \qquad (4)$$

where: O is the vector representation of the average face; $f_{veridical}$ is the representation of the exact exemplar; and $f_{maximum}$ is the representation of the probe face (along the extended line connecting the average and the veridical) that gives the highest response. This equation is equivalent to the percentage caricature term used when caricatures are generated for experiments. Two other values were also recorded for each of the faces. First was the exemplar *distinctiveness*, which was the distance between the exemplar and the average. Second was *the*

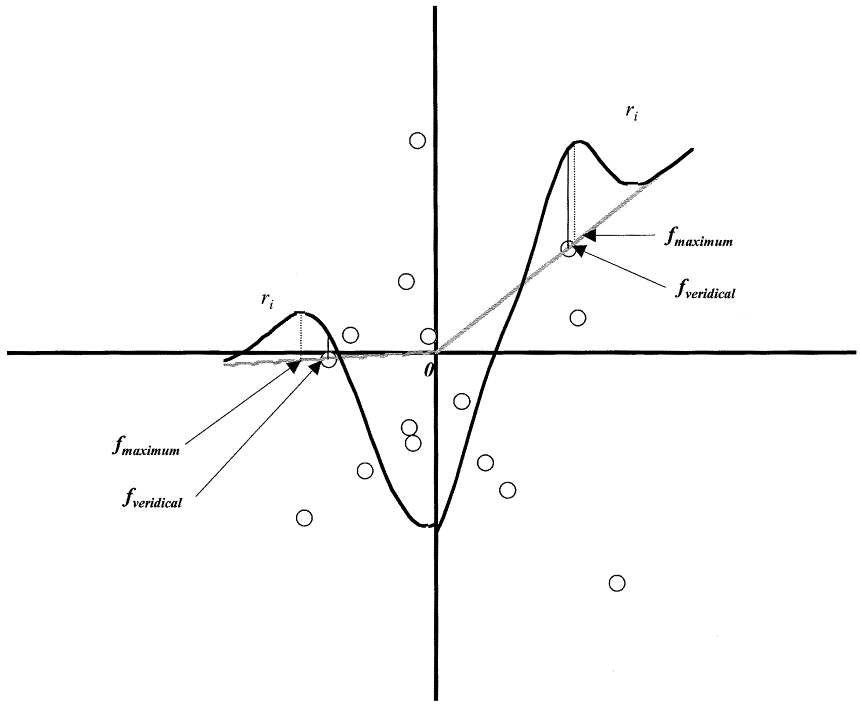


Figure 3. A two-dimensional representation of the type of implementation used to find the degree of caricature for best likeness. The circles represent 15 exemplars distributed along the two dimensions. The tests for just two of these exemplars are shown. The continua along which recognition is tested are shown as the two grey lines. The responses (r_i's) given by the identity along these continua are shown by the curved black lines: The higher the black curve is above the grey line the stronger the response. For each of the two faces it can be seen that the response of the veridical (i.e., r_i at $f_{veridical}$) is not the maximum response. The maximum response for these two items is at a positive displacement away from the centre on the space. For these two items, therefore, best likeness occurs for a positive caricatured form of the item.

veridical response, which was the target identity's response to the probe when it was exactly on the target representation.

The implementation took place using a range of errors (σ). For each of these, the number of dimensions were gradually reduced until some of the exemplars were not recognized (i.e., the necessity assumption was violated). The penultimate dimensionality for each series, therefore, provided the situation where the sufficiency assumption was not violated.

The average caricature for best likeness was recorded for all sets of parameters considered. The results are presented for the 1000 exemplar implementation in Figure 4a. From this figure it can be seen that the parameters that are consistent with the assumptions and give a caricature of best likeness at 4.4%

are (for number of faces = 1000): Error or $\sigma = 0.92$; and dimensionality = 15. The results are presented for the 5000 exemplar implementation in Figure 4b. This figure shows that the parameters that are consistent with the assumptions and give a caricature of best likeness at 4.4% are (for number of faces = 5000): error or $\sigma = 1.10$; and dimensionality = 22. Due to the large number of random data points, these calculations were extremely time consuming. To generate the data shown in Figure 4b required 20 machines (1 GHz each) running almost continuously for 3 weeks.

The conclusions from this modelling work are that, assuming that one knows somewhere between 1000 and 5000 people, the dimensionality of the face space is between 15 and 22. The order of magnitude of the dimensionality, therefore, appears relatively robust over a wide variety of parameter values. Indeed, changing the 4.4% caricature value, which after all is an experimental value and hence an estimate, has little effect on the estimate of dimensions. Doubling the caricature advantage increases the dimensionality by about two dimensions where halving decreases the dimensionality by about five dimensions in the 1000 faces case.

It may be suggested that the 4.4% caricature advantage on which these dimensionality estimates are based is actuality just an extrapolation. The figure comes from best likeness estimates using images caricatures in steps of 16% from −32% to +32%. In these judgements the image chosen most often was the veridical face. The implementation described above allowed any degree of caricature to be reported as the best likeness between −200% to +600% with an accuracy of 0.01%. Two typical patterns of responses over the range are shown in Figure 3 albeit for a two-dimensional face space. These reponses patterns show the similar asymmetrical pattern as shown in Benson and Perrett's (1991) choice data with increased choice at a caricature of the veridical face. Further, the model can be analysed according to the identity response at just the five levels of caricature used by Benson and Perrett. For each exemplar the responses to the five levels of caricature was recorded. Using the ideal parameters found earlier, the items that produce the highest responses most often were the veridical. Adding an amount of random binomial noise to the data the pattern of highest responders is very similar between the model (Figure 5a) and Benson and Perrett's data (Figure 5b). This demonstrates that if the face-space-R model is tested under the same constrained manner as that in which the participants were tested then similar patterns of data (in distribution as well as averages) can be found.

Rhodes et al. (1987) found a caricature advantage with line-drawn faces with best likeness occurring at 16% exaggeration. This value is a lot larger than the 4.4% employed here. The face-space-R model, however, can account for this discrepancy and even tell us something about line-drawn faces because of it. Line-drawn faces do not contain as much information as photographic faces and so the number of dimensions that are available will be less. Exactly how much less can be calculated by reducing the number of dimensions in the model while keeping the error fixed at the level appropriate for photographic faces. Of course, reducing the number of dimensions in this way will violate the necessity assumption but this is consistent with the fact that not all line-drawn faces would be recognized. Figure 4a demonstrates what happens for the situation with 1000 faces and it shows that a caricature advantage of approximately 16% is found when the number of dimensions is reduced to 12 (16.2% caricature for best likeness). The face-space-R model can, therefore, explain the increased car-

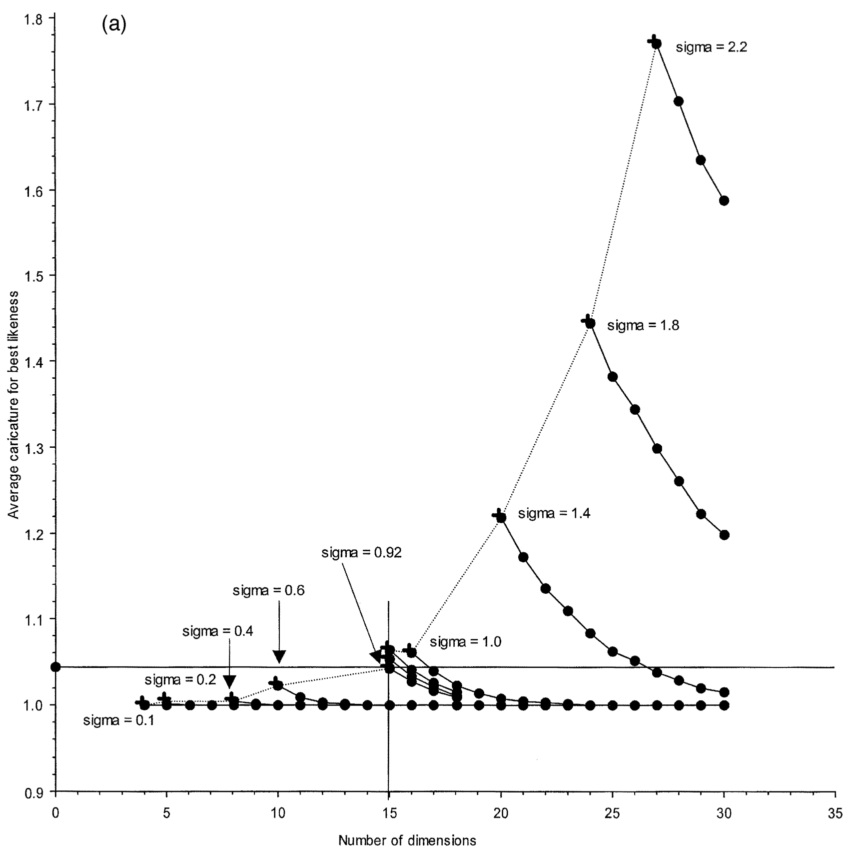
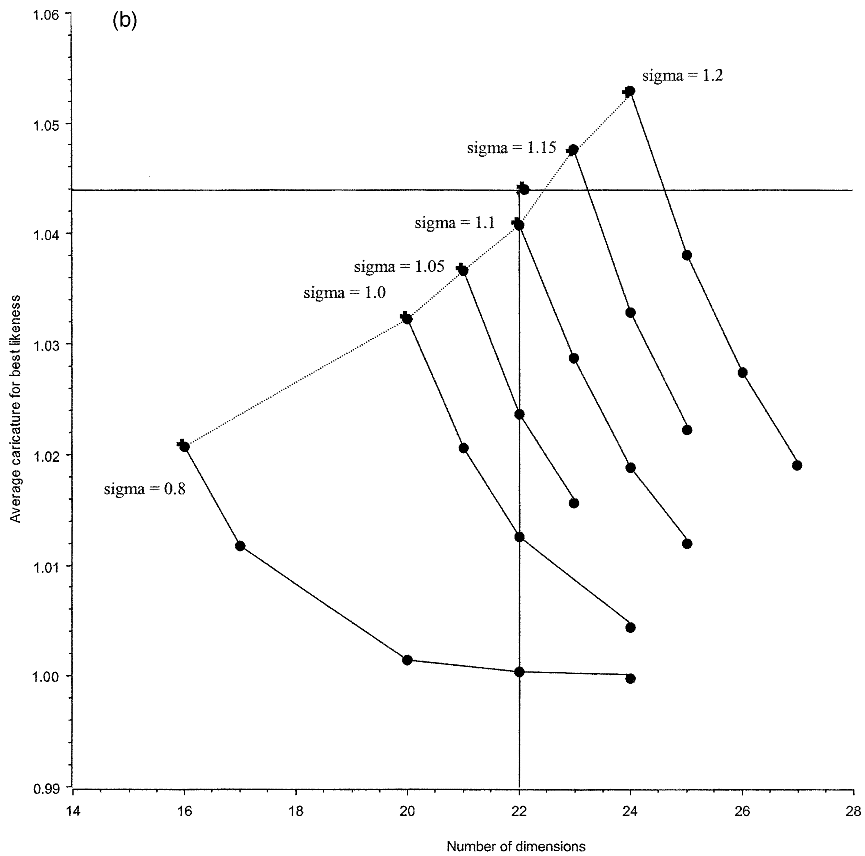
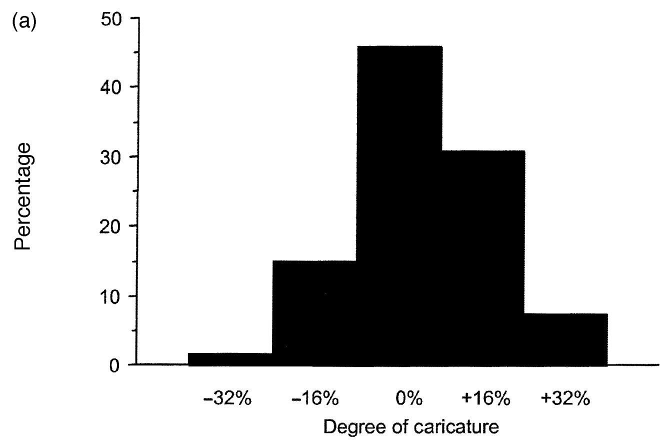


Figure 4 (above and opposite). Average caricatures required for best likeness as predicted by different values of error (σ, different lines) and different numbers of dimensions (on x axis). Only implementation where all items were recognized are included (i.e., necessity assumption holds). Conditions where the sufficiency assumption also holds are linked by the dotted line. The horizontal line indicates the empirical situation where the caricature for best likeness is 4.4%. Where the dotted line crosses the horizontal line provides an estimate for the dimensionality of a typical face space. In the case with 1000 exemplars (a), the dimensionality is 15. In the case with 5000 exemplars (b), the dimensionality is 22.



icature advantage for line-drawn faces and suggests that we use approximately 12 of the facial dimensions that can be extracted from line-drawn images.

A further concern over the estimate of a 4.4% caricature advantage is that more recent research has found that this advantage can vary up to 40% if the images are presented briefly (see Lee & Perrett, 1997, 2000). While it is acknowledged that this value of 4.4% is an empirically derived from averaged data, it is taken here to be an indication of the order of magnitude of the real value. If, in fact, the real value can vary up to 40% then this would be highly damaging for this type of estimate. Figure 4 shows that a ten-fold increase in the best likeness estimate would produce almost 100% increase for the estimated number of dimensions to face space. The face-space-R model, however, can accommodate Lee and Perrett's data. In fact, the face-space-R model (like the Voronoi model) would predict a larger caricature advantage for briefly presented faces. The following section describes why.



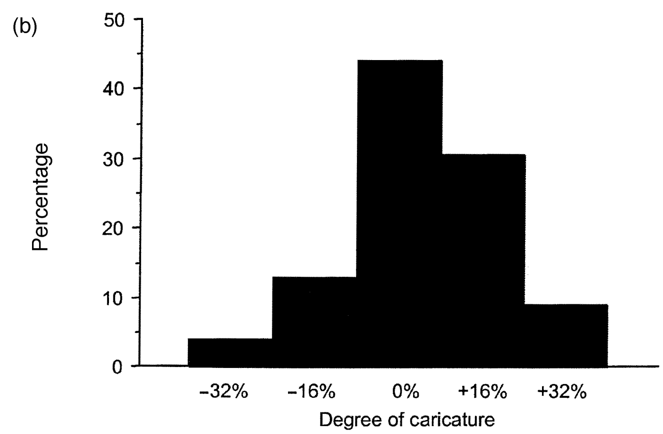


Figure 5. The distribution of items chosen as the best likeness for an individual from a set of five items varying in degree of caricature. The first figure (a) shows the distribution predicted by the face-space-R model whereas the second (b) shows the actual empirical data found by Benson & Perrett (1991—an approximate reproduction).

Dynamics during short presentations

The estimate of 15 to 22 dimensions comes from the assumption that the face is attended to for sufficiently long enough to extract the maximum amount of information available. This is indeed the nature of the task that Benson and Perrett (1991) used to obtain the 4.4% best likeness estimate employed here. In these situations, the σ value (error associated with encoding) in equation 1 is the minimum value possible and part of its calculation required that all known faces were recognizable using this σ value. If a face is presented very briefly, however, then this will introduce further error in the probe representation. The σ value used in the calculation will not be the minimum but might be considerably larger. A consequence of this is that such probes would be harder to recognize (as is indeed the case for briefly presented faces). A second consequence is that the advantage for caricaturing is greater. In order to illustrate this, the face-space-R model was implemented with 1000 faces and 15 dimensions. A variety of values of σ were employed (from 1.0 to 1.6 in steps of 0.1) to model the effect of reducing the presentation time. The degree of caricature for best likeness increased as the value of σ increased, but also the number of faces that were recognized (i.e., produced responses greater to zero) decreased. This pattern of data is shown in Table 1. If we interpret brief presentation as a situation where there is increased error, the face-space-R model predicts that brief presentation will lead to fewer faces being recognized and a larger degrees of caricature for best likeness estimate. The model is, therefore, consistent with people's performance.

Further assessment of the model can be made by examining the effects of short duration of presentations and caricaturing on recognition accuracy. Lee and Perrett (1997) measured recognition accuracy for a set of famous faces. The faces were either veridical or 40% caricatures and they were presented for 33, 67, 100, or 133 ms. Recognition performance increased as presentation duration increased—it was at chance for 33 ms. A recognition caricature advantage was found for 67 ms and 100 ms presentations. For 133 ms presentations, there was no significant difference between the veridical and the 40% caricature. The face-space-R model was tested to explore whether varying the error value alone could produce the pattern of results reported by Lee and Perrett. The model was implemented with 1000 exemplars and 15 dimensions. The error (σ) was given a range of values between 1.0 and 2.1 in steps of 0.1. For each value of the error,

TABLE 1
Performance of the face-space-R model with 1000 exemplars and 15 dimensions for different values of error

Error value (σ)	*Caricature percentage for best likeness estimate*	*Proportion of exemplars recognized*
1.0	8.9%	0.950
1.1	17.9%	0.746
1.2	29.4%	0.470
1.3	42.8%	0.227
1.4	58.7%	0.089
1.5	76.7%	0.026
1.6	96.8%	0.005

Increasing error increases the caricature advantage but the number of exemplars that are recognized decreases.

the number of the 1000 exemplars that gave a positive response to the veridical probe was found. Further, for each value of error, the number of exemplars that gave positive responses to the 40% caricature of the exemplar was calculated. These values (in terms of proportion of items recognized at different errors) are present in Figure 6. When the error is large (analogous to very brief presentation) recognition is at chance for both veridical and caricatures. At intermediate levels of error (analogous to 67 ms or 100 ms presentations) recognition is above chance and is better for the 40% caricature over the veridical. When the error is at the lowest (analogous to 133 ms presentation or longer) there is little difference between recognition performance for the veridical and the 40% caricature. The performance of the model, therefore, is equivalent to the findings by Lee and Perrett if we interpret shortening the duration of presentation as increasing the degree of error associated with the probe.

Fifteen to twenty-two dimensions of face space

The implementation provides estimated values to the dimensionality of face space and this is between 15 and 22 depending on the number of faces assumed to be encoded. These values are higher than some estimates (e.g., the MDS study by Shephard et al., 1977) but lower than have been employed in some models of

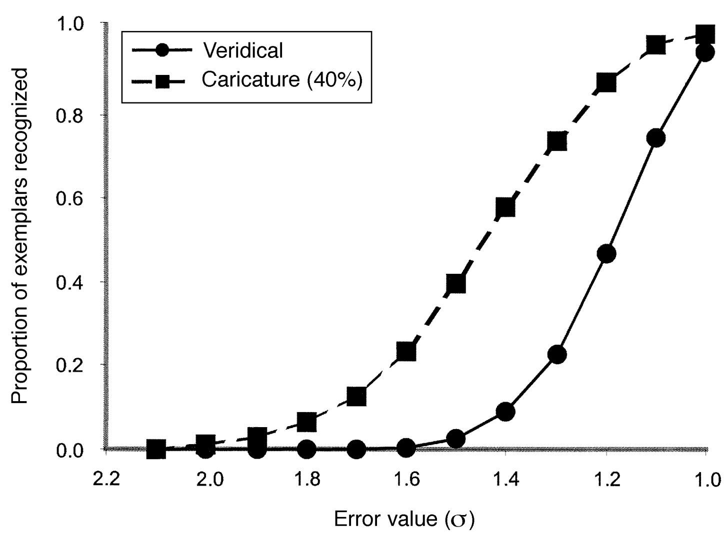


Figure 6. The proportion of exemplars recognised in the face-space-R model with 1000 exemplars and 15 dimensions. Probes were either veridical exemplars or 40% caricatures. The error value was varied to model limited duration presentations with shorter presentations relating to higher values of error.

face recognition (e.g., Burton et al., 1999). The values are of the same order of magnitude as the estimate based on the factor analysis of Shepherd and colleagues (i.e., 11).

The number of dimensions (15–22) is similar in order to the number of eigenfaces that need to be employed to reproduce a recognizable face (Turk & Pentland, 1991). This adds support to a number in this region. Indeed, 15 dimensions appear to be more than sufficient for the task of recognizing faces. Even if faces were encoded dichotomously on each dimension (i.e., as bits), a 15-dimensional face space would still be able to uniquely encode 2^{15} faces, which works out to be 32,768 (2^{22} gives a value of 4,194,304). A greater acuity along the dimensions would mean that even more faces could be uniquely encoded.

It is worth mentioning what this value means and the best way of doing this is to describe what it does not mean. It does not mean that there are just 15 (or 22) measurements that we extract from a face. Rather, it is the number of measurements that are encoded or stored to be compared with future presentations of faces. Further, this number of dimensions is based on the set of homogeneous faces (that is standard faces from a single race). Whether or not further dimensions are employed to encode faces of other races is a debatable point but it is probably dependent on one's familiarity with other-race faces (cf. the contact hypothesis—see Valentine, Chiroro, & Dixon, 1995). These estimates of dimensionality mean that if we understand the brain to be performing a principal-components analysis on faces then it makes use of just the first 15–22 components. This finding may be of importance in understanding how to best employ eigenfaces in automatic face recognition systems or in face reconstruction systems that explore the space defined by eigenfaces (e.g., Evo-Fit, developed by Hancock, 2000).

Distinctiveness effects

As well as providing a measure for the dimensionality of face space, the current model also allows other effects to be accounted for. One of the most important effects in the investigation of face recognition has been that of distinctiveness (Light et al., 1979; Valentine & Bruce, 1986). It is a robust finding that faces that are rated as being more distinctive are easier to recognize. Distinctiveness is believed to refer to (or at least correlate strongly with) the distance between the face's representation and the average face (e.g., see Burton et al., 1994). Distinctiveness, therefore, can be envisaged in the model as the distance between the exemplar and the average. Ease of recognition is interpreted as the size of the response of the target exemplar. This response was recorded as the veridical response in the implementation. Figure 7a shows the relationship between veridical response and distinctiveness for the 1000 exemplars using the

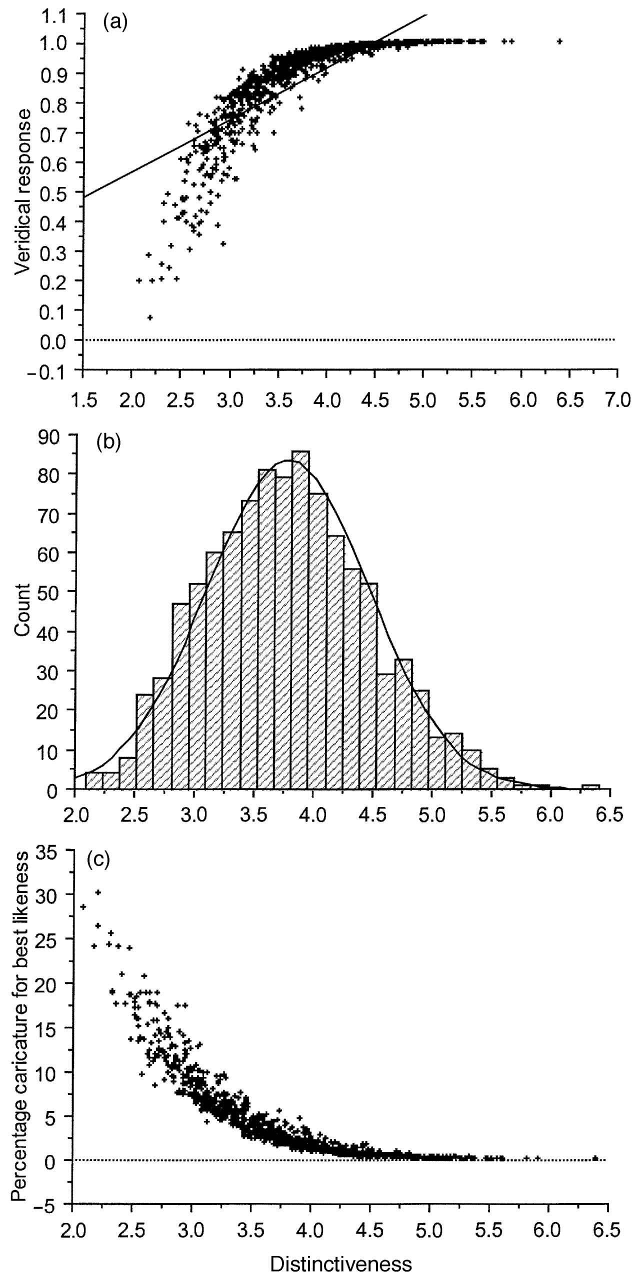


Figure 7. (a) The relationship between the veridical response of each exemplar and its distinctiveness using ideal parameters. This shows a positive, albeit curvilinear, relationship. (b) The distribution of distinctiveness of the 1000 exemplars using 15 dimensions. The curved line indicates the normal comparison. (c) The negative relationship between the size of the caricature advantage and the distinctiveness of each exemplar.

ideal parameters. This shows a strong positive correlation that is consistent with the findings of Valentine and Bruce and others.

Another important issue regarding distinctive is the typicality paradox as described and resolved by Burton and Vokey (1998). This paradox regards the fact that typical faces must be densely packed and distinctive faces are sparse. If this arrangement is considered in two dimensions then one would expect to see many more typical faces than distinctive faces and distinctiveness ratings should be skewed. Burton and Vokey demonstrated that this skew is produced only when two dimensions are considered: The larger the number of dimensions the smaller this skew is. If, however, the number of dimensions is too large then one gets very little variability in distinctiveness. The number of dimensions in face space is important, therefore, for the distribution of distinctiveness. The distribution of distinctiveness for the 15-dimensional face space is shown in Figure 7b, indicating a small degree of skew (0.110). A measure of skew found for distinctiveness rating on real faces is higher (0.768—based on a further analysis of data reported by Lewis & Johnston, 1997). This higher value of skew would be indicative of fewer dimensions but it should be remembered that Lewis and Johnston employed three-quarters views of faces. It is possible that the larger skew in the real ratings data is due to fewer dimensions being used when processing three-quarters view faces rather than full-frontal faces. Another account for this difference are that participants do not make use of the full range of dimensions employed for recognition when making a subjective distinctiveness rating. A further possibility is that the participants do not use the distinctiveness rating scale as a strictly linear indicator of perceived dissimilarity from the average. This last explanation is the most likely because, if we wished to match the skew of the model data to that of the ratings data (i.e., 0.768) then this would require a face space of between just two and three dimensions. Further, even if there were only two or three dimensions then, although the skew of the data would match, the kurtoses of the distributions from the model and the real data would not match.

A further and less obvious effect of distinctiveness has also been documented, this being the interaction between distinctiveness effects and caricature effects. Benson and Perrett (1994) demonstrated that, using line-drawn faces, the size of the caricature effect is larger for typical faces than it is for distinctive faces. It was a relatively simple matter to test whether the proposed model is consistent with this finding. Figure 7c shows the relationship between the distinctiveness of the 1000 exemplars and the caricature for best likeness for each. There is a clear and significant negative correlation between these variables. The relationship, in fact, follows a curvilinear pattern similar to a negative power function. This prediction of the model is consistent with the data found by Benson and Perrett, although no consideration of nonlinearity has been made.

Other-race faces

The race of a face has long been known to affect how well it is recognized by particular observers (see Bothwell, Brigham, & Malpass, 1989 or Valentine et al., 1995 for reviews). One of the most common findings is that faces of races other than that of the participants are harder to recognize than faces from their own race. This own-race bias is believed to be an effect developed from a higher degree of exposure or contact with one's own race (see Valentine et al.). Face-space models have been shown to be able to account for this own-race bias. Valentine (1991) illustrated that if other-race faces were all clustered at some distance from the own-race average face (but with a similar distribution) then distinguishing between the vectors from the norm to the other-race faces would be difficult (i.e., using a norm-based face-space model). Valentine and Endo (1992), however, illustrated that an exemplar-based face-space model better accounted for the own-race bias and, in particular, distinctiveness effects found with other-race faces. How other race faces are represented within a face space is, of course, important. In the following discussion we will describe and consider the implications of the other-race representations suggested by Valentine and Endo.

Faces from other races will be normally distributed but this distribution will be about a point at some distance from the own-race average. One is likely to know fewer other-race faces than same-race faces and so there will be fewer representations. Also, the dimensions that were learnt to distinguish between own-race faces will be of limited value when distinguishing between other-race faces (e.g., eye colour greatly varies for Caucasians but has a low variance among Japanese). In order to represent the fact that own-race dimensions are inappropriate for other-race faces, the spread of the items around the other-race average will be smaller. Figure 8 illustrates the type of distribution of own-race and other-race faces considered appropriate to model.

The face-space-R model was implemented to confirm that it shows an own-race bias. The dimensionality was set to 15 and σ to 1.0. The exemplars consisted of 900 own-race items and 100 other-race items. The own-race faces were generated as previously described with a normal distribution with mean of zero and standard deviation of 1.0. The other-race faces were generated from a normal distribution with a mean of 2.0 and a standard deviation of 0.75. Recognition of own and other race faces was tested using a veridical probe for each of the items.

The average recognition response to the own-race items was 0.934 (*SD* = 0.088), whereas for other-race items it was 0.899 (*SD* = 0.077). This own-race bias was statistically significant, $t(998) = 3.870$; $p < .001$. The size of this difference, however, is completely dependent on the parameters chosen—the effect appears because within the face space the exemplar density for the other race faces is greater than for the own-race faces. There are fewer other-race

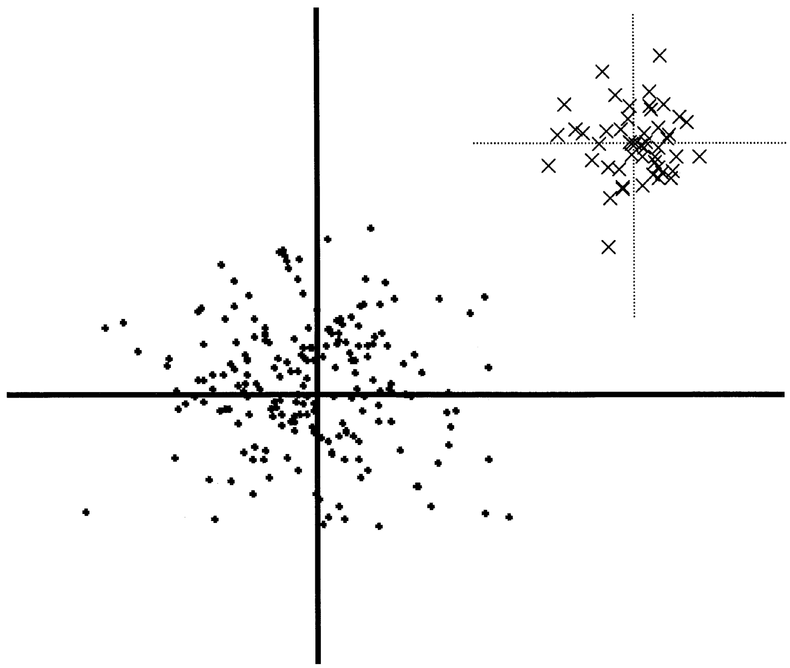


Figure 8. The theoretical distribution of own-race faces (+'s) and other-race faces (× 's) in a face space. Other race faces are clustered about their own average and have a narrower distribution on the face-space dimensions than own-race faces.

faces and so the density might be less but this is offset against a narrower distribution for those faces along the own-race appropriate dimensions.

The race of faces has also been shown to be important for caricature effects. Byatt and Rhodes (1998) explored how the caricature advantage was influenced by whether faces were from one's own race or from another race. Further, they investigated the effect of creating the caricature from a same-race average or a different-race average. The reasoning behind this manipulation was that a norm-based face space would predict a caricature advantage for other-race faces exaggerated from one's own-race average face. Alternatively, their absolute-coding face-space model predicts that the caricature advantage would be greater for other-race faces when they have been exaggerated away from the average of a set of faces from the same other-race category. Their data supported the latter model, in that a caricature advantage was found for other-race faces but only when the average face involved in making the caricature was of the same other-race category. If the average employed was of the same category as the participant then there was no clear caricature advantage. For own-race faces, a caricature advantage was observed for when an own-race average was employed but only an asymmetry (i.e., better performance for caricatures than equivalent anticaricatures) was observed when the other-race average was employed to generate the caricature. In summary, then, the race of the face and the race of the average face have to be the same to obtain a clear caricature advantage.

The face-space-R model was invoked to investigate whether it too could explain the pattern of results found by Byatt and Rhodes (1998). The same set of 900 own-race items and 100 other-race items was used as described above. The parameters of the model were the same. For each item two measurements were obtained. First, the degree of caricature from the own-race average required to produce the maximum response. This value was calculated by measuring the response to probes along and beyond a line connecting the average own-race item and the item its self. Second, the degree of caricature from the other-race average required to produce the maximum response. This value was calculated by measuring the response to probes along and beyond a line connecting the average own-race item and the item its self.

The means of the degree of caricature for best likeness for own-race and other-race items by own-race and other-race averages were calculated. When the caricatures were generated from the same race as the item itself the caricature advantage was 2.4% and 4.8% for own- and other-race items respectively. When a different race was used to produce the caricature there was an almost negligible caricature advantage (0.5% and 0.4% for own- and other-race items respectively). The differences between all four means where significant. The data also illustrate that the caricature advantage using the appropriate average face is larger for other-race items than for own-race items. Although this was surprising, it is consistent with the pattern of data reported by Byatt and Rhodes (1998), who found that the caricature advantage continues for higher degrees of caricature for other- than for own-race faces. The face-space-R model, therefore, is completely consistent with Byatt and Rhodes' data and can account for aspects of their data not previously considered. One further prediction can be generated from this analysis. That is, the face-space-R model predicts that the size of the caricature advantage should decrease as the distinctiveness of other-race faces increase (i.e., the same relationship as found for own-race faces). This prediction is yet to be tested.

Face-categorization tasks

Face-space models have been shown to be able to account for the typicality advantage observed in face-categorization tasks (e.g., Valentine & Bruce, 1986). In these tasks images are presented that could either be whole, intact faces, or be jumbled versions of the same faces. The participant's task is to respond as to whether each image is a whole face or not. It has been found that participants are consistently faster at this task for typical-rated faces than for distinctive-rated faces.

All of the original face-space models could account for this empirical finding. It was necessary, therefore, to demonstrate how the face-space-R model could also account for the data. In a face categorization decision, it is necessary to use one's knowledge of the faces one knows to evaluate the likelihood that an image is a face. All known faces or exemplars will contribute activation to a face-categorization decision. The activation elicited by a probe image will again produce activation according to the distance between its representation and each exemplar representation according to equation 1. A face-categorization decision is different to a recognition decision because there is not the same within-category competition. The activations in this case are all summed to produce a response measure that will influence the reaction time to the task. The equation governing this categorization response measure is:

$$r_{categorization}(x) = \Sigma_{\forall i}\{act_i(x)\} \tag{5}$$

That is, it is simply the sum of the activation of all the exemplars. This calculation will produce a value greater for zero for all probes and so some threshold would be required but this threshold would be different dependent on the nature of the nonface probes. The nonface probes are typically jumbled faces but similar typicality effects have been found when the decision was between human and cat faces (Lewis, 1997) albeit the overall reaction times were different indicating a different threshold was employed for this decision.

The face-space-R model was tested for a simulated face-categorization task. The model contained 1000 exemplars. The dimensionality was 15 and σ was 0.92. The value for $r_{categorization}$ was found for the 1000 exemplars and for a further 1000 ''unfamiliar'' probe items (these were generated from the same distribution as the exemplars but they were unfamiliar to the model). These values are plotted against distinctiveness (calculated as distance from the average exemplar) in Figure 9.

Figure 9 shows that there is a clear negative curvilinear relationship between distinctiveness and categorization response (the sum of all activations). The points for exemplars fall approximately 1 unit above those for unfamiliar items. This merely reflects the activation of 1 provided to the summation in equation 5 by including the exact known exemplar. Clearly, the face-space-R model can account for the typicality advantage seen in the face-categorization tasks.

The discussion of the model implementation of face categorization has been in terms of response size. The data being modelled, however, are reaction times. It can be assumed that there is some monotonic analogue between measured response of the model and observed reaction times. The face-space-R model, in fact, allows this analogue to be made explicit using the reduction of error over time. Let us assume, as before, that the longer a face is observed, the more accurate its encoding is and so the error (σ) becomes lower. Further, the decision that an image is a face involves the response ($r_{categorization}$) becoming positive

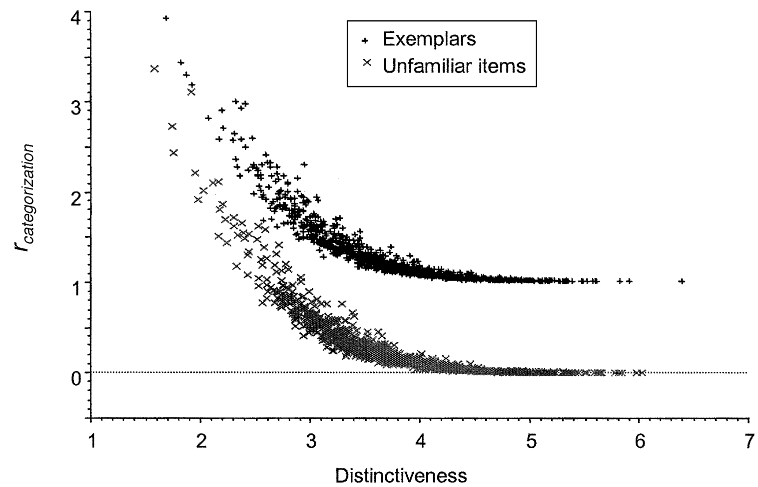


Figure 9. Results of the simulation of a categorization task. The upper set of data (+'s) show responses to known exemplars varying in distinctiveness. The lower set of data (×'s) show responses to unknown probes varying in distinctiveness. Each set shows an advantage for the more typical items.

once the activation of nonface possibilities have been subtracted as in the formula:

$$r_{categorization}(x) = \Sigma_{\forall i = faces}\{act_i(x)\} - \Sigma_{\forall i = nonfaces}\{act_i(x)\} \qquad (6)$$

Defining the nature of the nonface items remains problematic. The set of items that need to be considered, however, can be relatively small because equation 2 shows that the contribution of very nonface-like items will be minimal. Figure 9 shows that the first part of equation 6 will be more positive for typical than for distinctive faces. The larger the error, the larger the first part and also the second part will be. As there are more things that are not faces than are faces, for very large errors the response to a face will be negative. As the face is attended to for longer, both the faces part and non-faces part of equation 6 will decrease but the nonfaces part will decrease faster until the response become positive. The level of error required for the positive response will be larger for typical faces than for distinctive faces and so a response can be made to typical faces earlier than distinctive faces. The face-space-R model, therefore, can explain the typical face advantage for categorization in terms of reaction times.

The reaction times required to categorize a face as a face is considerably less than that required to recognize a face. Valentine and Bruce (1986) found an average reaction time of 684 ms for a familiarity task, whereas a face-categorisation task took only 529 ms. This is consistent with the predictions of the face-space-R model because the error required for categorization response to become positive is considerably greater than the error that will be required for the recognition response to become positive.[8] The same model, therefore, with the same parameters is able to account for the recognition data, the categorization data, and the differences between the two.

Unfamiliar faces

One criticism of the Voronoi model (Lewis & Johnston, 1999a) is that it has no obvious way to deal with unfamiliar faces. All probe faces will be recognized as someone familiar but at different speeds. One possible solution that has been proposed (see Valentine, 2001) is the introduction of a threshold of activation. Such an arbitrary threshold does not sit well with the explicit nature of the Voronoi model.

The face-space-R model is similar in many ways to the Voronoi model; thus, it too may have difficulties handling unfamiliar faces. The introduction of an arbitrary threshold value for recognition could have problematic ramifications for certain aspects of the model such as the sufficiency assumption. This assumption was based on the axiom that the recognition threshold was zero or close to zero.

In order to investigate the nature of the responses that unfamiliar probe faces produce in the current model, further implementations of the ideal model were conducted. The model was set up with the same 1000 exemplars in the 15-dimensional face space. A set of novel probe faces was presented to the model taken from the same Gaussian distribution as the original exemplars. No precautions were taken to ensure that the probe faces were different from the original faces. For each probe face, the response of the highest responding identity was recorded. The distinctiveness of the probe face was also calculated. This procedure was repeated for 5000 unfamiliar items. For the majority of these items, the highest responding identity produced a negative response. Just 282 of the 5000 probe faces produced positive responses to any known identity and even these produced very low responses (lower than all but one of the known faces).

The consequence of this pattern of data is that the model, using the ideal parameters, is capable of distinguishing unfamiliar faces from familiar ones using a zero or very low recognition threshold. This also means that only a small proportion of the volume of the face space (about 5%—based on the distribution

[8] Recognition involves distinguishing between items that are very similar. This requires a low level of error. Categorization, however, involves distinguishing between large sets of items. Activations of different identities, which will be high with large error values, add to the response in the categorization and so little is to be gained from very low levels of error. The error merely needs to be sufficient to distinguish between the larger categories.

of the population of faces) can be considered to be designated to particular identities. The remaining 95% is identity free. This is quite different from the Voronoi model, which is 100% designated to particular identities, or the GCM, that will always produce an identity.

Figure 10a shows the distribution of activation of the highest responding identity against the distinctiveness of the probes. This figure shows that there is a positive relationship between the two variables, which is similar in size and

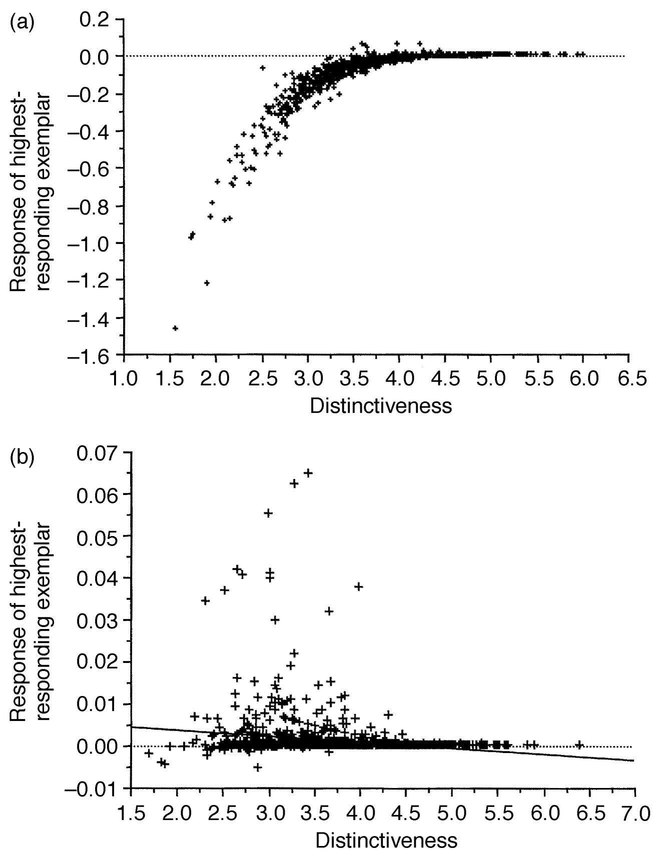


Figure 10. (a) The relationship between the response of the highest-responding exemplar (from the 1000 known exemplars) to a random probe and the probe's distinctiveness. Most responses are negative and there is a positive relationship between the two factors. This is a model of the situation where a new face is seen and one has to decide whether it is new or not. (b) The relationship between the response of the highest-responding exemplar of the 20 target items to each of 1000 random unfamiliar probes and the probes' distinctiveness. Most responses are positive and there is a negative relationship between the two factors. This is a model of the situation where a new face is presented and one has to decide whether it comes from a previous set of 20 faces or not.

shape to that seen for the familiar probe items. This relationship, however, is not in the direction that is consistent with the empirical data. It has been found that false-positive recognition is more likely to occur for typical faces than for distinctive faces (e.g., Hancock, Burton, & Bruce, 1996; Lewis & Johnston, 1997). The typical faces in the current model, however, produce much lower responses than distinctive faces. This is possibly a problem for the model; however, it can be addressed in at least two ways: First, in terms of the nature of the specific tasks employed in the experiments; and second in terms as a variable discrimination criterion. These two accounts are now described.

Hancock and colleagues (1996), and Lewis and Johnston (1997) each employed a face-learning memory task. In these tasks, a set of unfamiliar face was presented to the participants to be learnt. Later, test faces were presented that included the target faces and previously unseen distractor faces. The task, therefore, was to distinguish between items previously seen and unseen items. Subjects knew that famous people or personal acquaintances would not appear in the test and so similarity to the faces of these people did not need to be considered. In the implementation of the model just described, however, the novel probes were tested against being recognized as any known identity (previously seen items and famous individuals). This situation is effectively different from the empirical research that is being modelled. The human participants of the experiments knew that they only needed to distinguish between unknown faces and those in the training set. They could effectively remove all their previously known faces from the decision criteria. They would be working, therefore, with a very much more sparse face space consisting of just the target set they had recently seen. This paradigm is a valid method of experimentation but the way the experiment is modelled needs to be modified so that the face-space-R model only tests a probe test face against the subset of items that had just been learnt. That is, use a restricted set of exemplars when testing false recognition.

If this restricted-exemplar explanation is true then it should be possible to model it using the face-space-R model and produce data that simulate human results. In order to test this, the model was set up with the same ideal parameters but with the number of faces reduced to 20 (i.e., the number of target items employed by Lewis & Johnston, 1997). Probe items were tested within this restricted-exemplar face space for their highest responses. These highest responses were then correlated with their distinctiveness. Figure 10b shows the highest responses from the 20 targets for each of the probes. While the target's identity response is positively correlated with distinctiveness, the figure shows that the size of the strongest responses for the distractors is negatively correlated with distinctiveness. This illustrates how the mirror effects of distinctiveness (i.e., distinctive faces produce less hit errors *and* less false positive errors than typical faces) can be accounted for using the face-space-R model.

While this restricted-exemplar explanation can account for the false positive advantage for distinctive faces in some experiments, it does not necessarily apply to studies where the task is to distinguish between famous and nonfamous faces. Valentine (1991, Exp. 3) reports just such an experiment. He found that typical unfamiliar faces needed longer to be responded to correctly than did distinctive unfamiliar faces. Also, the typical faces produced more false positives. If we wish to apply the restricted exemplar account then it is necessary that the subjects were comparing each face to just a small set of famous individuals that they would predict would occur in such an experiment. This is unlikely to be the case. In order to understand this effect of distinctiveness on false positives it is necessary to consider the relationship between reaction times and the face-space-R model in more detail.

One clue to what is going on in Valentine's (1991) Experiment 3 comes from analysis of the d prime scores.[9] These measure the overall level of performance or how difficult the task is. For the distinctive faces and the typical faces the d prime score was 1.63 and 1.64 respectively. This means that both tasks are just as difficult and it is just as hard to make speeded familiarity decision to distinctive faces as to typical faces. This is unexpected because we know that, due to the distributions of the responses, it should be harder to distinguish familiar and unfamiliar typical faces than familiar and unfamiliar distinctive faces. The reason for the equivalence of the d prime scores, it is proposed, comes from the difference in the observed reaction time. Put simply, the decision to make the familiarity decision is made with the same amount of information for both typical and distinctive faces but the time required to reach this level of information is longer for typical faces. The implementation of this within the face-space-R model is now described.

It has been discussed earlier that during presentation of a probe image, there is an error associated with the probe. For brief presentation there is a large error but as the duration increases this error reaches an asymptote. There is, therefore, a dynamic change that occurs as the face is viewed. Early during the presentation there is a large error. If this produces a positive response by any of the exemplars then a positive familiarity decision can be made. This is more likely to occur for distinctive faces. Typical faces will require the error to be reduced further before any exemplar produces a positive response. This difference explains the reaction time differential observed for recognizing typical and distinctive faces. Figure 11 illustrates this distinctiveness effect on recognition time for recognition. The face-space-R model was implemented with 1000 exemplar, 15 dimensions, and the error value was gradually decreased from 1.5

[9] d prime scores are based on signal detection theory and refer to how difficult a task is independent of what decision criterion being employed by the subject. These are calculated based on numbers of hits and numbers of false positives. The higher the score, the easier the more accurate the decision.

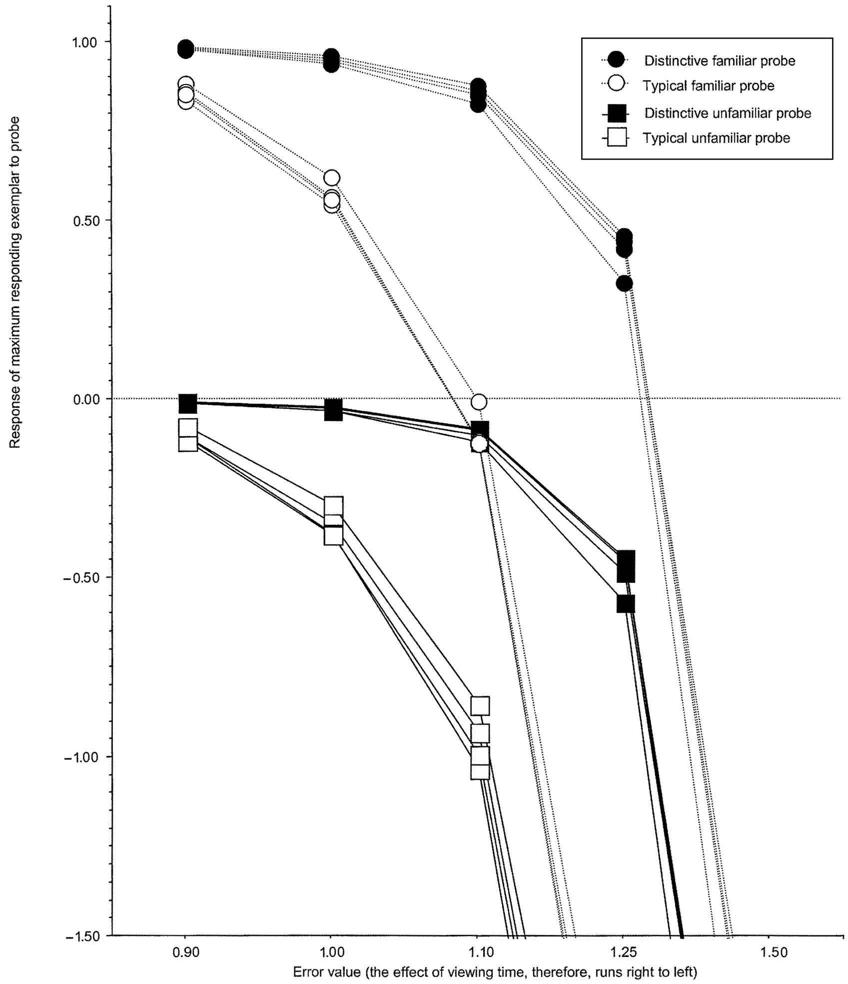


Figure 11. Performance of the face-space-R model on a random selection of typical and distinctive familiar and unfamiliar probes as the error value changes. As the error decreases familiar distinctive items are recognized earlier than typical items and for unfamiliar items the distinctive items reach asymptote earlier than typical items.

to 0.9. It can be observed that the distinctive familiar probes produce positive responses earlier (i.e., for higher levels of error) than typical familiar probes.

Valentine (1991) observed that determining that a face is unfamiliar requires an additional 500 ms to making a familiarity decision, but there is still an advantage for distinctive over typical faces of about 175 ms. During these longer presentation times, the error associated with the probe will continue to decrease towards the asymptote. Figure 11 shows the pattern of maximum response to distinctive and typical unfamiliar probes. These responses increase monotonically as the error decreases. The rate of this increase is faster and reaches asymptote earlier for distinctive than typical faces. In order to conclude that a face is not familiar it is not sufficient for it to produce no positive responses—indeed all faces produce no positive responses initially when the error is still large. To conclude that a face is unfamiliar it is necessary to observe that further decreases in error are unlikely to bring the maximum response above zero. As the gradient of this increase in response becomes shallower one would be able to reach the conclusion that the face is unfamiliar. Distinctive-face responses become shallow earlier than those for typical faces and so one would expect faster responses to distinctive faces than to typical faces, particularly if these responses were being made to a similar accuracy level (i.e., in terms of d prime).

SUMMARY OF THE SCOPE OF THE FACE-SPACE-R MODEL

The face-space-R model is based on the application of Spencian discrimination theory to a high dimensional space. This fully explicit model has allowed an investigation of its predictions. These predictions are most easily explored computationally owing to the mix of random distributions and multidimensional geometry. The resulting computational model was tested with regard to a range of findings in the face recognition literature. While many assumptions have been made in the generation of the computationally explicit model, these assumptions are clearly expressed and, for the most part, uncontroversial.

Using the finding that the caricature for best likeness is 4.4% (Benson & Perrett, 1991) it was possible to set parameters for the computation model based on the assumption that a typical adult can recognize between 1000 and 5000 people. The main parameter that came out of this modelling was that of the dimensionality of the face space which was between 15 and 22. These parameters were then used within the computational model to explore whether the model is consistent with a range of other findings in the face recognition literature.

The model with ideal parameters was found to be consistent with a range of findings from the face-recognition literature. It can explain why distinctive-rated faces are recognized faster than typical-rated faces but are categorized as faces more slowly. The same model can account for the caricature advantage and the relationship between this and distinctiveness. It can account for why the caricature effect is larger for line-drawn faces than for photographic faces and for briefly presented faces than for unlimited presentations. The model was applied to the situation where two races of faces where known and it was shown how it could account for the own-race bias and also the fact that caricature effects could be found with other-race faces but only when they had been exaggerated from the other-race average. The model also can explain why the majority of faces we see are not incorrectly recognized as someone familiar; why typical faces produce more false-positive errors than distinctive errors; and why distinctive unfamiliar faces can be rejected as unfamiliar faster than typical faces.

Face-space-R is based on one account of categorization. This account, however, has a long history (Spence, 1936; Thurstone, 1927) and is still an important and relevant framework. While the face-space-R model may ultimately prove to be based on flawed assumptions, it does offer a way of exploring the predictions of a multidimensional Spencian space. From this, much of the literature concerning the perception of faces can be accounted for. Further, testable predictions can be generated from the model that are not necessarily obvious from the low dimensional analogy. The testing of these predictions will not only inform us as to the utility of face-space-R but also to the generalizability of Thurstonian choice as model for categorization in the form described by Wills and colleagues (2000).

REFERENCES

Benson, P. J., & Perrett, D. I. (1991). Perception and recognition of photographic quality facial caricature: Implications for the recognition of natural images. *European Journal of Cognitive Psychology*, *3*, 105–135.

Benson, P. J., & Perrett, D. I. (1994). Visual processing of facial distinctiveness. *Perception*, *23*, 75–93.

Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *104*, 3–21.

Bothwell, R. W., Brigham, J. C., & Malpass, R. S. (1989). Cross-racial identification of faces. *Personality and Social Psychology Bulletin*, *15*, 19–25.

Brennan, S. E. (1985). The caricature generator. *Leonardo*, *18*, 170–178.

Burton, A. M., Bruce, V., & Dench, N. (1994). What's distinctive about a distinctive face. *Quarterly Journal of Experimental Psychology*, *47A*, 119–141.

Burton, A. M., Bruce, V., & Hancock P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science*, *23*, 1–31.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81*, 361–380.

Burton, A. M., & Vokey, J. R. (1998). The face-space typicality paradox: Understanding the face-space metaphor. *Quarterly Journal of Experimental Psychology*, *51*, 475–483.

Busey, T. A., & Tunnicliff, J. L. (1999). Accounts of blending, distinctiveness and typicality in the false recognition of faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1210–1235.

Byatt, G., & Rhodes, G. (1998). Recognition of own-race and other-race caricatures: Implications for models of face recognition. *Vision Research*, *38*, 2455–2468.

Carey, S. (1992). On becoming a face expert. In V. Bruce, A. Cowey, A. W. Ellis, & D. J. Perrett (Eds.), *Processing the facial image*. Oxford: Oxford University Press.

Craw, I. (1995). A manifold model of face and object recognition. In T. Valentine (Ed.), *Cognitive and computation aspects of face processing*. London: Routledge.

Fortune, S. (1992). Voronoi diagrams and Delauney triangulations. In D. Z. Du & F. K. Hwang (Eds.), *Computing in Euclidean geometry*. Singapore: World Scientific Publishing Company.

Goldstein, A. G., & Chance, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology*, *105*, 47–59.
Hancock, P. J. B. (2000). Evolving faces from principal components. *Behavior Research Methods, Instruments and Computers*, *32*(2), 327–333.
Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory and Cognition*, *24*, 26–40.
Johnston, R. A., & Ellis, H. D. (1995). Age effects in the processing of typical and distinctive faces. *Quarterly Journal of Experimental Psychology*, *48A*, 447–465.
Jones, F. W., Wills, A. J., & McLaren, I. P. L. (1998). Perceptual categorisation: Connectionist modelling and decision rules. *Quarterly Journal of Experimental Psychology*, *51B*, 33–58.
Lee, K. L., Byatt, G., & Rhodes, G. (2000). Caricature effects, distinctiveness and identification: Testing the face-space framework. *Psychological Science*, *11*, 379–385.
Lee, K. L., & Perrett, D. (1997). Presentation-time measures of the effect of manipulations in colour space on discrimination of famous faces. *Perception*, *26*, 733–752.
Lee, K. L., & Perrett, D. (2000). Manipulation of colour and shape information and its consequence upon recognition and best-likeness judgements. *Perception*, *29*, 1291–1312.
Lewis, M. B. (1997). *A computational and empirical investigation into face-space.* Unpublished PhD thesis, Cardiff University, UK.
Lewis, M. B. (1999). Age of acquisition in face categorisation: Is there an instance-based account? *Cognition*, *71*, B23–B39.
Lewis, M. B., & Johnston, R. A. (1997). Familiarity, target set and false positive in face recognition. *European Journal of Cognitive Psychology*, *9*, 437–459.
Lewis, M. B., & Johnston, R. A. (1998). Understanding caricatures of faces. *Quarterly Journal of Experimental Psychology*, *51A*, 321–346.
Lewis, M. B., & Johnston, R. A. (1999a). A unified account of the effect of caricaturing faces. *Visual Cognition*, *6*, 1–42.
Lewis, M. B., & Johnston, R. A. (1999b). Are caricatures special? Evidence of peak-shift in face recognition. *European Journal of Cognitive Psychology*, *11*, 105–117.
Light, L. L., Kayra-Stuart, F., & Hollander, S. (1979). Recognition memory for typical and unusual faces. *Journal of Experimental Psychology: Memory, Learning, and Cognition*, *5*, 212–228.
Nosofsky, R. M. (1986). Attention, similarity and the identification-categorisation relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.
Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 3–27.
Purtle, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, *80*, 408–421.
Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.
Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception*, *17*, 48–63.
Rhodes, G. (1996). *Superportraits: Caricatures and recognition*. Hove, UK: Psychology Press.
Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations. *Cognitive Psychology*, *19*, 473–497.
Rhodes, G., Carey, S., Byatt, G., & Proffitt, F. (1998). Coding spatial variations in faces and simple shapes: A test of two models. *Vision Research*, *38*, 2307–2321.
Rhodes, G., & Tremewan, T. (1994). Understanding face recognition: Caricature effects, inversion and the homogeneity problem. *Visual Cognition*, *1*, 275–311.
Shepherd, J. W., Ellis, H. D., & Davis, G. M. (1977). *Perceiving and remembering faces* [Report Pol/73/1675/24/1]. London: Home Office.
Shin, H. J., & Nosofsky, R. M. (1992). Similarity-scaling studies of dot-pattern classification and recognition. *Journal of Experimental Psychology: General*, *121*, 278–304.

Spence, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, *43*, 427–449.

Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, *44*, 430–444.

Stevenage, S. V. (1995). Can caricatures really produce distinctiveness effects? *British Journal of Psychology*, *86*, 127–147.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–246.

Thomas, D. R., Mood, K., Morrison, S., & Wiertelak, E. (1991). Peak shift revisited: A test of alternative interpretations. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 130–140.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, *34*, 273–386.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, *3*, 71–86.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race in face recognition. *Quarterly Journal of Experimental Psychology*, *43A*, 161–204.

Valentine, T. (2001). Face-space models of face recognition. In M. J. Wenger & J. T. Townsend (Eds.), *Computational geometric and process perspectives on facial cognition: Contexts and challenges* (pp. 83–113). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Valentine, T., & Bruce, V. (1986). The effects of distinctiveness in recognising and classifying faces. *Perception*, *15*, 525–535.

Valentine, T., Chiroro, P., & Dixon, R. (1995). An account of the own-race bias and the contact hypothesis based on a ''face space'' model of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition*. London: Routledge.

Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology*, *44A*, 671–703.

Valentine, T., & Ferrara, A. (1991). Typicality in categorisation, recognition and identification: Evidence from face recognition. *British Journal of Psyuchology*, *82*, 87–102.

Vernon, M. D. (1955). The function of schemata in perceiving. *Psychological Review*, *62*, 180–192.

Wills, A. J., & McLaren, I. P. L. (1997). Generalization in human category learning: A connectionist explanation of differences in gradient discrimination and non-discrminative learning. *Quarterly Journal of Experimental Psychology*, *50B*, 607–630.

Wills, A. J., & McLaren, I. P. L. (1998). Perceptual learning and free classification. *Quarterly Journal of Experimental Psychology*, *51B*, 235–270.

Wills, A. J., Reimers, S., Stewart, N., Suret, M., & McLaren, I. P. L. (2000). Tests of the ratio rule in categorisation. *Quarterly Journal of Experimental Psychology*, *53A*, 983–1011.

Zaki, S. R., & Nosofsky, R. M. (2001). Exemplar accounts of blending and distinctiveness effects in perceptual old–new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1022–1041.

Manuscript submitted January 2003
Manuscript accepted May 2003