Featural and configurational processes in the recognition of faces of different familiarity

Article (Published Version)

Collishaw, S M and Hole, G J (2000) Featural and configurational processes in the recognition of faces of different familiarity. Perception, 29 (8). pp. 893-909. ISSN 03010066

Perception, 2000, volume 29, pages 893 – 909

DOI:10.1068/p2949

Featural and configurational processes in the recognition of faces of different familiarity

Stephan M Collishaw, Graham J Hole
School of Cognitive and Computing Sciences, University of Sussex, Falmer, Brighton BN1 9QH, UK; e-mail: stefanc@cogs.susx.ac.uk
Received 21 June 1999, in revised form 17 April 2000

Abstract. Previous research suggests that face recognition may involve both configurational and piecemeal (featural) processing. To explore the relationship between these processing modes, we examined the patterns of recognition impairment produced by blurring, inversion, and scrambling, both singly and in various combinations. Two tasks were used: recognition of unfamiliar faces (seen once before) and recognition of highly familiar faces (celebrities). The results provide further support for a configurational – featural distinction. Recognition performance remained well above chance if faces were blurred, scrambled, inverted, or simultaneously inverted and scrambled: each of these manipulations disrupts either configurational or piecemeal processing, leaving the other mode available as a route to recognition. However, blurred/scrambled and blurred/inverted faces were recognised at or near chance levels, presumably because both configurational processing and featural processing were disrupted. Similar patterns of effects were found for both familiar and unfamiliar faces, suggesting that the relationship between configurational and featural processing is qualitatively similar in both cases.

1 Introduction

Previous research has shown that the information contained in faces, and which is used to perceive, store, and recognise them, does not simply consist of the sum of the information that can be gained from considering individual features. The spatial relationship between the facial features—their 'configuration'—also appears to be very important for recognition.[1] A variety of evidence supports this view. For example, Young et al (1987) showed that the top half of a celebrity's face can be recognised with reasonable accuracy when it is presented in isolation, but that recognition is markedly slower when it is combined with the bottom half of a different face. This effect is only found, however, when the two halves are closely aligned so as to create the impression of a new face, and not when the two face halves are misaligned. This chimeric face effect cannot therefore be explained in terms of interference from the processing of information from other, irrelevant features. Instead, it appears that features are processed interactively when they form a facial configuration (Suzuki and Cavanagh 1995). The identification of individual facial features also relies on the processing of configurational information to some extent. For instance, individual features presented in the context of the whole face are more easily identified than when they are presented in isolation or in the context of a scrambled face, or an upside down face (Tanaka and Farah 1993).

A variety of evidence suggests that featural information also provides an important source of information for face processing. Bruyer and Coget (1987) and Tanaka and Farah (1993) have shown that individual features can be recognised with moderate accuracy, even when presented in isolation or in the context of a jumbled face.

[1] See Searcy and Bartlett (1996) for a clear summary of the various theoretical interpretations of 'configurational', 'holistic', or 'relational' processing. The terminology in this area is rather muddled at present, with these terms being used interchangeably but also being used to refer to quite different types of processing. We are using 'configurational' processing to refer to a form of processing which is primarily based on information about the spatial relationships between local 'features' (eg eyes, nose, and mouth). This is in contrast to 'piecemeal' or 'featural' processing, which would be based primarily on details of the individual facial features rather than on their spatial relationships to each other.

The distinction between configurational and featural information is influential in many accounts of face processing (Sergent 1984; Diamond and Carey 1986; Rhodes et al 1993; Searcy and Bartlett 1996). However, a number of important questions remain. What exactly is the relationship between configurational and featural processing modes? Can face recognition be conceived of as a dual process mechanism, with a dissociation between the processing of featural and configurational information? Are both forms of information of similar importance in the recognition of faces? Are they organised hierarchically, with a feature-based analysis preceding the processing of configurational information, as suggested by Rhodes et al (1993)? Can information derived from the facial image be dichotomised as configurational or featural in nature, or does it instead fall on a configurational – featural continuum?

Tanaka and Farah (1993) and Farah et al (1998) have questioned the validity and usefulness of models of face recognition based on the configurational – featural distinction. Instead, they have proposed an alternative account, based on the extent to which image representations are decomposed into their component parts. According to their model, face perception is different from other forms of recognition in that faces are mainly represented as un-decomposed wholes (ie holistic representations).

The first aim of the present study was to examine these issues by studying the disruptive effects on face recognition of inversion, scrambling, and blurring. While each of these manipulations has been studied in isolation, our purpose here has been to gain further insight into the relationship between featural and configurational processing modes by examining how the effects of these manipulations might interact.

It is now well documented that upside-down faces are difficult to recognise (Yin 1969, 1970; reviews in Diamond and Carey 1986, and Valentine 1988). While the disruption of shape-from-shading relations in inverted faces contributes to this effect (eg Johnston et al 1992; Enns and Shore 1997), the prime reason for difficulty seems to be that inversion disrupts the processing of configurational information (eg Sergent 1984; Endo 1986; Young et al 1987; Rhodes et al 1993; Lewis and Johnston 1997). Numerous studies suggest that sensitivity to the relative positions of features within a face is reduced when they are inverted (eg Kemp et al 1990; Bruce et al 1991; Bartlett and Searcy 1993). Whilst it is debatable whether the disruption of configurational information is complete or partial (see Valentine 1988), it does appear that the processing of featural information is not particularly strongly affected by inversion. For example, Searcy and Bartlett (1996) showed that faces made to look grotesque by marked changes in the spatial relationship between their features no longer looked grotesque when inverted, whereas faces which had been made grotesque by isolated feature changes (eg the addition of fangs) remained grotesque when seen upside down (see also eg Endo 1986; Bruyer and Coget 1987; Rhodes et al 1993).

A number of researchers (eg Bruce 1986; Johnston et al 1996) have attempted to directly disrupt the perception of the basic configuration of the facial features by 'scrambling' a face (rearranging its component features). It should not be assumed, however, that scrambling a face eliminates its configurational properties completely, or that featural information necessarily remains unimpaired. The disruption to each process will depend on the precise nature and number of the changes made, but it can be reasonably assumed that configurational information will be disproportionately affected by scrambling.

In contrast, blurring (specifically, the removal of the high spatial frequencies in a face image) is likely to affect featural processing more than configurational processing at intermediate levels of blur (Costen et al 1994). At one extreme (blur of up to 16 cycles per face), neither configurational processing nor featural processing are likely to be particularly affected and recognition is likely to be unimpaired. At the other extreme (blur of more than 8 cycles per face), both featural and configurational processing will be affected, since the information remaining in the image is likely to be too coarse to allow even very general patterns to be perceived. Within this range, as a face is progressively blurred, featural information will be masked at an earlier stage and more completely than configurational information (Costen et al 1994). Previous studies of the effects of blurring on face recognition (eg Harmon 1973, Hayes 1988, Sergent 1986) have shown that recognition survives levels of blurring that remove fine-detail information about the facial features, suggesting that fairly coarse configurational information may be sufficient for recognition to be achieved.

If configuration and feature are analysed by two routes in face recognition, and if these are differentially affected by inversion, scrambling, and blurring, then two hypotheses can be formulated: first, any one of these three manipulations will lead to some impairment in recognition performance; second, these manipulations should show interactive effects if applied together. Scrambling and inversion each affect configurational processing and should therefore have relatively little (if any) additional impact together compared to the effects of either of these manipulations alone. Blurring plus scrambling, and blurring plus inversion, are expected to reduce recognition rates markedly, possibly to chance levels, as both configuration-based and feature-based processes are disrupted. This experiment is designed to test these hypotheses by comparing recognition accuracy for unmanipulated faces to accuracy for faces which have been blurred, scrambled, inverted, or subjected to combinations of these manipulations (scrambled and inverted; blurred and scrambled; or blurred and inverted). Response latencies for these different conditions will also be examined as these are able to provide additional information about the particular demands placed on processing by the manipulations employed here.

This study provides a long overdue test of a 'two-process' theory of face recognition where configurational and featural processing are posited as two dissociable routes to the identification of a face. According to such a theory, access to both routes is optimal, access to either route may be sufficient, but at least one route must be available for recognition to occur.

The second aim of this study is to provide further data on the relationship between the processing of familiar and unfamiliar faces. Surprisingly few studies have attempted to address questions relating to how a face becomes familiar, or how the representations used to store faces change with repeated exposure to (and recognition or recall of) such faces. First, it is possible that qualitatively different facial information or processing mechanisms are involved in processing familiar and unfamiliar faces. Second, an alternative account is that unfamiliar and familiar faces simply differ on a continuum of the strength of the underlying memory traces, but that these memory traces encode essentially the same types of facial information.

Support for a qualitative distinction between the processing of familiar and unfamiliar faces comes from a number of sources. First, experimental evidence shows that the internal features are of relatively greater importance to the processing of familiar faces than to the processing of unfamiliar faces (Ellis et al 1979; Young 1984). One explanation for these findings is that the internal features carry greater configurational information than the outer features, and that this information predominates in the processing of familiar faces (Campbell et al 1999). However, an alternative explanation can also account for this effect. The external features of the face are those that are most likely to change over time and with age (eg haircut, facial shape), and are therefore least likely to give reliable cues regarding the identity of faces stored for longer periods of time (Young 1984; Bruce and Young 1998).

Second, neuropsychological evidence supports a double dissociation between the matching of unfamiliar faces and the recognition of famous faces (Benton 1980; Malone et al 1982). However, such studies have in general confounded the familiarity of the stimuli with the nature of the task (face recognition versus face matching). It is therefore unclear whether a true double dissociation between the processing of facial stimuli of low and high familiarity does in fact exist. A second important point is that one must be clear whether one is comparing faces of high and low familiarity or faces of high and no familiarity. The present study is mainly concerned with the first question, whilst much of the neuropsychological evidence only directly bears on the second question.[2]

Few studies have directly addressed the question whether familiar and unfamiliar faces are processed differently in terms of the balance between configurational and featural analysis. Scapinello and Yarmey (1970) found that the effects of inversion did not interact with the number of exposures at study, nor did they differ for famous and non-famous stimuli (Yarmey 1971). Young et al (1987) and Hole (1994) demonstrated that the chimeric face effect operates with unfamiliar faces, as well as with highly familiar faces. These studies suggest that configurational and featural processes are both used in the processing of faces of varying familiarity. Further research is needed to establish whether the balance between the two forms of processing changes with the increasing familiarity of the face. In particular, no studies to date have allowed a formal test of an interaction between face familiarity and manipulations that reveal the relative involvement of configurational and featural processes.

The question whether recognition of unfamiliar faces relies to a greater extent on either configurational or featural facial information than does recognition of highly familiar faces can be addressed by comparing the pattern of results over the seven conditions for the two experimental tasks. For example, if configurational information becomes more important with the increasing familiarity of a face, and a featurally oriented processing mechanism is used to encode and recognise unfamiliar faces, then performance on a task of recognising celebrities should be relatively more affected by scrambling and/or inversion. In contrast, blurring would be expected to have a greater detrimental impact on subjects' recognition of unfamiliar faces. If, however, the processing strategies for unfamiliar and familiar faces do not differ in their use of featural or configurational information, then no difference in the pattern of the effects of the three manipulations is expected between the two tasks.

2 Method

2.1 *Subjects*

One hundred and forty subjects (sixty-eight men and seventy-two women), ranging in age from 18 to 50 years participated in the study. Subjects were randomly assigned to one of the seven experimental conditions. Seven further subjects had been tested, but were excluded from the study for various reasons.[3]

2.2 *Design*

A mixed-subjects design was used. Each participant was assigned to one of seven conditions, reflecting various combinations of the following manipulations: (1) normal (not manipulated); (2) blurred; (3) scrambled; (4) inverted; (5) scrambled and inverted; (6) blurred and scrambled; (7) blurred and inverted. All subjects completed two tasks. Task 1 tested the recognition of famous faces, and task 2 tested the recognition of previously learnt unfamiliar faces.

[2] For the sake of simplicity we use the term 'unfamiliar' to refer to faces of low familiarity. In this study, faces of low familiarity are ones which have been presented for study minutes earlier for a duration of several seconds.

[3] These included subjects who responded uniformly to one or more blocks of stimuli (either all "yes" or all "no"), and subjects who were excluded owing to computer failure part-way through the experiment.

Each task consisted of two blocks of faces: first a block of unmanipulated 'control' faces (which were identical for all participants), and then a block of manipulated 'experimental' faces (with the manipulation varying according to experimental condition). Performance on control faces was examined in order to equate the face recognition abilities of the different groups of subjects in subsequent analyses. The order of the two tasks was counterbalanced between subjects and across conditions. Subjects were assigned to the same condition for both tasks.

2.3 *Materials and apparatus*

All faces used in both tasks were of clean-shaven, Caucasian, adult men. For task 1, photographs of famous people were taken from a variety of magazines. These included well-known celebrities from the worlds of show business, sports, and royalty. Photographs of non-famous distractor faces were taken from a variety of Dutch magazines. Although unknown in the UK, these faces were well-known celebrities in the Netherlands, and were individually matched to the target faces on the basis of age, hair colour and hair length, and quality of image. The reason for using Dutch celebrities as distractors was to control for extraneous characteristics that might indicate that a particular face is famous (eg quality of images taken from magazines, hairstyle and/or makeup, pose, and expression). Twenty-two target faces and twenty-two distractor faces were used.

The majority of photographs used in the second task were taken of students and employees at University College London. In addition, a small number of faces were taken from sets 1 and 2 of the 'PICS' database at Stirling University. Two photographs of each person were used. These differed in view (full face or three-quarter view) and expression (smiling or 'neutral'). Twenty-two faces were randomly designated to be target faces. One photograph of each face was used in the study phase, and the other in the test phase of the experiment. An equal number of each view and expression were used in each phase of the experiment. Distractor faces were matched with the target faces used in the test phase on the basis of the following criteria: view, expression, quality and source of image, hair colour and hair length, and age.

Faces were prepared in Adobe Photoshop by replacing all the background information with a uniform grey background, orienting the face upright (if necessary), and scaling it to a standard size across the width of the face (450 pixels). For both tasks, half the faces served as control faces and were not manipulated further, whilst the other half were experimental faces. Seven versions of each experimental face were prepared for use in the seven conditions of the experiment. Blurred images were prepared by using the Gaussian filter available in Photoshop. A filter with a radius of 10 pixels was used. Scrambled faces were prepared by dividing them into five horizontal strips, and rearranging these in the following order from top to bottom: (a) mouth; (b) nose; (c) forehead and hair; (d) chin; (e) eyes.

Examples of the stimuli used in this study are shown in figure 1 (famous face recognition) and in figure 2 (unfamiliar face recognition).

Images were presented with purpose-written software for the PC. The software recorded the subject's choice (left or right mouse button), and the time between the onset of the stimulus and the mouse button press by the subject. The images measured 140 mm by (approximately) 165 mm. All subjects were tested at a viewing distance of 60 – 70 cm. The images thus subtended about 12.15 deg horizontally.

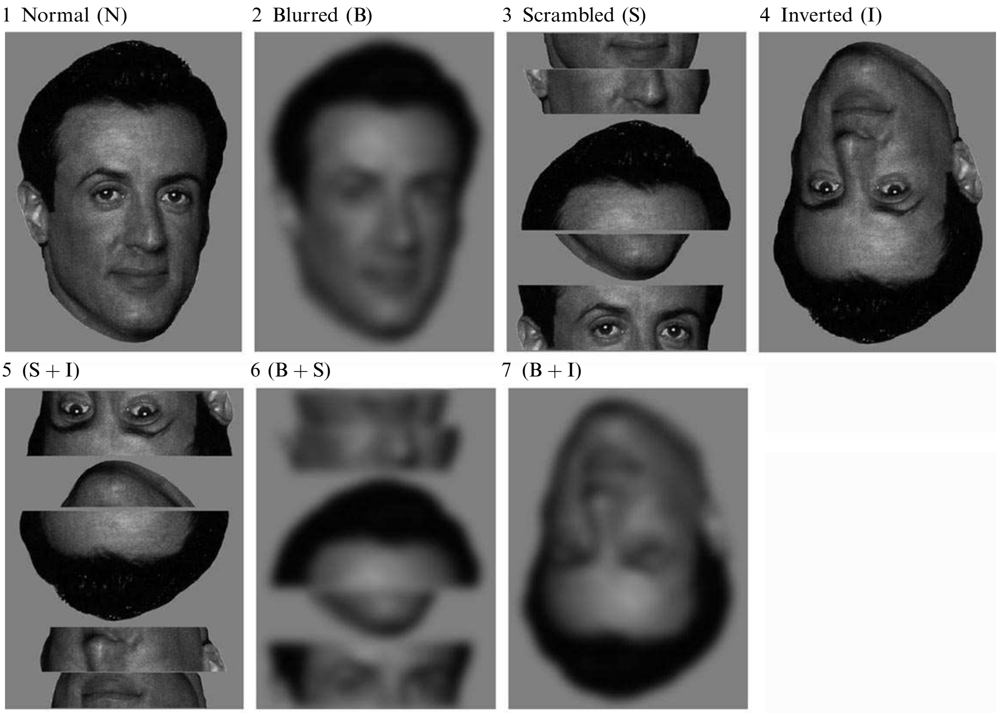

Figure 1. Recognition of celebrities. Sample stimuli.

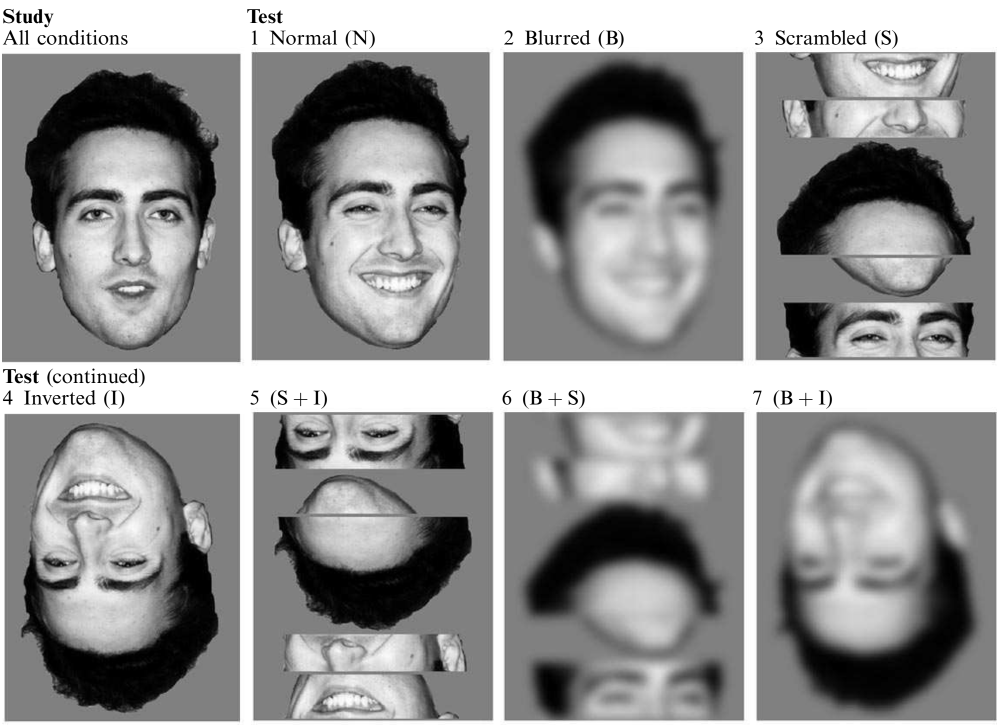

Figure 2. Learning and recognition of unfamiliar faces. Sample stimuli.

2.4 *Procedure*

2.4.1 *Task 1: Recognition of celebrities.* The task comprised two blocks of faces. Each block consisted of 20 faces: 10 celebrities and 10 nonentities. The first block of faces was made up of the set of unmanipulated control faces. All subjects were shown this block of faces first, regardless of experimental condition. Subjects were then shown the block of manipulated experimental faces. The order of presentation of the stimuli within each block was randomised for each subject.

Subjects were requested to respond as quickly and as accurately as possible and received practice trials to familiarise them with the experimental procedures. On each trial a face was displayed and the subjects pressed the left button on the mouse if they thought it was a celebrity, and the right button if they thought it was a nonentity. Stimuli were displayed for three seconds or until the subject responded. There was a delay of three seconds between the subject's decision and the onset of the next trial.

2.4.2 *Task 2: Recognition of unfamiliar faces.* The stimuli again consisted of one set of control stimuli and one set of (manipulated) experimental stimuli. Half of each set were designated as targets and half as distractor faces. As described above, a second photo (differing in view and expression) of each of the 'target' faces was shown to subjects in the study phases of this task. These stimuli were always unmanipulated. All subjects were first asked to learn the eleven control faces, received two practice trials (1 target, 1 distractor), and were then tested on the remaining control faces (10 targets, 10 distractors). Subjects were then asked to learn the 11 experimental faces, received two practice trials, and were then tested on the remaining experimental faces (10 targets, 10 distractors). In the learning phase, stimuli were displayed one at a time for 3 s and separated by delays of 3 s. The procedures used to display images and assess subjects' choices in the practice and test phases were identical to those used in task 1.

3 Statistical analysis

d' scores (Miller 1996) are used as a measure of recognition accuracy in all the analyses reported here. Perfect performance on these tasks corresponds to a d' of 3.29. A d' of 0 or below indicates chance level performance. The mean percentage of correct responses is also reported to allow for comparison with other studies.

Analyses of reaction time (RT) were restricted to latencies for correct responses. To reduce the error that often results from extreme scores within a RT paradigm (Ulrich and Miller 1994), an outlier analysis was performed for each task, and mean $+2$ SD cutoffs were calculated for each condition. Between 4% and 5% of responses were excluded when mean response times were calculated for each subject. As a direct measure of the effect of a particular manipulation on subjects' reaction times, 'RT-difference scores' were derived by subtracting subjects' mean RTs for control stimuli from mean RTs for experimental stimuli. In some analyses hit and correct rejection times are examined separately.

Results were analysed by one-way analyses of variance to test for main effects of condition on measures of accuracy and response speed separately for each task. For analyses showing significant main effects, Tukey HSD tests were used to compare the performance in the seven conditions. One-sample t-tests were used to compare recognition accuracy in conditions 5 (S+I), 6 (B+S), and 7 (B+I) with that expected by chance. Comparisons of effects between the two tasks were made by using mixed-design ANOVAs with condition as the between-subjects factor, and task as the within-subjects factor. Additional analyses covaried subjects' performance on control faces. Results are comparable, and for the sake of simplicity they are not reported here.

4 Results

4.1 *Access to configurational and featural information: An analysis of recognition accuracy*

4.1.1 *Task 1: Recognition of celebrities.* Table 1 shows mean and standard deviation d' scores for experimental stimuli for both tasks for subjects in conditions 1 to 7. Figure 3 illustrates the effects of the different manipulations on percentages of correct responses, relative to subjects' control baseline.

A highly significant main effect of condition on subjects' ability to discriminate famous from non-famous faces was found for experimental faces ($F_{6,133} = 23.6$, $p < 0.0001$). The results of the range test show that manipulated faces in general were harder to recognise than normal (unmanipulated) faces. Specific effects between types of manipulation were also found. d' scores in conditions 2 (B), 3 (S), 4 (I), and 5 (S+I) were significantly higher than in conditions 6 (B+S) or 7 (B+I), and there were no significant differences within these two subsets of conditions.

This pattern of results is in accordance with the experimental hypotheses. A further test of these hypotheses is whether subjects' performance in conditions with

Table 1. Recognition accuracy (d') by condition and task.

	Condition[a]						
	1 (N)	2 (B)	3 (S)	4 (I)	5 (S + I)	6 (B + S)	7 (B + I)
Task 1: celebrities[b]							
mean	1.93	1.11	1.14	0.81	0.61	−0.11	−0.18
SD	(0.62)	(0.80)	(0.79)	(0.63)	(0.72)	(0.55)	(0.63)
Task 2: faces seen once[c]							
mean	1.15	0.77	0.76	0.61	0.89	0.40	0.07
SD	(0.87)	(0.49)	(0.48)	(0.56)	(0.56)	(0.64)	(0.78)

[a] N = normal; B = blurred; S = scrambled; I = inverted; only responses for experimental stimuli are described here. See figures 3 and 4 for control responses.
[b] Task 1 (celebrity faces): $F_{6,133} = 23.6$, $p < 0.0001$; Tukey HSD: 1 > 2, 3, 4, 5 > 6, 7.
[c] Task 2 (unfamiliar faces): $F_{6,133} = 5.91$, $p < 0.0001$; Tukey HSD: 1, 2, 3, 5 > 7; 1 > 6.

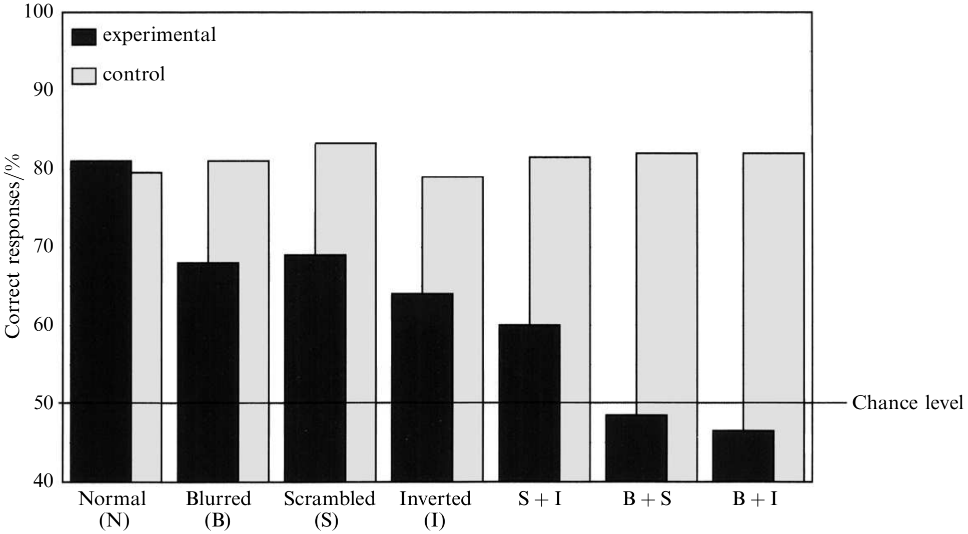

Figure 3. Recognition of celebrities. Effects on recognition accuracy of blurring, scrambling, and inversion alone, and in combination.

combinations of more than one manipulation differs significantly from chance. This was assessed by examining whether subjects' d' scores were significantly different from zero. Recognition performance in condition 5 (S+I) was significantly greater than chance ($t_{19} = 3.77$, $p = 0.001$), whilst recognition performance in conditions 6 (B+S) and 7 (B+I) did not differ from that expected by chance ($t_{19} = -0.86$, $p = 0.4$; $t_{19} = -1.30$, $p = 0.2$, respectively).

The seven groups of subjects did not differ in their ability to discriminate the unmanipulated control faces ($F_{6,133} = 0.41$, $p = 0.9$).

4.1.2 *Task 2: Recognition of unfamiliar faces.* Subjects' d' scores for experimental stimuli in the seven conditions are displayed in table 1. Figure 4 illustrates the effects of the different manipulations on percentages of correct responses. There was a similar pattern of effects to that shown for task 1. First, all manipulations appeared to result in some impairment in performance. Second, blurring, scrambling, and inversion resulted in rather similar decrements in performance without appearing to impair subjects' performance altogether. Third, performance in condition 5 (S+I) appears to be rather more similar to the recognition accuracy of subjects in conditions 2 (B), 3 (S), and 4 (I) than to that of subjects in conditions 6 (B + S) and 7 (B + I).

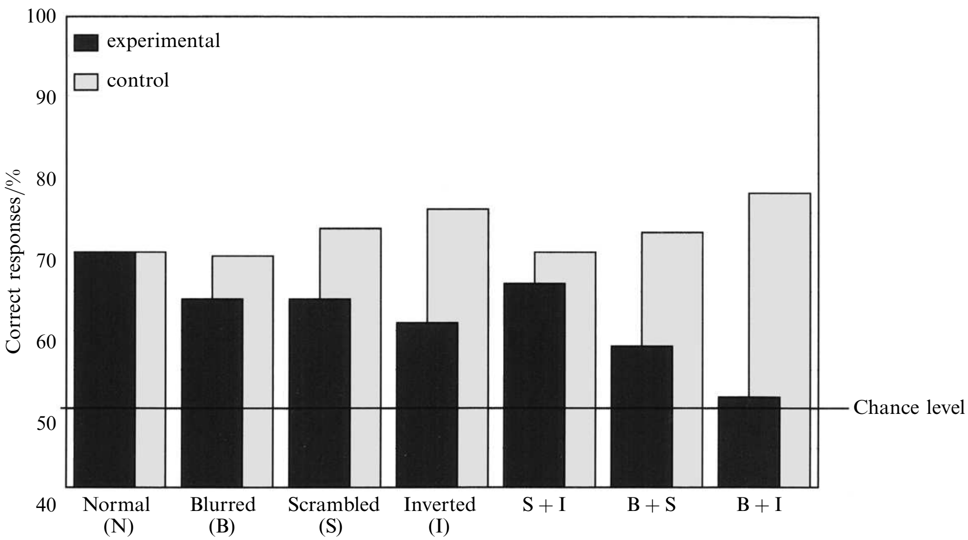

Figure 4. Recognition of faces seen once. Effects on recognition accuracy of blurring, scrambling, and inversion alone, and in combination.

Although a significant main effect of condition was found for subjects' discrimination for the experimental set of faces, ($F_{6,133} = 5.91$, $p < 0.0001$), the pattern of statistically significant differences between conditions was less clear-cut—the range tests showed that discrimination in condition 7 (B+I) was significantly lower than that in conditions 1 (N), 2 (B), 3 (S) and 5 (S+I), whilst d' scores in condition 6 (B+S) were also significantly lower than those in condition 1. No group differences in recognition performance were found for control faces ($F_{6,133} = 1.28$, $p = 0.3$).

Accuracy in condition 5 (S+I) was significantly greater than that expected by chance ($t_{19} = 7.12$, $p < 0.0001$), and performance in condition 6 (B+S) only just exceeded the level expected by chance ($t_{19} = 2.80$, $p = 0.01$), whilst performance in condition 7 (B+I) was at chance level ($t_{19} = 0.43$; $p > 0.6$).

4.1.3 *Accuracy of recognition of familiar and unfamiliar faces compared.* As expected, subjects in general found the second task (unfamiliar face encoding and recognition) rather more difficult than the first (discrimination of celebrities from nonentities), with overall accuracy rates of around 70% and 80% respectively. Indeed, a paired-samples t-test found a significant difference between subjects' d' scores for the unmanipulated sets of control faces used in the two tasks ($t_{139} = 7.34$, $p < 0.0001$). No interaction between task and condition was found for control faces in a mixed-design analysis of variance ($F_{6,133} = 1.34$, $p > 0.2$), suggesting that no adjustments need necessarily be made in subsequent models for subjects' control performance. Analyses that covaried subjects' scores for the control faces did not differ in their results. For the sake of simplicity, analyses without these terms are reported.

Differences in discrimination performance between the two experimental tasks and by experimental condition were examined in an analysis that included both the terms for the within-subjects factor (task) and the between-subjects factor (condition). This analysis showed a significant interaction between task and condition ($F_{6,133} = 5.39$, $p < 0.0001$). This interaction to a large extent reflects the reduced impact of any manipulation for the unfamiliar-face recognition task. Whilst unfamiliar-face recognition was clearly harder than familiar-face recognition in this study, the relative impact of the variety of manipulations investigated here was rather less for the unfamiliar-face recognition task (compare figures 3 and 4).

Given the interest in examining the possibility of a differential impact on the two tasks of particular manipulations compared against one another (rather than the comparison of any versus no manipulation implicit in the above analyses), a further mixed-design analysis of variance, restricted to subjects in conditions 2 to 6, was conducted (subjects in condition 7 were also excluded as they performed at chance level on both tasks). This analysis continued to show a significant interaction between task and condition ($F_{4,95} = 4.32$, $p = 0.003$). Examination of the means in table 1 suggests that this result reflects a greater difference between accuracy in conditions 2 to 5 and accuracy in condition 6 for famous faces than for unfamiliar faces. However, there does not appear to be any difference in the relative impact of blurring, scrambling, and inversion when applied alone. Indeed, no significant interaction is found between task and condition when the analyses are further restricted to conditions 2 (B), 3 (S), and 4 (I)—($F_{2,57} = 0.23$, $p > 0.7$).

4.2 *Processing demands: An analysis of response speeds*

4.2.1 *Task 1: Recognition of celebrities.* Mean response times (and standard deviations) for correct trials are shown in table 2 separately for responses to the experimental set of faces and for the unmanipulated control faces. Furthermore, as a direct measure of the effect of the various types of manipulation on response speed, mean RT-difference scores (ie experimental RT minus control RT) have been calculated and are also summarised in table 2. As performance on the experimental stimulus set in conditions 6 (B+S) and 7 (B+I) was no better than chance, the reaction time data for these conditions will not be examined.

A significant main effect of condition on subjects' RT-difference scores was found ($F_{4,95} = 24.13$, $p < 0.0001$), whether only hits ($F_{4,95} = 16.35$, $p < 0.0001$), or correct rejections ($F_{4,95} = 18.54$, $p < 0.0001$) are considered. A range test showed that subjects' responses were slowest in condition 5 (S+I). In addition, blurring, inversion, or scrambling alone, all delayed responses relative to no manipulation (condition 1).

Table 2. Response time (RT, in milliseconds) and RT-difference scores for tasks 1 and 2: effects of condition.

	Condition[a]					
	1 (N)	2 (B)	3 (S)	4 (I)	5 (S + I)	6 (B + S)[b]
Task 1: celebrities						
experimental stimuli						
mean	1259	1619	1947	1578	2497	-
SD	(293)	(368)	(430)	(436)	(843)	-
control stimuli						
mean	1319	1233	1187	1118	1209	-
SD	(334)	(211)	(170)	(198)	(312)	-
RT-difference[c]	−60	+386	+760	+460	+1288	-
Task 2: faces seen once						
experimental stimuli						
mean	1297	1573	1913	1612	2375	2129
SD	(287)	(339)	(506)	(426)	(683)	(640)
control stimuli						
mean	1370	1376	1282	1310	1345	1414
SD	(281)	(291)	(267)	(318)	(276)	(427)
RT-difference[d]	−73	+197	+631	+302	+1031	+715

[a] N = normal; B = blurred; S = scrambled; I = inverted.
[b] Condition 6 at chance level for task 1. Condition 7 at chance level for tasks 1 and 2.
[c] RT-difference (experimental − control): $F_{4,95} = 24.13$, $p < 0.0001$; Tukey HSD: 5 > 2, 3, 4 > 1.
[d] RT-difference (experimental − control): $F_{5,114} = 20.79$, $p < 0.0001$; Tukey HSD: 5, 6 > 1, 2, 4; 5 > 3 > 1, 2; 4 > 1.

It is worth noting that a disproportionate effect of scrambling on subjects' reaction times becomes apparent when correct rejections are examined separately. Mean RT-difference scores for correct rejections are as follows—condition 2 (blurring): +231 ms; condition 3 (scrambling): +787 ms; condition 4 (inversion): +252 ms. A range test on the analysis of RT-difference scores for correct rejections shows a significant difference between scrambling on the one hand and blurring and inversion on the other.

4.2.2 *Task 2: Recognition of unfamiliar faces.* The main effect of condition on subjects' RT-difference scores was again significant ($F_{5,114} = 20.79$, $p < 0.0001$). Range tests showed that responses were significantly slower in condition 5 (S+I) than in conditions 1 to 4. Increased latencies were also found for scrambling and inversion compared to no manipulation, and for scrambling compared to blurring. Scrambled faces were also found to take longer to reject as unfamiliar than inverted faces.

4.2.3 *Familiar and unfamiliar face recognition compared: reaction times.* The greater difficulty of the unfamiliar-face recognition task was also reflected in subjects' response latencies. Reaction times for control faces in this task were on average 152 ms longer than for recognition of celebrities, $t_{139} = -5.99$, $p < 0.0001$.

A mixed design analysis of variance, for conditions where performance was above chance on both tasks (conditions 1 to 5), showed a main effect of task ($F_{1,95} = 12.64$, $p = 0.001$), but no significant interaction between task and condition ($F_{4,95} = 0.92$, $p > 0.4$). As figure 5 shows, all manipulations have a greater effect on RTs for familiar face recognition, but the patterns of RTs across the different manipulations do not differ for the two tasks.

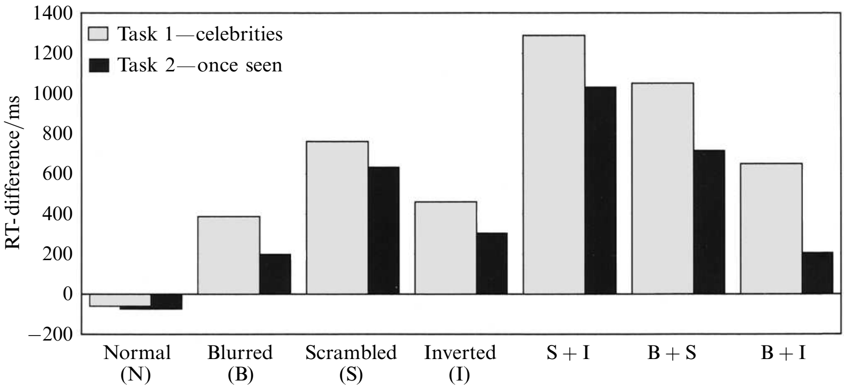

Figure 5. Familiarity effects. Effects on recognition times of the different conditions compared for the two experimental tasks. RT-difference scores (experimental – control) are a direct measure of the effect of each manipulation on response speed. The pattern of effects is similar for faces of high and low familiarity.

5 Discussion

The results of this study show that blurring, scrambling, and inversion each lead to the loss of important information, as shown by subjects' reduced recognition accuracy in these conditions. However, information carried predominantly by relatively low spatial frequencies is sufficient for faces to be recognised much of the time. Similarly, information available in scrambled faces, and preserved after inversion, allows many faces to be identified accurately. The conclusion that these are two quite different forms of information is based on the interaction between the three manipulations when applied in combination. Scrambled/inverted faces were as accurately recognised as upright/scrambled faces and intact/inverted faces. In contrast, recognition accuracy for blurred/inverted faces and for blurred/scrambled celebrities was at chance level, and blurred/scrambled unfamiliar faces were recognised only just above chance level. These data cannot be explained in terms of a greater difficulty with recognising blurred faces. The three manipulations on their own led to similar decrements in recognition accuracy. Instead, the results imply that blurring on the one hand, and scrambling and inversion on the other, disrupt two separate, but complementary sources of information that may be used to identify a face. At least one of these types of information appears to be necessary for face recognition to occur, whilst together they allow normal recognition performance.

What then are these two sources of information that underlie face recognition? Previous research and theory suggest that one describes local features and details of the face, whilst the other is concerned with the configuration of the facial features or holistic representation of the face. As outlined before, as a face is progressively blurred, local detail carried by higher spatial frequencies is lost earlier and more completely than global configurational information carried by lower spatial frequencies. Scrambling and inversion can be expected to differentially impair the process of extracting information relating to the configurations of features. Whilst this general conceptualisation of face recognition is, of course, not new, the present study suggests that the two forms of information allow separate routes in the recognition of faces, and thus supports a dual process conceptualisation of face recognition based on the configurational – featural distinction. Furthermore, many previous accounts have emphasised the importance, if not the dominance, of configurational processing (eg Diamond and Carey 1986; Bruce et al 1991). Whilst it is hard to conceive of a way to quantify and compare the extent of disruptions at local and configurational levels, the present results do show that both routes of processing are important in face recognition. There is no suggestion that disruption of one process leads to greater recognition impairment than disruption of the other.

Whilst the analysis of recognition accuracy reveals the availability of information at the outcome of processing, response latencies may give more detailed insights on the demands placed on processing by our manipulations. As pointed out by Tanaka (1999, personal communication), the equivalence of the effects of scrambling and inversion on recognition accuracy may mask important differences between the manner in which each manipulation impairs recognition. Scrambling directly disrupts the spatial configuration of facial features. In contrast, inverted faces retain the original configurational information, but it appears that people have great difficulty in retrieving this information from upside-down faces. Indeed, the results from the reaction-time analyses did not follow the same pattern as those for the accuracy data. Whilst the accuracy data point to some equivalence between the effects of scrambling and inversion, reaction times suggest an additive effect of the two manipulations. On both tasks, the RT-difference scores for scrambled/inverted faces are approximately equal to the sum of the RT-difference scores for the scrambled and inverted conditions. This suggests that scrambled faces and inverted faces each preserve the same information, but are subject to different processing demands (or strategies) that lead to this information being analysed in different lengths of time. One explanation consistent with this conclusion is that the scrambled faces were processed serially, feature-by-feature, whilst the information in the intact faces (whether upright or inverted) was processed in parallel. Second, inverted faces may need to be mentally righted before they can be analysed further (Rock 1973). Thus scrambled faces would be recognised somewhat slower than intact faces, due to the use of a serial processing strategy; inverted faces are processed slower than upright faces, due to the need to mentally rotate the stimuli first; whilst scrambled/inverted faces would yield the slowest RTs, due to additive effects of these two requirements. However, scrambled, inverted, and scrambled/inverted faces would each eventually yield the same information (local feature information), and are thus recognised equally accurately.

Further research is necessary to examine these possibilities. However, it is worth noting one aspect of the findings that supports an interpretation in terms of serial processing for the scrambled stimuli. Separate analyses of hits and correct rejections showed that rejection times were disproportionately affected by scrambling (but not by inversion or blurring). This is clear evidence that scrambled faces are processed with a serial feature-by-feature strategy. In particular, subjects are likely to examine the most salient face-strip first. Targets may well be identified without further need to examine the rest of the face. However, subjects are likely to examine all the regions of a face before it is rejected as unfamiliar.

One question of considerable importance is whether face recognition can best be characterised in terms of a distinction between featural and configurational processes as suggested here, or by a dissociation between whole and part-based representations (Tanaka and Farah 1993; Farah et al 1998). Farah et al (1998) argue that face recognition can be distinguished from other forms of recognition by the level to which representations are decomposed into separable parts. According to their theory, face recognition relies to a relatively greater extent on holistic unitary representations of faces. Evidence for this view comes from a number of studies. Tanaka and Farah (1993) showed that the context of the whole face aids identification of an individual feature. Their interpretation was that the whole-face condition maps more easily onto a stored holistic representation of the face. However, an alternative explanation is that emergent configurational information in the whole face condition aids recognition of the individual part. Farah et al (1998) demonstrated that irrelevant whole-face masks lead to greater interference than scrambled-face masks. Again, these results are consistent with both models: the whole-face masks more closely correspond to the posited holistic representations held by subjects, but also contain configurational information that is not contained in the scrambled masks. Whilst Farah et al (1998) allow for the finding that isolated parts can usually be recognised above chance levels by suggesting that a mixed population of holistic and local-part representations may contribute to face recognition, they also argue that part decomposition is minimal for faces. They thus suggest that a theory based on a distinction between holistic and part-based representations is preferable to one based on the configurational – featural distinction.

The dissociation between part-based and holistic representations can account for the interaction between the effects of scrambling, inversion, and blurring. Scrambling and inversion would both be expected to disrupt holistic processing, and lead to a reliance on part-based representations, whereas blurring would presumably disrupt part-based representations of local features more than a holistic representation of the whole face. However, Farah et al's (1998) model emphasises the dominance of holistic representations in face recognition, whereas the findings of this study suggest that two complementary processes of roughly equal importance underlie recognition performance. There is a need for further research that is able to examine and compare the efficacy of these two conceptualisations of face recognition within the same experimental design.

Most studies and theories of face recognition have focused on the processing involved in analysing the whole face. Less attention has been paid to how component features themselves are perceived, represented, and recognised. One possibility is that similar information is extracted at 'local' and 'global' levels, namely information about critical features and their configuration. Just as a face may be parsed as a combination of information about features such as the eyes, nose, and mouth, together with their spatial configuration, it is likely that a feature such as an eye is parsed as a combination of information about features such as sclera, iris, pupil, and eyelids, together with their spatial configuration. As Tanaka and Farah (1993) point out, "the shapes of individual parts are essentially within-part spatial relations" (page 242). Further research is needed to address directly the similarities and differences of the processing of faces and their features at local and global levels. Whilst a number of studies have compared the effects of inversion on the recognition of whole faces and of individual features, there is a dearth of studies of the role of spatial configuration at a local level.

Two further related questions are whether feature-level and configuration-level processes are organised hierarchically, and whether they operate independently or interactively. Given that the present results demonstrate that information can be extracted independently at each level, it is unlikely that the two are organised hierarchically in normal face perception. Note that this is not necessarily inconsistent with Rhodes et al's (1993) hierarchical account, according to which the basic (first-order) configuration of the stimulus is extracted to identify the face as a face, followed by a second stage of norm-based coding. At this stage, featural and spatial relation information are hypothesised to be extracted and coded with reference to some general norm.

The present results provide evidence that a theoretically useful distinction may be made between the processing of featural information and the processing of configurational information in faces. The results also show that each type of information can be used independently of the other to recognise the identity of a face. This does not, of course, address whether feature and configuration are processed independently in the course of normal face recognition. Evidence for interactive processing has been provided by Sergent (1984) and by Tanaka and Sengco (1997). Further research on this question might address to what extent featural encoding (eg scrambled faces) is able to prime retrieval of spatial relation information (eg blurred faces).

The second aim of this study was to compare the processing of familiar and unfamiliar faces. As outlined earlier, models of face recognition differ in whether and how unfamiliar and familiar faces are processed differently. One possibility is that similar information is encoded for all faces, and that the familiarity of a face determines the strength of the memory connections that encode this information. In contrast, Bruce and Young (1986) postulated semi-independent pathways, relying on qualitatively different processing mechanisms for famous and unfamiliar faces. However, existing experimental and neuropsychological studies showing support for a qualitative distinction have typically compared familiar faces with ones that are truly unfamiliar (ie never seen). Furthermore, many of these studies have also confounded familiarity (famous versus never seen) with experimental task (recognition versus perceptual matching). Few if any studies have examined whether configurational and featural processing are differentially involved in the processing of familiar and unfamiliar faces. Diamond and Carey (1986) provide convincing evidence that the use of configurational information is more pronounced for stimulus classes with which an observer has expertise. Does expertise with individual items within a stimulus class similarly lead to such processing differences?

The results provide evidence for quantitative differences between the two sets of faces—famous faces are easier to recognise, and the detrimental effect of any of the manipulations relative to baseline levels was greater for the recognition of famous faces. However, no evidence was found for a qualitative difference in processing strategy. Featural and configurational information appear to have been used to the same extent in both tasks, as the relative extent to which accuracy was reduced and responses were slowed by blurring on the one hand, and by scrambling and inversion on the other, did not differ between the two tasks. For example, blurring (an impairment of featural processing) and scrambling (an impairment of configurational processing) each reduced recognition accuracy to 68% for famous faces, and to 64% for unfamiliar faces. The RT data also demonstrate no evidence for any qualitative differences in the processing of configurational and featural information for faces differing in familiarity. Whilst there were some quantitative differences (familiar faces were recognised somewhat faster), the effects of the various manipulations on response times (relative to one another) were closely similar in the two tasks.

These findings have a number of implications. First, they provide further evidence for Diamond and Carey's view that it is familiarity with a stimulus class rather than with individual items that is crucial in altering the balance between featural and configurational processing. Second, Bruce and Young's (1986) model of face recognition makes a qualitative distinction between the processing of familiar and unfamiliar faces. According to this model the usual route for the identification of (familiar) faces is through successfully matching structural codes built up for the presented images with existing representations held in the face-recognition units. Unfamiliar faces (as in a standard unfamiliar-face recognition paradigm) are attended to by strategically directed visual processes, and subsequent recognition is achieved by retrieving the representation built up in this way from a temporary episodic memory store. The current study, however, finds no evidence for a qualitative difference between processing unfamiliar and familiar faces.

6 Conclusions

The results of this experiment suggest that blurring on the one hand, and scrambling and inversion on the other, affect two rather different types of information used by the face-recognition system. It is argued that an important distinction exists between the processing of configurational information and featural information in face recognition.

Comparison of the results for famous and unfamiliar face-recognition tasks suggests that these findings are likely to be generalisable across the spectrum of familiarity (from once-seen to celebrity). No evidence was found in this study for a dissociation between the processing of familiar and unfamiliar faces, as the relative importance of configurational and featural routes of processing did not differ between the two recognition tasks.

Acknowledgements. We would like to thank Simon Wallace for his comments on an earlier draft of this paper.

References

Bartlett J C, Searcy J, 1993 "Inversion and configuration of faces" *Cognitive Psychology* **25** 281 – 316

Benton A L, 1980 "The neuropsychology of facial recognition" *American Psychologist* **35** 176 – 186

Bruce V, 1986 "Influences of familiarity on the processing of faces" *Perception* **15** 387 – 397

Bruce V, Doyle T, Dench N, Burton M, 1991 "Remembering facial configurations" *Cognition* **38** 109 – 144

Bruce V, Young A, 1986 "Understanding face recognition" *British Journal of Psychology* **77** 305 – 327

Bruce V, Young A, 1998 *In the Eye of the Beholder. The Science of Face Perception* (Oxford: Oxford University Press)

Bruyer R, Coget M C, 1987 "Features of laterally displayed faces: saliency or top – down processing?" *Acta Psychologica* **66** 103 – 114

Campbell R, Walker J, Benson P J, Wallace S, Coleman M, Michelotti J, Baron-Cohen S, 1999 "When does the inner-face advantage in familiar face recognition arise—and why?" *Visual Cognition* **6** 196 – 216

Costen N P, Parker D M, Craw I, 1994 "Spatial content and spatial quantisation effects in face recognition" *Perception* **23** 129 – 146

Diamond R, Carey S, 1986 "Why faces are and are not special: an effect of expertise" *Journal of Experimental Psychology: General* **115** 107 – 117

Ellis H D, Shepherd J W, Davies G M, 1979 "Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition" *Perception* **8** 431 – 439

Endo M, 1986 "Perception of upside-down faces: an analysis from the viewpoint of cue saliency", in *Aspects of Face Processing* Eds H D Ellis, M A Jeeves, F Newcombe, A Young (Dordrecht: Martinus Nijhoff)

Enns J T, Shore D I, 1997 "Separate influences of orientation and lighting in the inverted-face effect" *Perception & Psychophysics* **59** 23 – 31

Farah M J, Wilson K D, Drain M, Tanaka J N, 1998 "What is 'special' about face perception?" *Psychological Review* **105** 482 – 498

Harmon L D, 1973 "The recognition of faces" *Scientific American* **229** 70 – 83

Hayes A, 1988 "Identification of two-tone images: some implications for high- and low-spatial-frequency processes in human vision" *Perception* **17** 429 – 436

Hole G J, 1994 "Configurational factors in the perception of unfamiliar faces" *Perception* **23** 65 – 74

Johnston R A, Barry C, Williams C, 1996 "Incomplete faces don't show the whole picture: repetition priming from jumbled faces" *Quarterly Journal of Experimental Psychology, Section A—Human Experimental Psychology* **49** 596 – 615

Johnston A, Hill H, Carman N, 1992 "Recognising faces: effects of lighting direction, inversion, and brightness reversal" *Perception* **21** 365 – 375

Kemp R, McManus C, Piggot T, 1990 "Sensitivity to the displacement of facial features in negative and inverted images" *Perception* **19** 531 – 543

Lewis M B, Johnston R A, 1997 "The Thatcher illusion as a test of configural disruption" *Perception* **26** 225 – 227

Malone D R, Morris H H, Kay M C, Levin H S, 1982 "Prosopagnosia: a double dissociation between the recognition of familiar and unfamiliar faces" *Journal of Neurology, Neurosurgery, and Psychiatry* **45** 820 – 822

Miller J, 1996 "The sampling distribution of d'" *Perception & Psychophysics* **58** 65 – 72

Rhodes G, Brake S, Atkinson A P, 1993 "What's lost in inverted faces?" *Cognition* **47** 25 – 57

Rock I, 1973 *Orientation and Form* (New York: Academic Press)

Scapinello K F, Yarmey A D, 1970 "The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli" *Psychonomic Science* **21** 329 – 331

Searcy J H, Bartlett J C, 1996 "Inversion and processing of component and spatial-relational information in faces" *Journal of Experimental Psychology: Human Perception and Performance* **22** 904 – 915

Sergent J, 1984 "An investigation into component and configural processes underlying face perception" *British Journal of Psychology* **75** 221 – 242

Sergent J, 1986 "Microgenesis of face perception", in *Aspects of Face Processing* Eds H D Ellis, M A Jeeves, F Newcombe, A Young (Dordrecht: Martinus Nijhoff)

Suzuki S, Cavanagh P, 1995 "Facial organization blocks access to low-level features: an object inferiority effect" *Journal of Experimental Psychology: Human Perception and Performance* **21** 901 – 913

Tanaka J W, 1999, personal communication

Tanaka J W, Farah M J, 1993 "Parts and wholes in face recognition" *The Quarterly Journal of Experimental Psychology A* **46** 225 – 245

Tanaka J W, Sengco J A, 1997 "Features and their configuration in face recognition" *Memory and Cognition* **25** 583 – 592

Ulrich R, Miller J, 1994 "Effects of truncation on reaction time analysis" *Journal of Experimental Psychology: General* **123** 34 – 80

Valentine T, 1988 "Upside-down faces: A review of the effect of inversion upon face recognition" *British Journal of Psychology* **79** 471 – 491

Yarmey A D, 1971 "Recognition memory for familiar 'public' faces: Effects of orientation and delay" *Psychonomic Science* **24** 286 – 288

Yin R K, 1969 "Looking at upside-down faces" *Journal of Experimental Psychology* **81** 141 – 145

Yin R K, 1970 "Face recognition: a dissociable ability?" *Neuropsychologia* **23** 395 – 402

Young A W, 1984 "Right cerebral hemisphere superiority for recognising the internal and external features of famous faces" *British Journal of Psychology* **75** 161 – 169

Young A W, Hellawell D, Hay D C, 1987 "Configurational information in face perception" *Perception* **16** 747 – 759