Perception, 2000, volume 29, pages 159 – 170

DOI:10.1068/p3012

The face-inversion effect as a deficit in the encoding of configural information: Direct evidence

Alejo Freire, Kang Lee¶, Lawrence A Symons
Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada;
e-mail: kang@psyc.queensu.ca
Received 27 September 1998, in revised form 12 July 1999

Abstract. We report four experiments leading to conclusions that: (i) the face-inversion effect is mainly due to the deficits in processing of configural information from inverted faces; and (ii) this effect occurs primarily at the encoding stage of face processing, rather than at the storage stage. In experiment 1, participants discriminated upright faces differing primarily in configuration with 81% accuracy. Participants viewing the same faces presented upside down scored only 55%. In experiment 2, the corresponding discrimination rates for faces differing mainly in featural information were 91% (upright) and 90% (inverted). In experiments 3 and 4, the same faces were used in a memory paradigm. In experiment 3, a delayed matching-to-sample task was used, in which upright-face pairs differed either in configuration or features. Recognition rates were comparable to those for the corresponding upright faces in the discrimination tasks in experiments 1 and 2. However, there was no effect of delay (1 s, 5 s, or 10 s). In experiment 4, we repeated experiment 3, this time with inverted faces. Results were comparable to those of inverted conditions in experiments 1 and 2, and again there was no effect of delay. Together these results suggest that an 'encoding bottleneck' for configural information may be responsible for the face-inversion effect in particular, and memory for faces in general.

1 Introduction

Research on face processing has addressed a number of questions concerning face perception and recognition. A central issue has been the distinction between two different types of information available in a face and their respective roles in face processing. Featural information pertains to face elements which can be referred to in relative isolation, such as the size and shape of the eyes, nose, and mouth; configural information relates broadly to spatial relationships among these major elements within a face. Although there is considerable debate about the nature of what is 'configural information', as opposed to 'featural information' (Carey and Diamond 1977), for the purposes of this paper we use the term 'configural information' narrowly to refer to information about distances between the eyes, nose, and mouth—information referred to by Searcy and Bartlett (1996) as spatial-relational.

It has been proposed that face configuration and features are processed differently. This proposal is largely based on research on face-inversion effects. Many studies have shown that inverting a face stimulus has two important outcomes: the face becomes more difficult to recognise (eg Yin 1969; see Valentine 1988, and Searcy and Bartlett 1996, for reviews) and distortion in the face is less perceptible (eg Bartlett and Searcy 1993; Thompson 1980). The face-inversion effect is thought to result mainly from a disruption to processing of configural information that is sensitive to face orientation. By contrast, inversion does not affect processing of featural information because the latter is thought to be processed similarly regardless of orientation (eg Bartlett and Searcy 1993; Farah et al 1995; Tanaka and Sengco 1997).

Although there is converging evidence to support this contention, direct empirical evidence is limited for several reasons. First is the fact that the majority of face-processing studies have used photographs of different individuals in the stimulus set; these faces

¶ Author to whom correspondence and requests for reprints should be addressed.

differed simultaneously in configural and featural information (eg Carey and Diamond 1977; Diamond and Carey 1986; Goldstein and Chance 1980; Yin 1969). Second, studies that have not confounded configural and featural manipulations have sometimes used somewhat unnatural face stimuli. One example is faces that are lacking shading information (Kemp and McManus 1990; Sergent 1984), but these may not be processed in a manner similar to gray-scale photographs (see Bruce et al 1991). Some have used gray-scale photographs, but have created stimuli in which facial components are exchanged (eg the nose is moved to where the mouth normally belongs), again resulting in unnatural-looking face stimuli (eg Baenninger 1994). Bartlett and Searcy (1993; Searcy and Bartlett 1996) used spatially distorted gray-scale photographs of face stimuli. Participants found these altered faces relatively easy to discriminate, which suggests that configural information alone may be sufficient to encode faces. However, this deduction is premature given the very dramatic manipulations of configural information in Bartlett and Searcy's studies: original photographs were altered such that some resultant faces appeared abnormal and grotesque. Participants may have therefore relied not only on configural information to make discriminations between face pairs, but also on aesthetic or affective cues (see Searcy and Bartlett 1996, for discussion).

There are studies that have both avoided the confound of simultaneously altering configurations and features, and used faces that appear normal (eg Haig 1984; Rhodes et al 1993, experiment 3). However, these studies have primarily focused on recognition memory. For example, participants in the Rhodes et al (1993) study viewed a set of faces, and subsequently identified these faces when they were paired with new faces differing primarily in configurations. The target and distractor had either a very similar configuration ('hard' pairs), or a vastly different one ('easy' pairs). Faces could also be seen right side up or upside down. An inversion effect was obtained for both the easy and hard pairs, with recognition for inverted hard faces being reduced to near-chance levels (55%). These findings are consistent with the general notion that configural information and featural information are processed differently.

However, research to date has for the most part failed to ascertain at which stage the purported differences in processing of featural and configural information occur. It can be suggested that there are at least two possible points in the processing of face information where these differences might arise. One point is the perceptual encoding of the face [this is referred to as the 'structural encoding' phase in the model proposed by Bruce and Young (1986)], while the second is the storage of face representations. Most of the previously mentioned studies concerned the nature of the storage of face information. Differences found to exist at this storage stage might be the result of limitations already present in the 'structural encoding' stage. It can be suggested that front-end encoding mechanisms may create an 'encoding bottleneck' for configural information. This would lead to the differences found between configural and featural face-information processing in face recognition studies. In particular, with regard to the face-inversion effect, the encoding-bottleneck hypothesis is consistent with previous suggestions that inversion disrupts configural processing. However, this hypothesis further suggests that this effect occurs at a perceptual stage, rather than at a storage stage. Thus, the face-inversion effect is really a perceptual phenomenon, rather than a memory phenomenon [as suggested by the model proposed by Valentine (1991), for instance].

In the present study we specifically examined this encoding-bottleneck hypothesis. In experiment 1, a computer graphics program was used to manipulate a photograph of a face such that the configuration of the face was altered by varying slightly the location of the eyes and mouth. The resulting face stimuli thus differed primarily in configural information. In experiment 2, the graphics program was again used to manipulate the same original face photograph in a different manner. This time, main facial features (eyes, nose, and mouth) of other faces replaced those of the original face, while similar spatial relationships were maintained between features. The resulting face images thus differed mainly in facial features. All manipulations were subtle, resulting in face stimuli that resemble normal black-and-white portraits. Participants in experiments 1 and 2 were asked to discriminate between face images in upright or inverted displays. In experiment 3, participants were asked to recognise configurally or featurally altered upright face stimuli after delays of 1 s, 5 s, and 10 s. In experiment 4, the procedure of experiment 3 was repeated, this time with inverted displays.

If the encoding-bottleneck hypothesis is correct, and it is true that inversion selectively disrupts configural-information processing, the following patterns of results are expected:

(i) The documented inversion effect should be replicated in the discrimination task of experiments 1 and 2 with a specific pattern of results. That is, discrimination accuracy for the configurally altered inverted faces should be at near chance levels, while discrimination accuracy for the featurally altered inverted faces should be substantially above chance. Further, discrimination accuracy for featurally altered face images in experiment 2 should be similar in both orientations.

(ii) Patterns of memory retention (and therefore decay) over time should be similar for configurally and featurally altered upright faces (experiment 3). Also, memory for the inverted featurally altered faces (experiment 4) should be similar to those for upright faces in experiment 3. For the inverted configurally altered faces in experiment 4, because discrimination of the face stimuli is expected to be already at near-chance levels owing to the encoding bottleneck, recognition should remain extremely poor.

2 Experiment 1

2.1 *Method*

2.1.1 *Participants.* Eight participants completed a rating task. Sixteen additional observers participated in a discrimination task. Participants were students in an introductory-level psychology course and received course credit. Equal numbers of men and women completed each task. All participants had normal or corrected-to-normal vision.

2.1.2 *Materials.* The stimulus set consisted of eight 8.8 cm × 11.5 cm face images, each a modification of the same original photograph. The original portrait was a gray-scale photograph of a white adult male and was not included in the stimulus set. This photograph was altered with the Adobe Photoshop graphics program. Eyes were moved slightly up, down, in, or out relative to the original image, and the mouth was moved up or down: vertical movement of eyes and mouth was 3 pixels either up or down from the original image; horizontal movement of the eyes was 2 pixels for each eye, either closer together or farther apart than the original image. The eight permutations of four eye positions and two mouth positions made up the stimulus set (see figure 1).

2.2 *Procedure*

2.2.1 *Rating task.* A rating task was completed to evaluate face images in the stimulus set in terms of grotesqueness, distinctiveness, and familiarity. The task was included to ensure that face stimuli were not extreme in any of these dimensions. Booklets consisting of eight 8.8 cm × 11.5 cm gray-scale images of the different faces in the stimulus set, in random order, were presented to participants. Three questions assessed participants' perceptions of each face: grotesqueness (how grotesque is this face?: 1—not grotesque, 7—very grotesque; Searcy and Bartlett 1996); distinctiveness (how difficult would it be to pick out this face in a busy shopping mall?: 1—very easy, 7—very difficult; Valentine 1991); familiarity (how familiar does this face appear to you?: 1—not familiar, 7—very familiar). Participants were tested individually and no time restrictions were indicated.

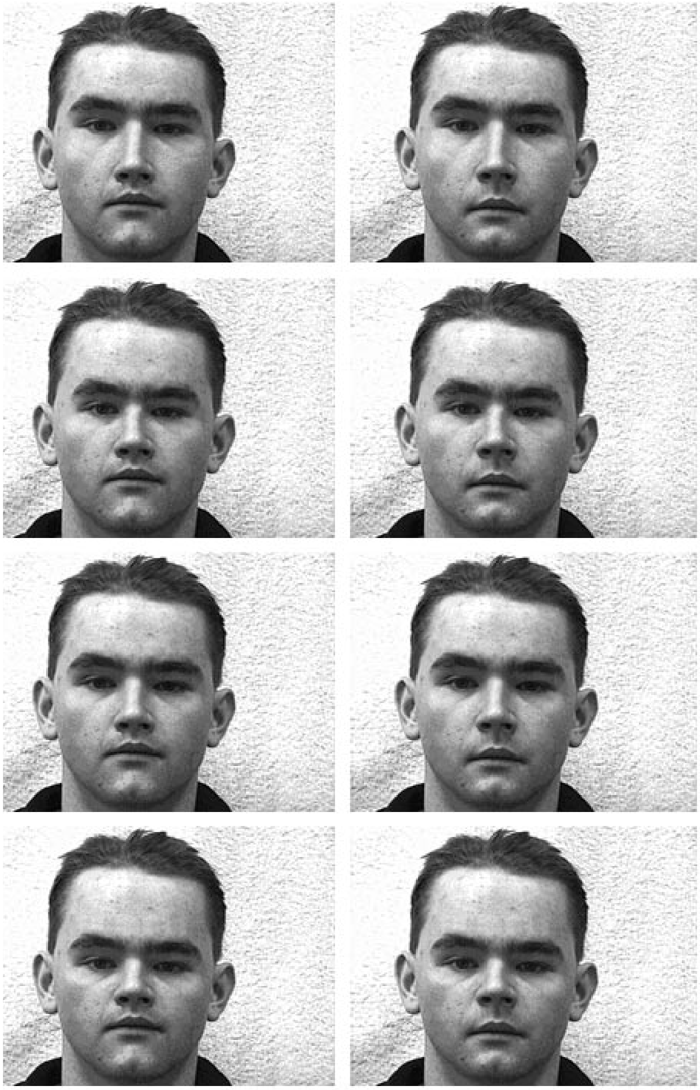

Figure 1. Stimulus faces for experiment 1.

2.2.2 *Discrimination task.* A Macintosh LCIII computer was used to present stimulus displays on a high-resolution 13-inch Apple monitor. Displays consisted of pairs of face photographs from the stimulus set in figure 1, presented simultaneously and side by side. Individuals were seated approximately 95 cm away from the monitor. Displays measured approximately 12.6 deg at their highest point and 6.9 deg at the widest point. The space-averaged luminance of the faces was 76.4 cd m^{-2}.

Participants were tested individually. Each was shown 112 pairs of face images on the computer monitor. Each face in the stimulus set was paired with itself 7 times (total of 56 trials) and with each of the other seven faces twice (once each on left and right sides, for a total of 56 trials). Thus, half the displays consisted of identical faces and half were composed of faces differing in configuration. Vscope software (Enns et al 1990) was used to present displays randomly, and to record responses. Half the observers viewed upright displays; the other half viewed the same displays presented in an inverted orientation. Participants were randomly assigned to a condition.

Observers were instructed to indicate that the two presented faces were photos of different persons by pressing one keyboard key, and that they were the same by pressing another key. They were told to be fast but accurate, and were informed that the faces would be quite similar in appearance. Keys used to indicate same and different pairs were reversed for half the participants in each experimental condition.

2.3 *Results*

2.3.1 *Rating task.* Table 1 shows participants' ratings of stimulus faces on dimensions of grotesqueness, distinctiveness, and familiarity. Table 1 also shows values of two-tailed t-tests conducted on each face, comparing each to the middle score of 4 on each of these dimensions (all subsequent reported t-tests were also two-tailed). Four of the eight faces were judged to be average in terms of grotesqueness; the remaining four faces were rated as less grotesque than average. None of the faces in the stimulus set was judged to be distinctive or familiar.

Table 1. Mean grotesqueness, distinctiveness, and familiarity ratings for the face stimuli used in experiment 1. Also presented are t-tests comparing each face stimulus to a middle score of 4 on each rated dimension.

Face	Grotesqueness		Distinctiveness		Familiarity	
	mean (SD)	t	mean (SD)	t	mean (SD)	t
1	2.63 (1.60)	−2.43*	4.13 (1.89)	0.19	4.38 (1.69)	0.63
2	2.63 (1.60)	−2.99*	3.75 (1.75)	−0.40	4.00 (1.60)	0.00
3	2.75 (1.75)	−2.02	4.00 (1.69)	0.00	3.88 (2.17)	−0.16
4	2.88 (1.55)	−2.05	4.25 (1.67)	0.42	3.25 (2.19)	−0.97
5	2.50 (1.31)	−3.24*	3.88 (2.03)	−0.17	3.88 (1.96)	−0.18
6	3.13 (1.36)	−1.83	3.88 (2.03)	−0.17	3.13 (2.03)	−1.22
7	2.88 (1.64)	−1.94	4.38 (2.00)	0.53	3.75 (2.05)	−0.34
8	2.63 (1.60)	−2.43*	4.13 (2.17)	0.16	4.00 (1.93)	0.00

* $p < 0.05$.

2.3.2 *Discrimination task.* Figure 2 shows the means and standard errors of accuracy scores for discrimination in upright and inverted conditions. Participants viewing upright-face pairs were significantly more accurate than were those viewing the same face pairs presented upside down: $t_{14} = 6.66$, $p < 0.001$. Inversion of displays reduced percentage correct discrimination from 81% to 55%. A one sample t-test comparing the accuracy in the inverted condition to chance performance (50%) showed that correct discrimination was at a higher than chance level: $t_7 = 3.67$, $p < 0.01$.

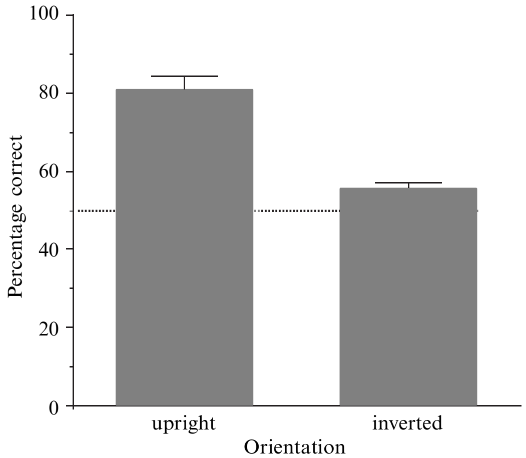

Figure 2. Percentage of correct discrimination of face pairs in 'upright' and 'inverted' groups for experiment 1.

Although none of the faces in the stimulus set was judged as particularly grotesque, distinct, or familiar, it is still possible that some face images were easier to discriminate than were others. To investigate this possibility, the percentage of correct scores was calculated for each face, that is accuracy scores only on trials when each face was part of the test pair. The percentage of correct scores on upright trials including faces 1 through 8 was: 82.14%, 78.57%, 80.95%, 79.76%, 81.55%, 79.17%, 80.36%, and 80.95%. Corresponding scores for faces 1 through 8 in the inverted orientation were: 51.19%, 55.36%, 53.57%, 54.17%, 55.36%, 55.36%, 59.52%, and 54.17%. Two separate one-way repeated-measures ANOVAs were carried out on these percentages for the upright and inverted faces, respectively. These analyses showed no significant difference in discrimination rate between the eight faces, in either upright or inverted orientations (upright: $F_{7,49} = 0.37$, ns; inverted: $F_{7,49} = 0.41$, ns).

2.4 *Discussion*

Experiment 1 provides evidence that inverting a face disrupts perceptual processing of configural information in a discrimination task. Participants were able to discriminate upright faces differing primarily in configural information, but, when the same face pairs were inverted, discrimination was extremely poor.

The strength of our manipulations is exemplified by the fact that, in recognition studies, face inversion reduces accuracy to about 70% (eg Carey and Diamond 1977; Diamond and Carey 1986), whereas, in the present study, the discrimination of upside-down faces was only 55%. In contrast, the upright-face discrimination rate was comparable to recognition rates typically reported for upright faces (eg Yin 1969). The discrepancy between our data and previous reports is likely the result of the fact that previous recognition studies typically used stimulus sets comprised of photographs of different people, in which configuration and features varied simultaneously. Results of the present task indicate that if this confound is reduced, such that it is primarily configural information that differentiates faces, face processing is more drastically reduced by stimulus inversion. Indeed, the 55% discrimination rate for inverted faces is almost identical to that obtained by Rhodes et al (1993, experiment 3) for memory of inverted faces differing subtly in configurations, which suggests that their results may be completely attributable to deficits in encoding of configural cues.

This suggestion is consistent with that of Searcy and Bartlett (1996), who reported an inversion effect in a similar discrimination task in which faces were distinguished by their configurations. However, Searcy and Bartlett's subjects were accurate in 70% of inverted trials, a much higher rate than that obtained in experiment 1. The most notable difference in the two studies is the nature of the stimuli used. Configural manipulations in the present study were much subtler than those in Searcy and Bartlett's work: they intentionally created stimulus faces with a grotesque appearance. Results of the rating task included in experiment 1, however, indicate clearly that, when configural variations are restricted to a normal range, discrimination of faces can be reduced to near-chance levels. Our results therefore provide strong evidence that the detrimental effects of inversion on face processing likely result from the disruption of *encoding* of configural information when a face is inverted.

3 Experiment 2

Experiment 1 demonstrates that a dramatic inversion effect can be found in a discrimination task involving faces that differ mainly in configural information. However, claims that the face-inversion effect is largely due to a disruption of configural encoding also require that a similar inversion effect *not* be found in a task requiring featural discrimination. In the second experiment we examined the effect of inversion on the ability to process faces that differ primarily in featural information. Faces were constructed with similar configural information (the configuration was that of the original face photograph from experiment 1), but with features from other faces replacing the original features.

3.1 *Method*

3.1.1 *Participants.* Eight participants completed a rating task and sixteen completed a discrimination task. All were students in an introductory-level psychology course and received course credit for their involvement. Equal numbers of men and women completed each task. All participants had normal or corrected-to-normal vision.

3.1.2 *Materials.* The stimulus set consisted of eight face photographs. The original photograph from which the stimuli in experiment 1 were derived served as one stimulus. Seven new face images were created with a graphics program, by pasting the eyes, nose, and mouth of photographs of seven different white adult males over the corresponding features of the original face (figure 3), thereby maintaining a similar configuration. Eyes from the additional face photographs were resized such that the diameter of the iris was equivalent to that of the original photograph. Similarly, the widths of the nose

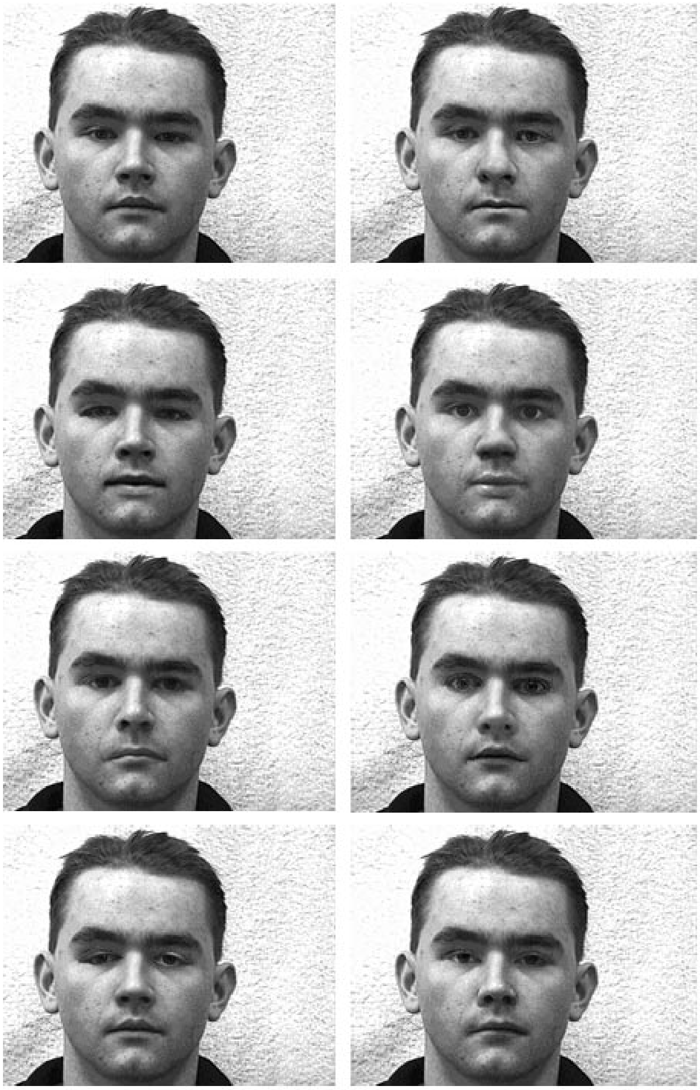

Figure 3. Stimulus faces for experiment 2.

and mouth were equalised prior to feature replacement. Features were then replaced so that the location of the irises remained constant, nostrils were at the same height, and the line at the boundary of the upper and lower lip appeared in the same location.

3.2 *Procedure*
Rating and discrimination tasks equivalent to those of experiment 1 were completed with experiment 2 stimuli.

3.3 *Results*
3.3.1 *Rating task.* Table 2 shows participants' ratings of stimulus faces in experiment 2 on dimensions of grotesqueness, distinctiveness, and familiarity. Table 2 also shows values of two-tailed *t*-tests conducted on scores reported for each face, comparing them to the middle score of 4 on each dimension. Six of the eight faces were judged to be average in terms of grotesqueness; the remaining two faces were rated as less grotesque than average. Two faces were judged as below average in distinctiveness and none was judged familiar.

Table 2. Mean grotesqueness, distinctiveness, and familiarity ratings for the face stimuli used in experiment 2. Also presented are *t*-tests comparing each face stimulus to a middle score of 4 on each rated dimension.

Face	Grotesqueness		Distinctiveness		Familiarity	
	mean (SD)	*t*	mean (SD)	*t*	mean (SD)	*t*
1	3.38 (1.92)	−0.92	4.00 (1.51)	0.00	3.75 (1.49)	−0.48
2	4.63 (1.06)	1.67	4.13 (1.73)	0.21	3.25 (1.67)	−1.27
3	4.13 (1.55)	0.23	3.50 (1.85)	−0.76	3.13 (1.73)	−1.43
4	3.38 (1.85)	−0.96	4.75 (1.04)	2.05	3.13 (1.81)	−1.37
5	2.75 (1.39)	−2.55*	5.25 (1.17)	3.04*	3.88 (1.55)	−0.23
6	2.75 (1.28)	−2.76*	5.25 (0.89)	3.99*	4.00 (1.77)	0.00
7	4.50 (1.69)	0.84	3.38 (0.92)	−1.93	2.63 (1.77)	−2.20
8	3.50 (1.60)	−0.88	3.25 (1.98)	−1.07	3.13 (1.955)	−1.59

$^* p < 0.05$.

To examine whether there were overall differences in grotesqueness, distinctiveness, or familiarity ratings between the faces used in experiments 1 and 2, three separate one-way ANOVAs were conducted on the grotesqueness, distinctiveness, and familiarity ratings. These analyses revealed no significant differences on these dimensions between faces in experiments 1 and 2: grotesqueness: $F_{1,14} = 1.85$, ns; distinctiveness: $F_{1,14} = 0.04$, ns; familiarity: $F_{1,14} = 0.29$, ns, respectively.

3.3.2 *Discrimination task.* Figure 4 shows means and standard errors of accuracy scores for discrimination in upright and inverted conditions. Percentage correct discrimination did not differ between observers viewing upright displays and those viewing inverted ones: $t_{14} = 0.23$, ns.

As with faces in experiment 1, the percentage of correct scores was calculated for individual faces in experiment 2. The values for upright faces were: 89.88%, 89.88%, 89.28%, 91.67%, 92.26%, 91.67%, 90.48%, and 92.86%; corresponding values for inverted faces were 86.90%, 89.29%, 88.10%, 87.50%, 88.69%, 87.50%, 89.88%, 87.50%. Two separate one-way repeated-measures ANOVAs on the percentage of correct scores for the faces again indicated no significant differences in discrimination rate between the eight faces, in either orientation of stimulus presentation (upright: $F_{7,49} = 0.50$, ns; inverted: $F_{7,49} = 0.35$, ns).

3.4 *Discussion*
Inverting face stimuli has no appreciable effect on discrimination ability when face pairs differ mainly in featural information. Observers performed comparably on the

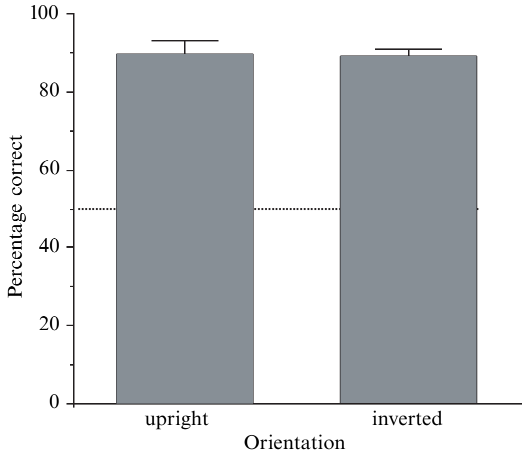

Figure 4. Percentage of correct discrimination of face pairs in 'upright' and 'inverted' groups for experiment 2.

accuracy measure for upright and inverted faces in experiment 2, in contrast to the difference obtained in experiment 1 when faces differed mainly in configuration. This result provides support for the contention that featural information of a face is less susceptible to inversion than configural information. Results from experiments 1 and 2 confirm the first critical prediction of the encoding-bottleneck hypothesis, that, when a face stimulus is inverted, the processing of configural information is severely and selectively disrupted during encoding.

4 Experiments 3 and 4

Experiments 3 and 4 were conducted to test the second critical prediction of the encoding-bottleneck hypothesis, that the face-inversion effect mainly reflects a 'front-end' encoding deficit of configural information from inverted faces. If this is the case, then a similar inversion effect should be observed in a memory task. That is, for inverted faces, memory of featural information should remain at a very high level and memory for configural information should be at near-chance levels. A related prediction of the encoding-bottleneck hypothesis is that, for upright faces, memory for both featural and configural information should be high. Further, the retention pattern of both types of information should be similar.

A delayed match-to-sample memory paradigm was used to test these predictions. Participants were first shown a face. After a 1 s, 5 s, or 10 s interval, the previously seen face was paired with a new featurally or configurally altered face. They were asked to determine which face was that previously seen. Hence, except for the memory component, this task was identical to the discrimination task of experiments 1 and 2. In experiment 3 all face stimuli were presented upright, while in experiment 4 displays were inverted.

4.1 *Method*

4.1.1 *Participants.* Sixteen introductory psychology students (ten females and six males) participated for course credit in experiment 3, and sixteen different students in experiment 4 (ten females and six males). All had normal or corrected-to-normal vision.

4.1.2 *Materials*

The design of the memory task required only nine face images, obtained in the following manner: First, three configurally altered faces used in experiment 1 were randomly selected. Second, three sets of features (eyes, nose, and mouth) used in experiment 2 were randomly selected. These features were then placed on the three configurally altered faces, resulting in nine stimulus faces. Within each configuration, the eyes, noses, and mouths of the three faces were matched as described for experiment 2.

For each stimulus face, there were thus two other faces that differed primarily in features and two that differed primarily in configuration. Monitor and software used were the same as for the previous experiments.

4.2 *Procedure*

Participants, tested individually, completed 108 trials divided into two blocks of 54. Participants had the option of taking a short break between blocks or continuing directly to the second block. For each trial, participants were first shown a target face (one of the nine faces from the stimulus set) for 5 s, followed by an interval during which the screen was black. Then, the target face was shown again side by side with another face. Participants' task was to indicate with the use of keyboard keys whether the target was that on the left or the right of each test pair. The target face appeared randomly on either the right or the left side of the screen.

The 108 trials were divided evenly into three sets that differed only in the interval between the initial display of the target face and the presentation of the test pair: 1 s, 5 s, and 10 s. In addition, each of these sets was divided evenly into two trial types (18 trials each). For the first type, the target face differed from the distractor primarily in configuration (configuration trials); for the second, the target and distractor faces differed primarily in features (feature trials).

Individuals were seated 95 cm away from the monitor. Initial presentation of the target was at the centre of the screen, the display measuring approximately 12.6 deg at its highest point and 3.5 deg at the widest point. The space-averaged luminance of the target faces was 76.4 cd m^{-2}. Characteristics of test pair displays were identical to those for face pairs in the discrimination tasks of experiments 1 and 2.

Participants pressed the ← key when indicating the target was on the left side of the screen and the → key to indicate the target was on the right side. Use of these keys was not counterbalanced between participants, given that it was deemed confusing to use directional keys to indicate an opposite direction (eg using the ← key to indicate the target is on the right). Identification of the target from a test pair terminated the trial. The next trial followed 3 s later.

The same face stimuli were used in experiments 3 and 4 except that faces were presented upright in experiment 3 and inverted in experiment 4.

4.3 *Results*

A score for number of correct identifications (out of 18) was calculated for each participant for each combination of face type (configuration or features) and interval (1 s, 5 s, or 10 s).

4.3.1 *Experiment 3.* A 2 (face type) × 3 (interval) repeated-measures ANOVA of percentage identification accuracy indicated a significant effect of face type: $F_{1,15} = 23.58$, $p < 0.01$. There was no main effect of interval: $F_{2,14} = 2.53$, ns; nor a significant interaction of face type and interval: $F_{2,14} = 1.58$, ns. Participants were more accurate on feature trials than on configuration trials (upright feature: mean = 15.42, SD = 1.97, or 86% correct recognition; upright configuration: mean = 13.25, SD = 2.61, or 74% correct recognition).

4.3.2 *Experiment 4.* A 2 (face type) × 3 (interval) repeated-measures ANOVA of percentage identification accuracy indicated a significant effect of face type: $F_{1,15} = 209.86$, $p < 0.01$. There was no main effect of interval: $F_{2,14} = 0.15$, ns, nor a significant interaction of face type and interval: $F_{2,14} = 1.54$, ns. Participants were more accurate on feature trials than on configuration trials (inverted feature: mean = 15.50, SD = 1.70, or 86%; inverted configuration: mean = 10.27, SD = 2.47, or 57%). Figure 5 shows percentage of correct identifications for experiments 3 and 4.

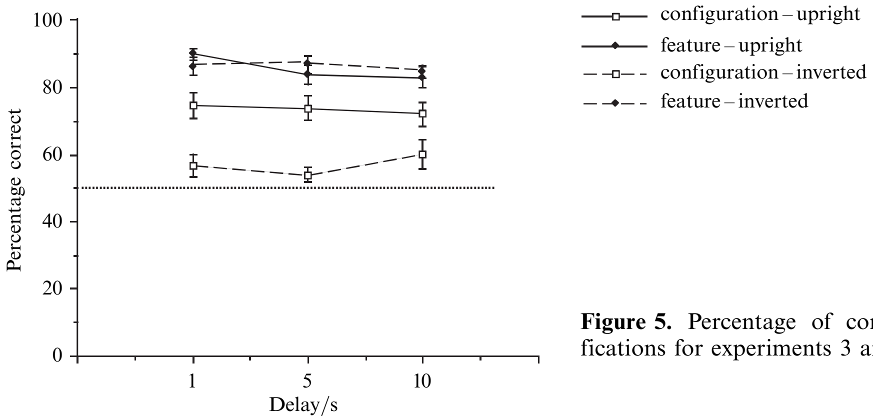

Figure 5. Percentage of correct identifications for experiments 3 and 4.

4.4 *Discussion*

Experiments 3 and 4 revealed that the expected advantage for feature recognition relative to recognition of configurations paralleled the similar advantage observed in the discrimination tasks of experiments 1 and 2. Accuracy measures in experiments 3 and 4 were qualitatively similar to those in experiments 1 and 2. In experiment 1, upright faces differing in configuration were discriminated with 81% accuracy; correct recognition rate for configurally altered upright faces in experiment 3, collapsed across delay intervals, was 74%. In experiment 2, participants were accurate on about 90% of upright face pairs distinguished by their features; in experiment 4, memory for featurally altered faces was 86% across all intervals. Results for inverted faces in discrimination and memory tasks was also highly similar: 89% and 86% for featurally altered faces in discrimination and memory tasks respectively; 55% and 57% for the corresponding configurally altered discrimination and memory trials. Combined, these data suggest that once configural and featural information pass the encoding stage, they remain largely unaffected by a short delay. This finding is consistent with the encoding-bottleneck hypothesis. Patterns of retention after longer intervals need to be addressed in future studies.

5 General discussion

The four experiments presented here provide direct evidence to suggest that inverting face stimuli disrupts the processing of configural information. Further, this disruption occurs at the encoding stage of face processing. Once the configural or featural information is encoded, retention of these two types of information remains similar, at least over short delays. These results provide support for the hypothesis that perceptual processing creates an 'encoding bottleneck' that limits the input of configural face information into memory.

Our findings suggest that models of face processing must account for the critical role of perceptual mechanisms. Some models of face processing suggest that inverted faces are more difficult to recognise because they are outliers in a distributed face-memory architecture (eg Valentine 1991). Models of this nature suggest that there is nothing special about the effects of inversion of faces, relative to any other visually perceived object. It has also been suggested that face inversion effects might be better explained as examples of mental rotation (eg Valentine and Bruce 1988). The present data would suggest that inversion effects are best explained by perceptual limitations specific to configural information in face stimuli. These perceptual limitations might play a very critical role in how a face is represented in memory. Further, mental rotation seems a limited explanation, given that inversion effects were so minimal for the faces distinguished largely by featural information in experiments 2 and 4.

The nature of an encoding bottleneck may be a useful explanation for a number of face-related phenomena. For instance, recently it has been demonstrated that the other-race effect (in which there is an impaired ability to recognise faces of a race other than one's own—see Bothwell et al 1989) can occur at a perceptual level. O'Toole et al (1996) showed that the ability to discriminate face gender was impaired for other-race faces. Symons et al (1997) found similar results in a task in which participants had to find happy-face targets amongst sad-face distractors. These findings, paired with those of the present study, suggest that an encoding bottleneck may be a powerful explanation for a variety of established phenomena in face processing. This hypothesis is consistent with the model suggested by Bruce and Young (1986) and emphasises the importance of their 'structural encoding' stage in the mechanisms of face processing.

Acknowledgements. This research was supported by a Natural Sciences and Engineering Research Council of Canada grant to Kang Lee. The authors would like to thank Nikolaus Troje for his suggestions concerning experiment 2. We also thank Nikolaus Troje, Darwin Muir, and three anonymous reviewers for their constructive comments on earlier versions of the manuscript.

References

Baenninger M, 1994 "The development of face recognition: Featural or configurational processing?" *Journal of Experimental Child Psychology* **57** 377 – 396

Bartlett J C, Searcy J, 1993 "Inversion and configuration of faces" *Cognitive Psychology* **19** 473 – 497

Bothwell R K, Brigham J C, Malpass R S, 1989 "Cross-racial identification" *Personality and Social Psychology Bulletin* **15** 19 – 25

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

Carey S, Diamond R, 1977 "From piecemeal to configurational representation of faces" *Science* **195** 312 – 314

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

Enns J, Ochs E, Rensink R, 1990 "Visual Search: Macintosh software for experiments in visual search" *Behavioral Research Methods in Instrumentation and Computers* **22** 118 – 122

Farah M J, Tanaka J W, Drain H M, 1995 "What causes the face inversion effect?" *Journal of Experimental Psychology: Human Perception and Performance* **21** 628 – 634

Goldstein A G, Chance J E, 1980 "Memory for faces and schema theory" *Journal of Psychology* **105** 47 – 59

Haig N D, 1984 "The effect of feature displacement on face recognition" *Perception* **13** 505 – 512

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

O'Toole A J, Patterson J, Deffenbacher K A, 1996 "An 'other-race effect' for categorizing faces by sex" *Perception* **25** 669 – 676

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

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

Symons L A, Clarke M P, Muir D W, 1997 "Inversion effects and cross-race face perception in a visual search task" *Investigative Ophthalmology & Visual Science* **38**(4) 4666

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

Thompson P, 1980 "Margaret Thatcher—A new illusion" *Perception* **9** 483 – 484

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

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

Valentine T, Bruce V, 1988 "Mental rotation of faces" *Memory & Cognition* **16** 556 – 566

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