Perceptual and Motor Skills, 2011, 113, 3, 859-874.

THE FIRST 100 MILLISECONDS OF A FACE: ON THE MICROGENESIS OF EARLY FACE PROCESSING[1, 2]

CLAUS-CHRISTIAN CARBON

Department of General Psychology and Methodology
University of Bamberg

Summary.—Face recognition involves both processing of information relating to features (e.g., eyes, nose, mouth, hair, i.e., featural processing), as well as the spatial relation between these features (configural processing). In a sequential matching task, participants had to decide whether two faces that differed in either featural or relational aspects were identical or different. In order to test for the microgenesis of face recognition (the development of processing onsets), presentation times of the backward-masked target face were varied (32, 42, 53, 63, 74, 84, or 94 msec.). To test for specific processing onsets and the processing of different facial areas, both featurally and relationally modified faces were manipulated in terms of changes to one facial area (eyes or nose or mouth), two, or three facial areas. For featural processing, an early onset for the eyes and mouth was at 32 msec. of presentation time, but a late onset for the nose was detected. For relationally differing faces, all onsets were delayed.

One major assumption in face research is that we recognize faces by processing the facial features (e.g., eyes, nose, mouth, hair) in addition to the relation between these features (Tanaka & Sengco, 1997). The literature encompasses an extensive series of articles on the differential influence of "featural" and "relational" (configural) facial information (Leder & Carbon, 2006) regarding face-specific effects, e.g., inversion (Yin, 1969; Carbon & Leder, 2005) or the composite effect (Young, Hellawell, & Hay, 1987). Notwithstanding these research efforts, a unique definition of "featural" versus "configural" is still missing despite many attempts to remedy this problem (e.g., Maurer, Le Grand, & Mondloch, 2002; Leder & Carbon, 2006). In the present paper, featural processing is defined as processing facial features, whereas configural processing is defined as the processing of the spatial relationship between those features, the so-called relational information, following Leder and Carbon's definition (2006). There is also a lack of deeper knowledge of the temporal aspects of processing such information in the literature, especially for limited presentation times of less than 100 msec., which are referred to as "early" face processing in this paper.

---

[1] Address correspondence to Claus-Christian Carbon, University of Bamberg, Markusplatz 3, D-96047 Bamberg, Germany or e-mail (ccc@experimental-psychology.com).

[2] The author would like to thank Thomas Ditye and Valentin Gattol for helping to conduct the experiments, Helmut Leder for providing help on pre-studies conducted for this experiment, and three anonymous reviewers who provided very helpful valuable information for improving the paper.

DOI 10.2466/07.17.22.PMS.113.6.859-874 ISSN 0031-5125

In order to understand the "microgenesis" or the time course of the first moments of face recognition, a presentation limitation paradigm with variable presentation time was employed. Microgenesis refers to the sequence of events which are assumed to occur in the period between the stimulus onset and the formation of a single, relatively stable cognitive response (percept or thought) to this stimulus (Flavell & Draguns, 1957). A modern framework for microgenetic research was mainly developed by Bachmann (2000), where the concept was defined, historically embedded, reflected, and integrated into possible research questions and domains of high relevance. Only recently, the microgenetic approach prompted a series of studies in the domain of aesthetic processing (Leder, Carbon, & Ripsas, 2006; Bachmann, 2007; Augustin, Leder, Hutzler, & Carbon, 2008), object recognition (Grill-Spector & Kanwisher, 2005), face recognition (Carbon & Leder, 2005; Veres-Injac, Hofer, & Schwaninger, 2007), and perception of real-world scenes (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2000; Bacon-Mace, Mace, Fabre-Thorpe, & Thorpe, 2005), making it a promising source of insight into the processing of visual stimuli.

In the present study, the processing onset of face recognition was investigated, employing a self-paced sequential matching task with backward masking at varying presentation times. The systematic variation of presentation times using featural versus relational faces enables testing specific onsets for featural versus configural processing, respectively. To test hypotheses on the processing of different facial areas, variations of the stimulus material were further employed by manipulating one, two, or three facial areas within one test face. The present paper aims to test the following hypotheses: (1) Within the first 100 msec. of when a face is presented, a clear microgenesis is detected, thus the performance of detecting changes to a face increases with prolonging the presentation time; (2) The processing of featural and relational facial information shows different performances and onsets in terms of presentation time needed to detect changes to a face; (3) The processing of different facial areas shows different performances and different onsets in terms of presentation time needed to detect changes to a face; and (4) The more changes to different facial areas, the better the rate of detecting them.

METHOD

In the following experiment, artificial photo-quality faces with changes to featural and relational information were constructed, either with low or highly distinctive characteristics. These manipulated faces were evaluated for distinctiveness and plausibility in a pre-study in order to ensure that faces of both information classes (featural versus relational) were equally distinctive and plausible. Participants had to detect differences between faces in a sequential matching paradigm with varying presentation times (32 msec.–94 msec. on seven levels of approx. 10-msec. difference).

Materials

A set of 22 female faces from the DADA Face Databank (consisting of true color photographs with neutral expressions), created by the author, was used in the present study (see details in Carbon, 2003). In Pre-study 1, 16 participants (12 women; *M* age = 23.8 yr.) individually rated the distinctiveness of the complete set of all 22 faces on an adjective-based scale, with anchors of 1: Very indistinctive and 7: Very distinctive, by pressing buttons on a keyboard. Afterwards, they all evaluated the individual eyes, noses, mouths, and hairstyles in sub-blocks, which they again had to rate for distinctiveness on the same scale [“How easily would the face (the facial feature) stand out in a crowd?”]. Facial features were split into two distinct pools of facial parts for every inner-facial component (eyes, nose, and mouth) based on distinctiveness data. The “low distinctiveness group” contained the lower half of all distinctiveness ratings cut by the median of all scores, the “high distinctiveness group” comprised the upper half. As the aim was to construct standardized artificial faces for the subsequent studies, facial features were selected from the different groups that fit the best alternative facial contexts in terms of skin colorization and quality. Two sets of faces were then generated. First, four basic faces were selected, which comprised the facial context and facial parts that were average in terms of distinctiveness. Based on these basic faces, all additional faces were constructed. Two classes of manipulations were performed: (a) manipulations by replacements of the components (“featural faces”), and (b) manipulations by changing the spatial relations between the cardinal components (“relational faces”). Furthermore, within both manipulation classes, a low and a highly distinctive version of each face were generated by varying the distinctiveness of these faces. For the featural faces this was achieved by replacing the original parts of the face with either a low or high distinctiveness member from the different pools of facial parts. For the relational faces, the components of mouth and nose were shifted downward, and the eyes were shifted inward, either in small (7 pixels) or larger steps (14 pixels). These manipulations were performed for all three facial areas plus the combinations of them, resulting in seven different versions. These were only for the eye region, the nose area, the mouth area, the eyes and the nose area, the nose, and the mouth area, the eyes and the mouth area and combined for the eyes, nose and mouth area.

For an illustration of typical manipulations of both classes, Fig. 1 shows an example of the basic faces (Face 2) and the respective highly distinctive featural and relational versions of it. For each of the faces this resulted in 2 (class: featural face versus relational face) × 2 (distinctive-

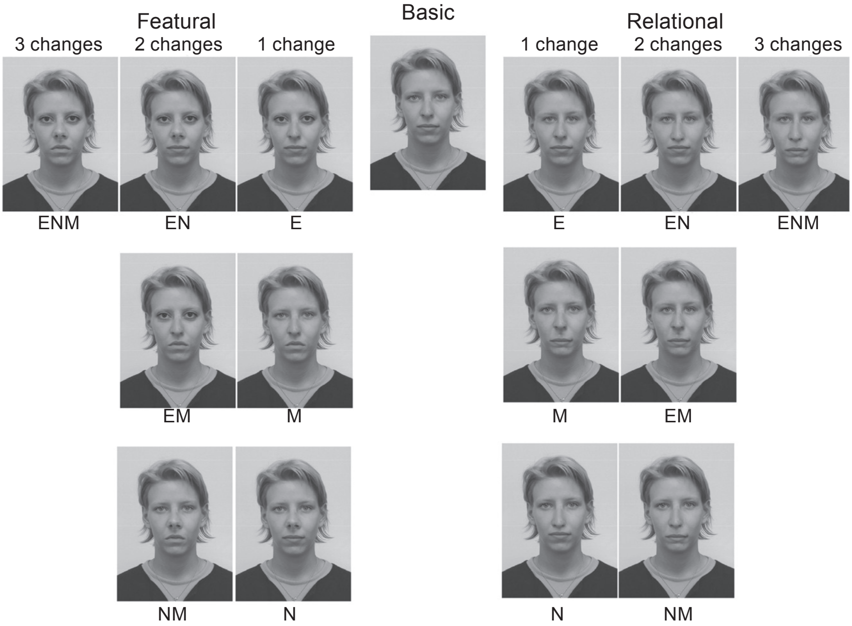

FIG. 1. Illustration of the (highly distinctive) manipulations to facial areas on one of the basic faces (Face 2). On the left side, featurally manipulated versions (“featural faces”) are shown, where cardinal features (Eyes = E, Nose = N, Mouth = M) have been changed. On the right side, relationally manipulated versions (“relational faces”) are shown, where second-order spatial relations between the cardinal features have been changed. For the 1-change manipulations, only one single facial area was changed; for the 2-change manipulations, a combination of two areas were changed (e.g., EM = Eyes and Mouth areas); and for the 3-change manipulations, all three areas were changed (ENM).

ness: low versus high) × 7 (area: E: Eyes, N: Nose, M: Mouth, EN, EM, NM, ENM) = 28 face versions plus the basic face.

Another pre-study (Pre-study 2) was conducted in order to ensure that these artificially constructed faces revealed comparable levels of distinctiveness. Sixteen undergraduate students (11 women; *M* age = 26.1 yr.) who were given course credit to fulfill course requirements took part in a rating task to evaluate the distinctiveness and the plausibility of all stimuli (see, for a detailed description, Carbon, 2003). The participants individually rated, in two successive sub-blocks, the distinctiveness and then the plausibility of each facial version on adjective scales with anchors of 1: Very indistinctive/implausible and 7: Very distinctive/plausible, by pressing buttons on a keyboard. To test for possible differences, two different three-way repeated measurement analyses of variance (ANOVAs; dependent measures: distinctiveness or plausibility) were conducted with class (featural faces versus relational faces), distinctiveness (low or highly distinctive), and area (E = Eyes, N = Nose, M = Mouth, EN = Eyes and Nose, EM = Eyes and Mouth, NM = Nose and Mouth, ENM = Eyes and Nose and Mouth) as within-subject factors. Class was not significant in any analysis, $F_{1,15}$s < 1.85, *p*s > .19, ns. Thus, no difference was observed between featurally and relationally manipulated items with respect to distinctiveness and plausibility.

Concerning the factor distinctiveness, both analyses indicate significant main effects. The faces categorized by highly distinctive ratings in Pre-study 2 were found to be more distinctive than faces of low distinctiveness ($F_{1,15} = 51.12$, $p < .0001$; $\eta^2_p = 0.77$). Moreover, high distinctiveness was associated with lower plausibility of the associated faces ($F_{1,15} = 14.90$, $p = .002$; $\eta^2_p = 0.50$). The combination of highly distinctive features might be more uncommon and thus reduces plausibility. Interestingly, this relationship between distinctiveness and plausibility had differential effects on the different face classes. As the interaction between class and distinctiveness for plausibility ratings revealed ($F_{1,15} = 15.58$, $p = .001$; $\eta^2_p = 0.51$), only the distinctiveness of relational faces influenced plausibility. Although natural faces differ in terms of featural as well as relational information, simultaneous alteration of several relational aspects seems to generate an uncommon appearance (see detailed craniofacial data in Hreczko & Farkas, 1994). Note, the present study did not try to match the featural and relational faces in terms of a physical method (e.g., equating the actual pixel changes), but in terms of a psychological method (perceived distinctiveness and plausibility). The size of stimuli was always 288 (horizontal) × 384 (vertical) pixels, with participants seated about 60 cm away from the screen resulting in a visual angle of approx. 3.4° × 4.6°.

Participants

Fifty-two participants (39 women; *M* age = 23.0 yr.) volunteered in the experiment. Participants were undergraduate students who were given course credit to fulfill course requirements. None of them had participated in one of the evaluative pre-studies to construct and evaluate the material used for this experiment.

Procedure

Participants sat approximately 60 cm away from the monitor. The experiment was controlled using PsyScope PPC 1.25 (Cohen, MacWhinney, Flatt, & Provost, 1993). A CMU ButtonBox was used for recording the participants' responses, allowing measurement of reaction times with a resolution of less than 1 msec. (Cohen, *et al.*, 1993).

In each trial, participants were shown two facial pictures in sequence (first: basic face; second: target face), separated by a black-and-white random dot mask (as in Carbon & Leder, 2005) presented for 200 msec. They were instructed to decide whether the target and the basic face were identical or different; the following dependent measure was the rate of correct matches. Answers were assigned to the keys in a counter-balanced manner across all participants. The participants had to give answers as fast and accurately as possible. Presentation time of the basic face was held constant across all conditions at 1,500 msec.; the presentation time of the target varied in seven steps (32, 42, 53, 63, 74, 84, and 94 msec.). After the presentation of the target, a second black-and-white random dot mask, physically the same as the first one, was presented for 200 msec. to reduce afterimages. The presentation times were matched to the refresh rate of the CRT monitor (possible presentation times were multiples of 1,000/95 msec. = 10.53 msec., i.e., 32, 42, 53, 63, 74, 84, and 94 msec.).

Half the trials consisted of identical matching pairs, the other half of different matching pairs. In total, this resulted in 2 (response: same/different) × 7 (presentation time) × 2 (class: featural/relational) × 2 (distinctiveness: low/high) × 7 (areas: E/N/M/EM/EN/NM/ENM) = 392 experimental trials, fully randomized across all experimental factors, which were divided in four blocks to give participants the chance to rest in short breaks in between. In order to familiarize the participants with the matching task, participants completed a special training phase before the experiment, consisting of 14 training trials with a presentation time for the targets fixed at 94 msec. In this phase, an alternative set of facial stimuli was used. The participants received acoustic feedback about the correctness of their responses. The training trials and the first two extra trials added to each test block were used as practice runs and were excluded from data analyses. The total duration of the experiment including all breaks and the post-experimental interview was between 50 and 60 min.

Written consent was obtained from each participant prior to the experimental session. As this was a nonclinical study without any harm inflicting procedure, and as all data were collected anonymously, ethical approval was not necessary according to the university's guidelines. Furthermore, written consent was obtained from each model photographed for this study including publication rights for their pictures.

Results

The Results section is structured as follows: First, I will analyze the matching performance (rate of correctly detecting differences) of the participants for all individual areas of manipulation (E = Eyes, N = Nose, M = Mouth) to assess the performance gain across presentation times to test Hypothesis 1 and to assess the onsets of matching performance for featural versus relational faces (Hypothesis 2) and for each facial area (Hypothesis 3). Second, I will present comparisons of multiple changes

(2 changes: EN, EM, NM; 3 changes: ENM) with the respective individual changes (E, N, M), e.g. EM, EN and ENM with E, to test for performance gains by processing multiple areas in line with Hypothesis 4.

Overall Performance

The overall performance (rate of correctly detecting differences) was 69.2%. Participants significantly detected manipulations within the given time frame, indicated by a directed one-sample t test against the base chance of 50% ($t_{51} = 71.6$, $p < .0001$; $d = 9.93$). The accuracy data were analyzed further by submitting matching performance to a three-way repeated-measurement ANOVA with the factors *class* (featural faces versus relational faces), area (E, M, N, EM, EN, NM, ENM), and presentation time (32, 42, 53, 63, 74, 84, and 94 msec.). The results indicated a main effect of class ($F_{1, 51} = 51.9$, $p < .0001$; $\eta^2_p = 0.50$), with percentage of recognizing manipulations in relational faces (65.2%) being lower than in featural faces (73.1%), partly supporting Hypothesis 2. As both classes of manipulation, featural faces and relational faces, were pre-experimentally matched for distinctiveness, the facial distinctiveness as such does not seem to be directly linked to matching performance, at least under the given presentation time restrictions. Additionally, main effects of area ($F_{6,306} = 44.1$, $p < .0001$; $\eta^2_p = .46$) and presentation time ($F_{6,306} = 34.7$, $p < .0001$; $\eta^2_p = 0.41$) were both significant, partly supporting Hypothesis 3. Regarding area, all levels were tested against each other via planned comparisons (Table 1) revealing all comparisons against the nose area were significant as well as most other comparisons.

The relationship between presentation time and accuracy was a monotonically increasing one: a two-tailed t test indicated ($t_{51} = 11.5$, $p < .0001$; $d = 1.59$) that even the low performance of 62.4% for the brief 32 msec. presentation condition was above chance; this actually holds true for performance at all other presentation times as well (t_{51}s > 12.8, ps $< .0001$;

TABLE 1

Differences in Matching Performance (Percentage of Correct Answers) Between All Possible Levels of Area

Area	1	2	3	4	5	6	7
1. Eyes		7.49	4.33	6.70	(−1.55)	−10.61	(0.34)
2. Eyes & mouth	−7.49		−3.16	(−0.79)	−9.03	−18.10	−7.14
3. Eyes & nose	−4.33	3.16		2.37	−5.87	−14.94	−3.98
4. Eyes, nose, & mouth	−6.70	(0.79)	−2.37		−8.24	−17.31	−6.35
5. Mouth	(1.55)	9.03	5.87	8.24		−9.07	(1.89)
6. Nose	10.61	18.10	14.94	17.31	9.07		10.95
7. Nose & mouth	(−0.34)	7.14	3.98	6.35	(−1.89)	−10.95	

Note.—All but the differences in parentheses were qualified as significant via planned comparisons.

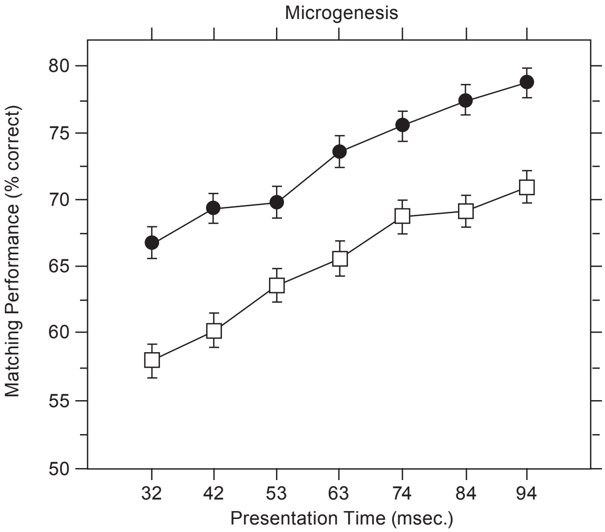

FIG. 2. The gain of matching performance over increasing presentation times, split by class: Featural (—●—); Relational (—□—).

ds > 1.78). This result extends Bachmann's finding (1991) that a face is perceived within only 100 msec. The longer the presentation time the better the performance in recognizing differences in the faces (Fig. 2). Significant differences between the performance of successive presentation times were always obtained when they differed at least 20 msec. (ps < .004), revealing clear performance gains when longer presentation times were used, thus supporting Hypothesis 1.

The main effects were further qualified by two significant interactions: between class and area ($F_{6,306} = 3.0$, $p < .01$; $\eta^2_p = 0.07$) and between area and presentation time ($F_{36,1836} = 1.4$, $p < .05$; $\eta^2_p = 0.03$). The significant interaction between area and presentation time indicates specific performance gains for matching different facial areas, providing further support for Hypothesis 3. As analyzing matching performance of each area for every presentation time is rather complex, first the performance gains were analyzed over the presentation times of matching single areas only. Subsequently, the matching performance of manipulated single areas was compared with that of manipulated multiple areas to test whether there was an interaction of processing of different facial areas.

Performance Gains Analyzing Individual Areas

To test Hypothesis 4, matching performance by means of (eta-squared) effect sizes of two-tailed t tests against a 50% chance is reported. This helps to identify onsets of performance across presentation times. As shown in Fig. 3, matching performance was always good when either the

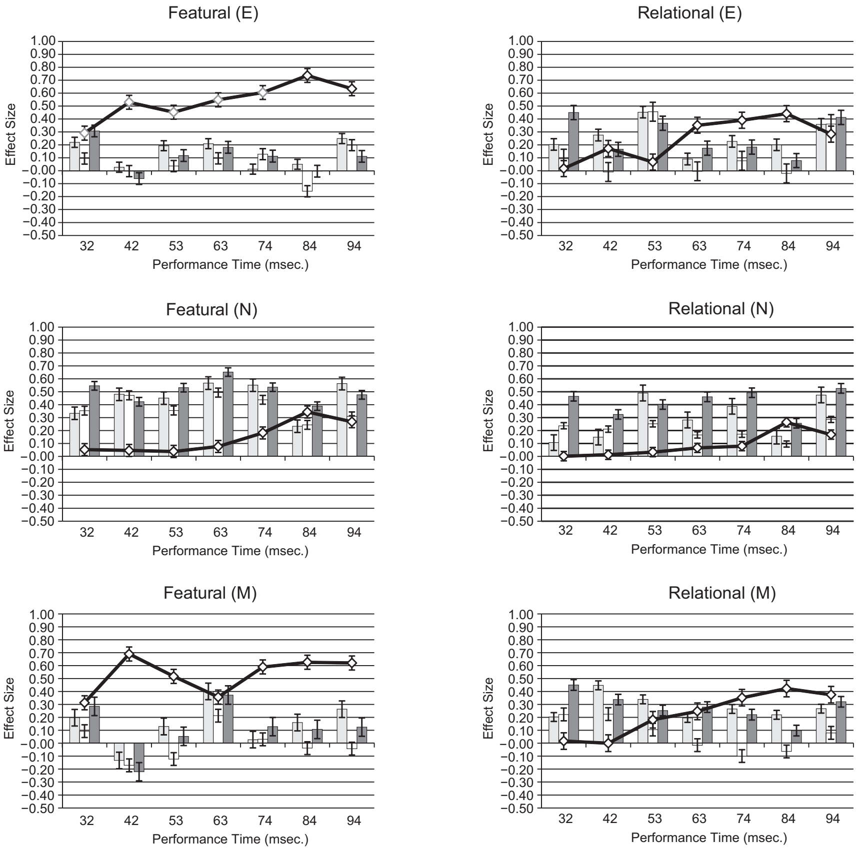

FIG. 3. Matching performance given by eta-squared effect size, whereas effect sizes for single areas result from *t* tests of matching performance against a 50% chance level, and effect sizes for multiple areas result from differences of such *t* tests and the respective effect size of the target (single) facial area. Lines show the performance gain for single areas (Eyes, Nose, Mouth) split by classes of faces (featural versus relational faces). In order to illustrate the specific gains of processing multiple facial areas, additional bars indicate the size of differential effects when a combination of the target facial area with other facial areas is presented (Eyes & Nose: ░; Nose & Mouth: □; Eyes, Nose, & Mouth: ■). Error bars indicate standard errors of the mean.

eye (E) or the mouth (M) had been changed, even within very limited presentation times. This holds true particularly for featural changes: the absolute performances were higher and the onsets of good performance were earlier than for relational changes.

For both manipulation classes, manipulations of the nose area were detected only at the longer presentation times. As changes were matched in terms of distinctiveness across all facial areas, realized by pre-studies, the late onset of matching the nose area does not seem to be an artifact of distinctiveness. Rather, there appears to be a characteristic microgenesis of matching faces within the first 100 msec. of presentation where the nose is apparently not attended to. This would make sense as the nose area contains less perceptual and social information: Noses neither move nor change their size in a situation of social communication, nor do they contain many perceptual cues for identification; therefore, mainly processing the eye and mouth region seems efficient, especially under very restricted presentation conditions.

Fig. 3 reveals another important finding: although there was a general trend of overall increasing performance across presentation times, some discontinuities were revealed. For instance, detecting featural changes to the mouth region was possible even at a presentation time of 32 msec., with a starting increase of performance and a following drop at a presentation time of 63 msec.

Performance Gains Analyzing Multiple Areas

The section above focused on the relative processing performance of matching for manipulated single areas. Manipulated multiple areas were analyzed to test Hypothesis 4. First, the effects due to the number of areas changed were analyzed (*N* changes) on matching performance, with Areas E, N, and M being 1 change, EN, EM, NM being 2 changes, and ENM being 3 changes. Then, the impact of *N* changes on matching performance of all manipulations in which the eyes or nose or mouth area were involved (e.g., for eyes: E: 1 change; EN, EM: 2 changes; ENM: 3 changes) was analyzed.

The matching performance data (rate of correctly detecting differences) were analyzed by a repeated-measurement ANOVA with the factors class (featural versus relational faces), *N* changes (1, 2, 3), and presentation time (32, 42, 53, 63, 74, 84, and 94 msec.). The results indicated main effects of class ($F_{1,51} = 48.6$, $p < .0001$; $\eta^2_p = 0.49$), *N* changes ($F_{2,102} = 81.1$, $p < .0001$; $\eta^2_p = .61$), and presentation time ($F_{6,306} = 28.4$, $p < .0001$; $\eta^2_p = .358$). Performance of matching 1-change manipulations was weakest (64.1%), followed by 2 changes (72.3%), and 3 changes (74.9%). Furthermore, main effects were qualified by an interaction between class, *N* changes, and presentation time ($F_{12,612} = 1.8$, $p < .05$; $\eta^2_p = .034$). Analysis of simple main effects of *N* changes for both relational faces ($F_{2,50} = 39.0$, $p < .0001$; $\eta^2_p = .61$) and featural faces ($F_{2,50} = 54.3$, $p < .0001$; $\eta^2_p = .69$) provided evidence for performance gains for matching faces with multiple versus single manipulated areas for featural as well as relational faces, which supports Hypothesis 4.

The specificity of performance gains for different facial areas was further investigated by conducting three independent ANOVAs for every cardinal facial area (eyes, nose, mouth) involved, thus, one ANOVA for the (E)yes area with E as 1 change, EN, EM as 2 changes, and ENM as 3 changes, a second one for the (N)ose area and a last one for the (M)outh area. Note that all three analyses show overlap in data, for instance, level ENM reflects the 3-changes condition for each cardinal feature.

For all three ANOVAs (Table 2), medium to large effects of *N* changes were found, which again supports Hypothesis 4 that the more areas were manipulated, the larger the frequency of detecting these changes was. For a deeper understanding of possible performance gains by adding manipulated facial areas for each *presentation time*, Fig. 3 shows not only the performance gains for each single cardinal facial area but also the concordant manipulations of multiple areas. Following the logic of the above analysis, performance gains of multiple areas were denoted as differences of effect sizes (eta squared) for the associated differences in performance of multiple and single areas.

TABLE 2

Results From Three Separate Repeated Measurement Three-way ANOVAs With Class, *N* Changes, and Presentation Time as Within-subject Variables and Matching Performance as Dependent Variable

	df	Eye Area			Mouth Area		
		F	*p*	η^2_p	*F*	*p*	η^2_p
Class	1/51	59.9	<.001	0.54	46.2	<.001	0.48
N changes	2/102	35.7	<.001	0.41	64.7	<.001	0.56
Presentation time	6/306	15.3	<.001	0.23	7.7	<.001	0.13
N changes * Class	2/102	2.0	ns		<1.0	ns	
N changes * Presentation time	12/612	2.1	<.05	0.04	<1.0	ns	
N changes * Class * Presentation time	12/612	1.8	<.05	0.03	2.0	<.05	0.04
		Nose Area					
Class		44.6	<.001	0.47			
N changes		168.5	<.001	0.77			
Presentation time		9.4	<.001	0.16			
N changes * Class		5.9	<.01	0.10			
N changes * Presentation time		1.1	ns				
N changes * Class * Presentation time		<1.0	ns				

Note.—Effect sizes are shown as partial eta-squared effect sizes (η^2_p).

Discussion

The first and most important finding of the present experiment is that processing of faces, here matching of sequentially presented faces, is possible when faces are presented for only 100 msec. or less (cf. Willis & Todorov, 2006). Changes to the eyes are detectable after a presentation time of only 32 msec. (note, all statements on presentation times should not be interpreted in an absolute way, depending on the specific masking, the task, and the material, these times will probably strongly vary). These findings add support to others that indicate fast processing of scenes (e.g., Biederman, 1981), objects (e.g., VanRullen & Thorpe, 2001), artwork (Augustin, *et al.*, 2008; Augustin, Defranceschi, Fuchs, Carbon, & Hutzler, 2011), and faces (e.g., Carbon & Leder, 2005; Veres-Injac, *et al.*, 2007).

The second finding is that this early processing does not occur at once, but as a so-called microgenesis of recognition as proposed by Bachmann (2000). The analysis indicates priority for processing the areas of the eyes and the mouth, and a significantly later processing onset for the nose area. One possible reason for this priority of features might be the differential information content or the diverse saliency of facial areas. In fact, saliency of a feature is known to have a strong influence on recognition performance (see Walther, Rutishauser, Koch, & Perona, 2005).

Third, the matching performance and the onset of processing were dependent on the type of facial manipulation. The participants' task was to detect changes to faces which were achieved by replacement of facial features ("featural faces") or by relational alterations of them ("relational faces"). Matching performance was more accurate and participants did not need long presentation times for matching featural faces. This might not be very intriguing if featural faces had shown different levels of distinctiveness, which can be excluded as explanation as the levels of distinctiveness and plausibility were matched via two pre-studies. Thus, this result seems to indicate a face-specific type of processing, but not a mere physical difference of stimulus material. Furthermore, in one of the pre-studies, highly distinctive relational faces had been evaluated as being less plausible and less attractive than their featural faces of comparison. It is conceivable that highly implausible faces have a greater tendency to be conspicuous, probably due to a lower frequency of real-world existence (see Hreczko & Farkas, 1994). Within the frame of the time-limited matching task of the present experiment, however, implausible faces were *not* detected more easily than plausible ones. On the contrary, even highly distinctive relational faces that gained least plausibility were processed later and with lower accuracy than featural faces. Thus, although relational information plays a key role in face recognition (Diamond & Carey, 1986; Rhodes, Tan, Brake, & Taylor, 1989), apparently this type of information is *not* being processed very efficiently within the first 100 msec. of presentation. This underlines the importance of featural information (Rakover, 2002; Carbon, 2008), at least for short presentation times (cf. Carbon & Leder, 2005). We should not forget that the present study is only capable of testing the processing of faces presented very briefly, so the possible importance of "configural processing" (Cabeza & Kato, 2000; Carbon, Grüter, Weber, & Lueschow, 2007) for faces presented much longer cannot be tested here. The whole argument regarding late configural processing is, nevertheless, based on the data of the pre-studies demonstrating that featural and relational information were not differently assessed in terms of perceived distinctiveness and plausibility. Should a more physical basis of attentional guidance be followed, a different inference could be drawn based on the fact that featural manipulations evoked larger pixel-based changes of the pictures. It is also important to note that any kind of featural change in faces also changes, at least in a subtle way, the whole configuration as such (Leder & Carbon, 2006; Carbon, Grüter, Grüter, Weber, & Lueschow, 2010); thus, featural manipulation of one facial area in the current case cannot be interpreted as a strictly local change independent of other areas.

Importantly, matching performance for featural as well as relational faces increased when more than one facial area was manipulated between prime and target face. For nearly all presentation times, higher, or more accurate, detection of differences for matching multiple facial areas were obtained compared with a single one. This supports the idea of holistic processing of faces based on parallel processing of different facial areas. Note that this definition of holistic processing is not the only possible one. For example, Wenger and Townsend (2006) discussed other possible definitions of holistic processing such as exhaustive processing of all parts of a face (see also Townsend & Wenger, 2004). Although one can find strong evidence for holistic processing in face recognition (Tanaka & Farah, 1993; Leder & Carbon, 2005), there are also some sources that found nonholistic processing when interactive processing was considered as the definition of holistic processing (e.g., Macho & Leder, 1998). Macho and Leder (1998) provided evidence for such nonholistic processing from data of a task requiring identification classification of faces in which confusion matrices were analyzed by a logit model. There are important differences between Macho and Leder's experiments and the present study, which might account for the discrepancy in results and conclusions. First, the authors used a classification task instead of a same-different matching task. Second, they used a linear logit model to test interactivity, when data did not follow linear trends. This is rather problematic for properly testing the hypothesis. Furthermore, the absence of an interactive influence was inferred when the logit model was not rejected, which gives a clear definition, without adequately reflecting possible disordinal interactions.

The results of the present experiment provide new evidence for early microgenesis of face processing which is not based on serial strategies but on parallel integration of visual information available from different facial areas. As the specificity of the employed experimental design was rather restricted in terms of the task (sequential matching task) and the material (frontal faces with neutral expression), this can certainly only be seen as the beginning of exploring so-called "early processing" of face recognition beside research efforts already done, mainly in the research domain of facial attractiveness (e.g., Locher, Unger, Sociedade, & Wahl, 1993) or trustworthiness of a face (e.g., Willis & Todorov, 2006). The use of a same-different matching task could specifically have facilitated the detection of manipulation of facial features as it might induce a feature processing mode like in a children's search game. It is also questionable whether the usage of stimuli based on the same basic faces triggers structural face processing strategies (processing the structure as such, see Bruce & Young, 1986) being more compatible with everyday life processing or trigger simplified pictorial/iconic (processing the mere picture of the face, see Carbon, 2008) processing notoriously found in lab research. Another restriction of interpretation is the usage of one, unique backward mask. There is no guarantee that this specific mask would not have different effects on featural versus relational changes or changes to the eyes/mouth versus nose areas due to subtle effects of the visual field's eccentricity. Thus, further studies could also use alternative procedures for manipulating configural processing, such as described in McKone, Martini, and Nakayama (2001).

In sum, the understanding of the microgenesis of identification and emotional recognition needs much more research (Derntl, Seidel, Kainz, & Carbon, 2009), but is needed to understand "expertise-based" as well as "prosopagnosic" face processing (Behrmann & Avidan, 2005; Grüter, Grüter, & Carbon, 2008; Grüter & Carbon, 2010).

REFERENCES

Augustin, M. D., Defranceschi, B., Fuchs, H. K., Carbon, C. C., & Hutzler, F. (2011) The neural time course of art perception: an ERP study on the processing of style versus content in art. *Neuropsychologia*, 49, 2071-2081.

Augustin, M. D., Leder, H., Hutzler, F., & Carbon, C. C. (2008) Style follows content: on the microgenesis of art perception. *Acta Psychologica*, 128, 127-138.

Bachmann, T. (1991) Identification of spatially quantised tachistoscopic images of faces: how many pixels does it take to carry identity? *European Journal of Cognitive Psychology*, 3, 87-103.

Bachmann, T. (2000) *Microgenetic approach to the conscious mind*. Amsterdam: John Benjamins.

Bachmann, T. (2007) When beauty breaks down: investigation of the effect of spatial quantisation on aesthetic evaluation of facial images. *Perception*, 36, 840-849.

Bacon-Mace, N., Mace, M. J. M., Fabre-Thorpe, M., & Thorpe, S. J. (2005) The time course of visual processing: backward masking and natural scene categorisation. *Vision Research*, 45, 1459-1469.

Behrmann, M., & Avidan, G. (2005) Congenital prosopagnosia: face-blind from birth. *Trends in Cognitive Sciences*, 9(4), 180-187.

Biederman, I. (1981) On the semantics of a glance at a scene. In J. R. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum. Pp. 213-253.

Bruce, V., & Young, A. (1986) Understanding face recognition. *British Journal of Psychology*, 77, 305-327.

Cabeza, R., & Kato, T. (2000) Features are also important: contributions of featural and configural processing to face recognition. *Psychological Science*, 11, 429-433.

Carbon, C. C. (2003) Face processing: early processing in the recognition of faces. Doctoral thesis, Berlin, DARWIN, Freie Universität Berlin, retrieved from http://www.diss.fu-berlin.de/2003/35/.

Carbon, C. C. (2008) Famous faces as icons: the illusion of being an expert in the recognition of famous faces. *Perception*, 37, 801-806.

Carbon, C. C., Grüter, T., Grüter, M., Weber, J. E., & Lueschow, A. (2010) Dissociation of facial attractiveness and distinctiveness processing in congenital prosopagnosia. *Visual Cognition*, 18, 641-654.

Carbon, C. C., Grüter, T., Weber, J. E., & Lueschow, A. (2007) Faces as objects of non-expertise: processing of Thatcherised faces in congenital prosopagnosia. *Perception*, 36, 1635-1645.

Carbon, C. C., & Leder, H. (2005) When feature information comes first! Early processing of inverted faces. *Perception*, 34, 1117-1134.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993) PsyScope: a new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257-271.

Derntl, B., Seidel, E. M., Kainz, E., & Carbon, C. C. (2009) Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, 38, 1849-1862.

Diamond, R., & Carey, S. (1986) Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, 115(2), 107-117.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. J. (2000) A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13, 171-180.

Flavell, J. H., & Draguns, J. D. (1957) A microgenetic approach to perception and thought. *Psychological Bulletin*, 54, 197-217.

Grill-Spector, K., & Kanwisher, N. (2005) Visual recognition: as soon as you know it is there, you know what it is. *Psychological Science*, 16(2), 152-160.

Grüter, T., & Carbon, C. C. (2010) Escaping attention: some cognitive disorders can be overlooked. *Science*, 328, 435-436.

Grüter, T., Grüter, M., & Carbon, C. C. (2008) Neural and genetic foundations of face recognition and prosopagnosia. *Journal of Neuropsychology*, 2, 79-97.

Hreczko, T. A., & Farkas, L. G. (1994) Norms of the craniofacial asymmetries in North American Caucasians. In L. G. Farkas (Ed.), *Anthropometry of the head and face.* (2nd ed.) New York: Raven. Pp. 359-380.

Leder, H., & Carbon, C. C. (2005) When context hinders! Context superiority versus learn-test-compatibilities in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 58A, 235-250.

Leder, H., & Carbon, C. C. (2006) Face-specific configural processing of relational information. *British Journal of Psychology*, 97, 19-29.

Leder, H., Carbon, C. C., & Ripsas, A. L. (2006) Entitling art: influence of different types of title information on understanding and appreciation of paintings. *Acta Psychologica*, 121, 176-198.

Locher, P. J., Unger, R. K., Sociedade, P., & Wahl, J. (1993) At first glance: accessibility of the physical attractiveness stereotype. *Sex Roles*, 28, 729-743.

Macho, S., & Leder, H. (1998) Your eyes only? A test of interactive influence in the processing of facial features. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1486-1500.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002) The many faces of configural processing. *Trends in Cognitive Sciences*, 6(6), 255-260.

McKone, E., Martini, P., & Nakayama, K. (2001) Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 573-599.

Rakover, S. S. (2002) Featural vs. configurational information in faces: a conceptual and empirical analysis. *British Journal of Psychology*, 93, 1-30.

Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989) Expertise and configural coding in face recognition. *British Journal of Psychology*, 80, 313-331.

Tanaka, J. W., & Farah, M. J. (1993) Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, 46, 225-245.

Tanaka, J. W., & Sengco, J. A. (1997) Features and their configuration in face recognition. *Memory & Cognition*, 25, 583-592.

Townsend, J. T., & Wenger, M. J. (2004) A theory of interactive parallel processing: new capacity measures and predictions for a response time inequality series. *Psychological Review*, 111, 1003-1035.

VanRullen, R., & Thorpe, S. J. (2001) Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, 30, 655-668.

Veres-Injac, B., Hofer, F., & Schwaninger, A. (2007) The time course of processing external and internal features of unfamiliar faces. *Perception*, 36(Suppl.), 155. [Abstract]

Walther, D., Rutishauser, U., Koch, C., & Perona, P. (2005) Selective visual attention enables learning and recognition of multiple objects in cluttered scenes. *Computer Vision and Image Understanding*, 100, 41-63.

Wenger, M. J., & Townsend, J. T. (2006) On the costs and benefits of faces and words: process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 755-779.

Willis, J., & Todorov, A. (2006) First impressions: making up your mind after a 100-msec. exposure to a face. *Psychological Science*, 17, 592-598.

Yin, R. K. (1969) Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987) Configurational information in face perception. *Perception*, 16, 747-759.

Accepted November 24, 2011.