

Structural Encoding of Human and Schematic Faces: Holistic and Part-Based Processes

Noam Sagiv¹ and Shlomo Bentin

Abstract

■ The range of specificity and the response properties of the extrastriate face area were investigated by comparing the N170 event-related potential (ERP) component elicited by photographs of natural faces, realistically painted portraits, sketches of faces, schematic faces, and by nonface meaningful and meaningless visual stimuli. Results showed that the N170 distinguished between faces and nonface stimuli when the concept of a face was clearly rendered by the visual stimulus, but it did not distinguish among different face types: Even a schematic face made from simple line fragments triggered the N170. However, in a second experiment, inversion seemed to have a different effect on natural faces in which face components were available and on the

pure gestalt-based schematic faces: The N170 amplitude was enhanced when natural faces were presented upside down but reduced when schematic faces were inverted. Inversion delayed the N170 peak latency for both natural and schematic faces. Together, these results suggest that early face processing in the human brain is subserved by a multiple-component neural system in which both whole-face configurations and face parts are processed. The relative involvement of the two perceptual processes is probably determined by whether the physiognomic value of the stimuli depends upon holistic configuration, or whether the individual components can be associated with faces even when presented outside the face context. ■

INTRODUCTION

Accumulated findings from neuropsychology as well as primate electrophysiology point to some degree of domain specificity in the functional organization of the ventral pathway of the visual system (for reviews, see Farah, 1996; Logothetis & Scheinberg, 1996; see also Farah, Wilson, Drain, & Tanaka, 1998; Logothetis & Pauls, 1995; Logothetis, Pauls, & Poggio, 1995). In particular, there is considerable evidence indicating that the perception of faces is mediated by extrastriate mechanisms specifically tuned to process physiognomic information (Allison, Ginter, et al., 1994; Allison, Puce, Spencer, & McCarthy, 1999; McCarthy, Puce, Gore, & Allison, 1997; McCarthy, 1999; McCarthy, Puce, Belger, & Allison, 1999; Puce, Allison, & McCarthy, 1999; Kanwisher, McDermott, & Chun, 1997). This hypothesis is supported by single-cell recordings in primates (for a review, see Desimone, 1991), by intracranial recordings of event-related potentials (ERPs) in humans (e.g., Allison, McCarthy, Nobre, Puce, & Belger, 1994), as well as by hemodynamic imaging studies using PET (e.g., Haxby et al., 1993; Sergent, Ohta, & MacDonald, 1992), and fMRI (e.g., Kanwisher et al., 1997; Puce, Allison, Gore, & McCarthy, 1995).

Selective activation of face-specific neural networks at a relative early stage of visual processing may increase

face recognition efficiency by restricting subsequent identification stages to a particular class of internal models—that of faces (Ullman, 1996). Indeed, several models of face recognition incorporate an initial “structural encoding” mechanism with physiognomic features as input and an integrated, but not fully identified, representation of a face as output (e.g., Moses, Edelman, & Ullman, 1993; Bruce & Young, 1986). The resulting structural code allows the identification of particular exemplars of faces despite changes in viewing angle, expression, lighting, age, or paraphernalia. According to such models, face identification is achieved by the combined operation of face recognition units and semantic nodes. The face recognition units select the prestored face model that best fits the currently structured perceptual representation and the semantic nodes provide the entire knowledge associated in semantic memory to the particular face model selected. The domain specificity of the structural encoder and its categorized output representation prevent the face recognition system from attempting to match the resulting representation to all possible 3-D object models. Indeed, in the absence of such specificity, face recognition is inhibited and becomes inefficient (Bentin, Deouell, & Soroker, 1999).

To date, the research into the rules governing the encoding of faces has focused mainly on the role of structural encoding in face identification (see reviews in Bruce & Young, 1998; Ullman, 1996; Rhodes, 1995). For

example, these studies examined how variation in viewing conditions influence face recognition (e.g., viewing position and illumination direction, Moses, Adini, & Ullman, 1994) or stimulus configuration (e.g., vertical orientation, Rhodes, Brake, & Atkinson, 1993; Tanaka & Farah, 1991, for a review, see Valentine, 1988; spatial relationship between components, Sergent, 1984; or analytic vs. holistic encoding, Tanaka & Farah, 1993). Relatively less effort has been directed toward investigating the functional architecture of the neural mechanism that mediates the structural encoding of faces in the visual system. Initial steps in this direction have been taken, however, by recording ERPs elicited by faces and objects in human volunteers (e.g., Bentin, Allison, Puce, Perez, & McCarthy, 1996; George, Evans, Fiori, Davidoff, & Renault, 1996).

Bentin et al. (1996) described a negative potential with a mean peak latency of 172 msec (N170) that was elicited by human faces but not by animal faces, cars, or butterflies. This component was larger over the right than the left hemisphere, and largest in the posterior temporal areas. Because the familiarity of the face was found inconsequential for the N170 (Bentin & Deouell, 2000), Bentin and his colleagues suggested that although essential for efficient face recognition, the mechanism associated with the N170 acts on basic physiognomic features and precedes within-category identification. Hence, the N170 may well constitute an electrophysiological manifestation of the neural analogue to the structural encoder proposed by Bruce and Young (1986)—a mechanism specialized to detect physiognomic features and extracts the visual characteristics needed to form an internal representation of a human face.

Several experimental manipulations were aimed at understanding how the structural face-encoding mechanism functions. Isolated eyes or combinations of inner components presented without the face contour elicited an N170 significantly larger than that elicited by full faces (Bentin et al., 1996; Bentin & McCarthy, unpublished). In the initial studies of N170, the response to inverted human faces was slightly more negative (i.e., larger) and significantly delayed relative to that elicited by upright faces (Bentin et al., 1996). More recent studies showed that, indeed, the N170 elicited by inverted faces is significantly larger than that elicited by upright faces (Rossion et al., 1999, 2000). A possible interpretation of this pattern is that the amplitude enhancement, and particularly the delay of the N170 latency, reflects a more difficult encoding of inverted than of upright faces. This suggestion is based on the assumption that face encoding is heavily based on holistic processing of the face gestalt, which is distorted by inversion (Rossion et al., 1999). Surprisingly, however, distortion of the inner-face component configuration did not affect the response significantly; in fact, the amplitude of the N170 elicited by faces with a distorted configuration of inner components was slightly smaller than for normal faces (Bentin

et al., 1996). Furthermore, the addition of the nose and the lips to isolated eyes (in the absence of the face contour) enhanced rather than reduced the amplitude of the N170, and a face without inner components elicits a significantly attenuated response (Bentin & McCarthy, unpublished). These data speak against the assumption that the enhancement of the N170 amplitude is associated with a more difficult perceptual process. Alternatively, they suggest the possibility of having a complex neural system for structural encoding, which involve both a component processing system and a holistic face processor. According to this view, the scalp-recorded N170 is influenced by the face component processor more than by the holistic face perception processor. Support for this hypothesis comes from recordings of the intracranial analogue of the N170, the N200 potential. These data show that, recorded from the face areas in the fusiform, the N200 is bigger in response to faces than in response to face components. Laterally from the fusiform, however, there are regions where the N200 is larger for isolated eyes than for faces (McCarthy et al., 1999). The present study was designed to explore some of the basic characteristics of the structural encoding mechanisms and provide evidence pertinent to the dual mechanisms hypothesis.

EXPERIMENT 1

Schematic faces (e.g., Smiley) are instantly perceived as faces. Even a set of fruits, vegetables, or other objects may be perceived as a face when organized in a face configuration. Is the visual information conveyed by such stimuli sufficient to trigger the face-specific structural encoder? Note that, alternatively, it is possible that the perception of schematic faces relies on a general-purpose visual mechanism and the face representation is formed at a higher-level of perceptual integration. In the present experiment, we addressed this question by comparing the ERPs elicited by four categories of faces: photographs of natural faces, realistically painted portraits, sketches of faces, and schematic faces. Photographs of flowers, as well as equiluminant scrambled faces and scrambled schematic faces were used as control stimuli to determine face specificity (Figure 4). Participants were engaged in an oddball paradigm, silently counting butterflies that were sporadically presented on the screen.

Assuming that the N170 is, indeed, associated with the structural encoding process, we hypothesized that the minimal requirements for eliciting the N170 should also constitute the minimal requirements for the perceptual system to determine that a face exists in the visual field and activate the face-specific perceptual module. Hence, the modulation of the N170 by different face or face-like stimuli could indicate whether this domain-specific module in the visual system is constrained to process rigidly defined visual primitives, or whether it can adapt

itself, and be triggered by a variety of visual stimuli which may, potentially, represent a face.

The components of a schematic face, such as Smiley, do not usually convey physiognomic information in isolation. Their interpretation as face parts depends on the configuration in which they are embedded. Therefore, apart from testing the range of the face structural encoder sensitivity, comparison of the evoked responses to real faces with the responses to schematic faces was expected to yield new insights regarding the functional organization of the structural encoding system with which the N170 ERP component is probably associated.

Results

Replicating previous studies, natural faces elicited a robust N170 component, maximal at the posterior temporal sites. At the same time, a positive peak was measured at the anterior sites (P190, cf. Jeffreys & Tuckwell, 1992). The negative (or negative going) ERPs elicited by equiluminant nonsense stimuli, as

well as photographs of flowers within the same time range, were significantly smaller (Figure 1). The outstandingly large amplitude of the N170 elicited by faces compared to flowers and meaningless stimulus categories in the present study, as well as similarly big differences between the response to faces and the response to other types of stimuli reported in previous studies, support our hypothesis that the N170 is a manifestation of an early face-specific visual processing mechanism. Given our present interest in this neural mechanism, we will not discuss differences between stimulus conditions that occurred later than the N170.

The major manipulation in this experiment involved presenting different face types at different levels of physical resemblance to photographs of natural faces. The consequences of this manipulation are presented in Figure 2, which compares the N170 elicited at the right and the left mastoid scalp locations by the different face categories. Relative to nonface stimuli (Figure 1), all four face types elicited more conspicuous, but not identical, N170 components.

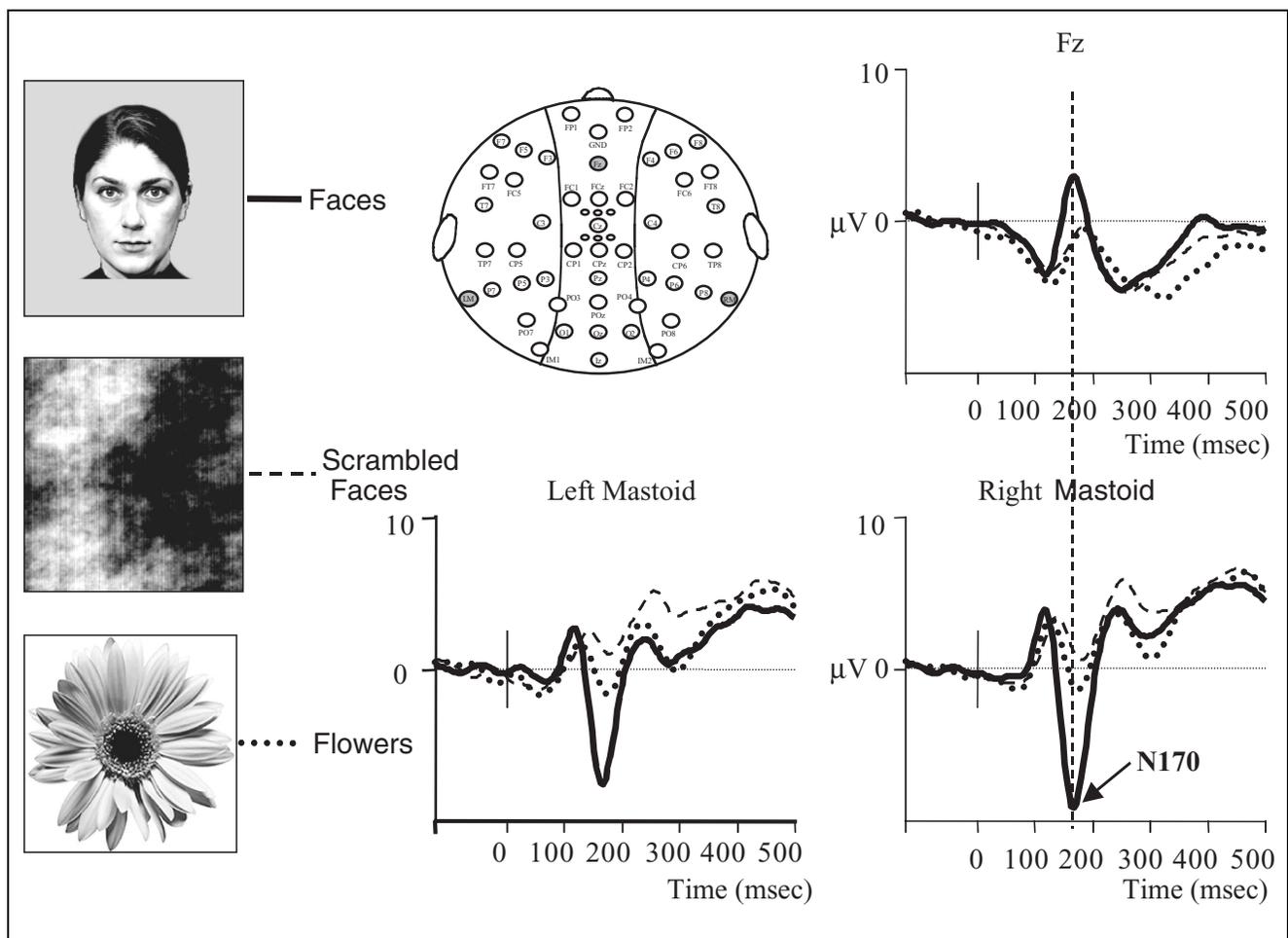


Figure 1. Demonstration of the basic difference between the N170 elicited by faces and that elicited by nonface stimuli. Note the polarity inversion at 170 msec between the posterior-lateral and the anterior sites.

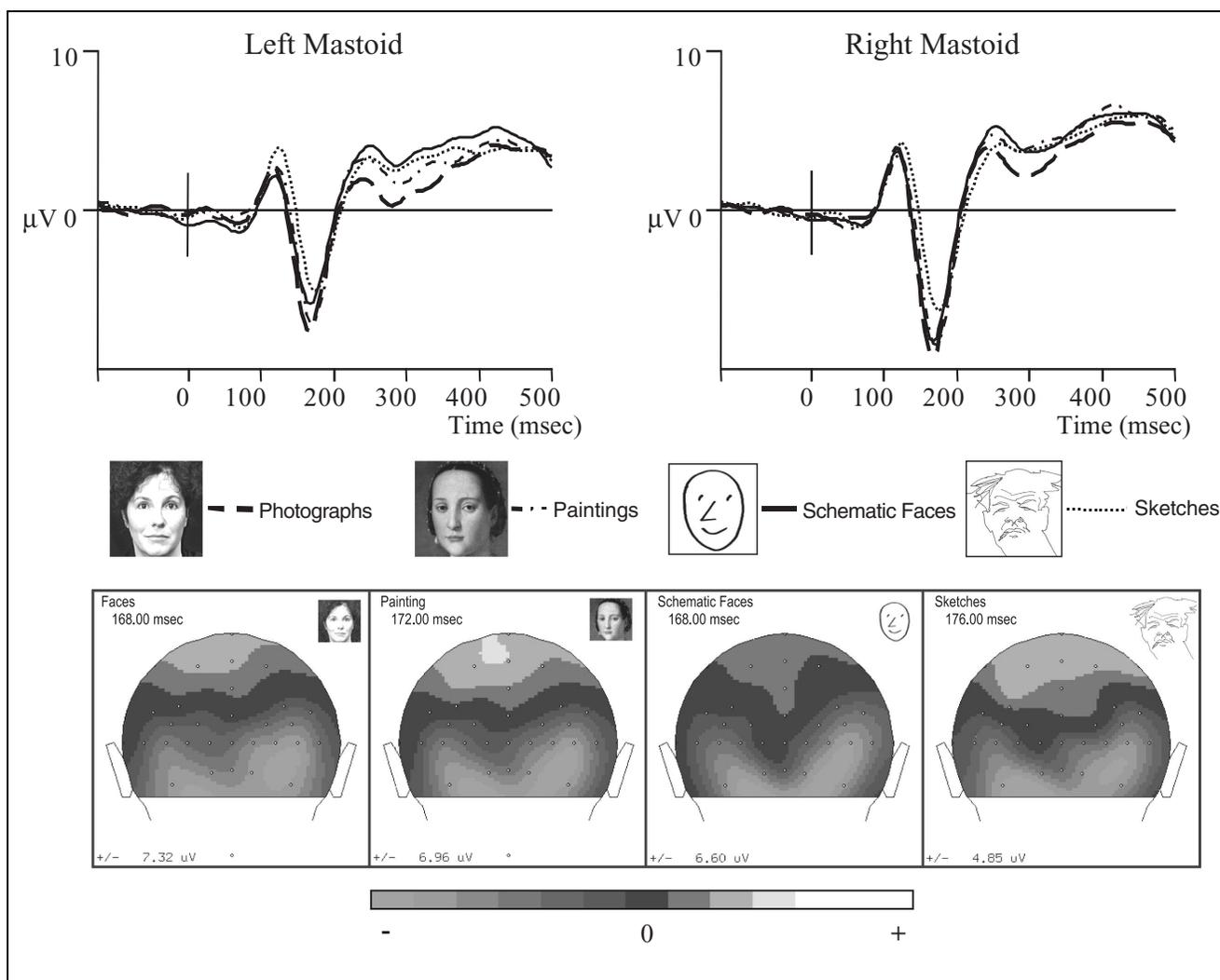


Figure 2. The N170 potentials elicited at the left and right mastoids by the four face types, and the scalp distribution of the potentials at this latency (spline interpolation).

The statistical reliability of the face/nonface difference and of the apparent differences between the N170 elicited by each face type was assessed by a series of within-subject, three-factor analysis of variance (ANOVA) with repeated measures. The factors were stimulus type (four face types, flowers, scrambled faces),² hemisphere (left, right), and site (P7/8, left and right mastoids, PO7/8, IM1/2).³ The dependent variables were the N170 amplitude and its latency (Table 1).

The ANOVA of the amplitude showed that there were significant differences among the different stimulus types [$F(5,145) = 85.6, p < .001$], that the N170 was larger at right ($-6.52 \mu\text{V}$) than at left ($-4.92 \mu\text{V}$) hemisphere sites [$F(1,29) = 19.6, p < .001$], and that there was a significant effect of site [$F(3,87) = 59.8, p < .005$]. Post hoc comparisons revealed that the N170 elicited by faces (weighted across all four face categories) was significantly more negative than that elicited by nonfaces (weighted across scrambled faces and flowers) [$F(1,29) = 183.6, p < .001$]. The ERPs elicited by flowers

were significantly less negative than those elicited by sketches [$F(1,29) = 56.1, p < .001$], but significantly more negative than those elicited by scrambled faces [$F(1,29) = 22.3, p < .001$]. Because our present experiment focused on the potential difference in processing different face types, and because of the clear distinction between the N170 elicited by faces and the ERPs elicited by nonface stimuli in the same time range, our subsequent analyses focused only on the four face types. The design of these ANOVAs was similar to the one described above except that the stimulus type factor included only the four face categories.

The ANOVA of the amplitude showed that all three main effects were significant [$F(3,87) = 11.0, p < .001$; $F(3,87) = 3.7, p < .025$; $F(1,29) = 20.3, p < .001$, for stimulus type, site, and hemisphere, respectively]. Post hoc contrasts showed that the amplitude of the N170 elicited by sketches was significantly less negative than that elicited by the three other face types [$F(1,29) = 46.8, p < .001$]. There was no significant difference

Table 1. (A) Mean Amplitudes (in μV) and (B) Mean Latencies (in msec) of the N170 Elicited by Natural Faces, Painted Portraits, Sketches, Schematic Faces Flowers, and Scrambled Faces at the Posterior Temporal Scalp Sites

	<i>Left Hemisphere</i>				<i>Right Hemisphere</i>			
	<i>Left mastoid</i>	<i>P7</i>	<i>PO7</i>	<i>IM1</i>	<i>IM2</i>	<i>PO8</i>	<i>P8</i>	<i>Right mastoid</i>
<i>A</i>								
Natural Faces	-7.2	-6.2	-6.0	-7.6	-8.9	-9.3	-9.1	-9.7
Painted Portraits	-7.3	-6.1	-6.2	-7.8	-8.8	-8.8	-9.2	-9.8
Sketches	-5.8	-4.9	-4.1	-6.1	-7.3	-6.4	-7.0	-7.6
Schematic Faces	-6.8	-6.3	-6.0	-7.0	-8.3	-8.8	-9.6	-9.4
Flowers	-2.9	-3.2	-2.3	-2.8	-3.0	-2.9	-3.3	-3.4
Scrambled Faces	-1.9	-1.7	-0.7	-1.3	-1.5	-0.7	-1.6	-2.2
<i>B</i>								
Natural Faces	170	168	170	168	169	169	170	170
Painted Portraits	171	171	172	171	171	171	172	172
Sketches	177	171	171	174	175	173	176	178
Schematic Faces	170	165	165	168	167	167	169	169
Flowers	174	166	168	171	172	168	171	176
Scrambled Faces	177	167	166	173	175	166	168	178

between the amplitudes of N170 elicited by photographs, paintings, and schematic faces. The N170 amplitudes across face categories were largest at the mastoids and, in descending order, IM1/2, P7/8, and PO7/8. Significant differences were found only between the mastoids and the P7/8 and PO7/8 locations [$F(1,29) = 5.0$, $p < .05$ and $F(1,29) = 4.8$, $p < .05$, respectively], and between the IM1/2 and the PO7/8 locations [$F(1,29) = 4.9$, $p < .05$].

All three first-order interactions were significant [$F(9,261) = 4.8$, $p < .001$; $F(3,87) = 3.1$, $p < .05$; $F(3,87) = 6.3$, $p < .01$ for the Stimulus Type \times Site, Stimulus Type \times Hemisphere, and Site \times Hemisphere, respectively]. The second-order interaction between all three factors was also significant [$F(9,261) = 3.4$, $p < .001$]. Post hoc examination of these interactions revealed the following pattern: The difference between the N170 elicited by sketches and that elicited by the other face types was larger at the right hemisphere sites than at the left hemisphere sites, and slightly smaller at IM sites than at the other three sites. The interaction between hemisphere and scalp location is probably a result of the smaller (but clear) right hemisphere advantage at scalp site IM1/2 [$F(1,29) = 9.0$, $p < .01$], compared to other sites [$F(1,29) = 15.3$, $p < .001$; $F(1,29) = 18.2$, $p < .001$; $F(1,29) = 24.2$, $p < .001$ for PO7/8, mastoids, and P7/8, respectively]. The second-order interaction reflected (a) a smaller (Site \times Hemi-

sphere) interaction for the sketched faces category [$F(3,87) = 2.3$ vs. $F(3,87) > 5.9$ for each of the other categories]; (b) an insignificant (Stimulus Type \times Hemisphere) interaction at the IM scalp site [$F(3,87) < 1.00$] versus a significant or close-to-significant effect for other sites [$F(3,87) > 2.4$]; and (c) a smaller (Stimulus Type \times Site) interaction in the right hemisphere [$F(15,435) = 3.0$] than in the left hemisphere [$F(15,435) = 5.5$].

As evident in Table 1B, the mean N170 latency was shortest for schematic faces (167.6 msec), followed by photographs (169.3 msec), paintings (171.4 msec), and sketches (174.3 msec). Scalp sites also differed on N170 latency; the site at which the shortest N170 latency was recorded was PO7/8 (169.7 msec), followed by P7/8 (170.2 msec), IM1/2 (170.4 msec), and R/L-Mast (172.3 msec). ANOVA showed that both the stimulus type and the site effects were significant [$F(3,87) = 16.7$, $p < .001$; $F(3,87) = 5.0$, $p < .005$, respectively]. Post hoc contrasts revealed that differences between the mean N170 latencies of all stimulus type pairs were highly significant [$F(1,29) = 9.2$, $p < .005$], except for the difference between face photograph and schematic face latencies [$F(1,29) = 3.0$, $p < .1$]. In addition, the N170 latency at the mastoid scalp locations (172.3 msec) was significantly delayed [$F(1,29) = 17.3$, $p < .001$] compared with the three other locations (170.1 msec), which did not significantly differ from each other [$F(1,29) = .5$, $p = .6$].

Hemispheric differences failed to reach significance [$F(1,29) < 1.0$]. A Site \times Hemisphere interaction was, however, reliable [$F(3,87) = 4.0, p < .005$]. This interaction is possibly the result of relatively late right hemisphere mean latency at P7/8 and PO7/8 scalp locations. The interaction between stimulus type and site was also significant [$F(9,261) = 3.5, p < .001$], reflecting a greater site effect within the sketched face category. The Stimulus Type \times Hemisphere interaction was not significant [$F(3,87) = 1.4, p < .25$]. The interaction between all three factors did not reach significance [$F(9,261) = 1.7, p = .1$].

Discussion

Experiment 1 demonstrated that face-specific N170 components can be recorded with scalp electrodes not only when photographs of natural faces are presented, but also when paintings, sketches, and even highly schematic drawings of faces are used as stimuli. The amplitude of the N170 elicited by all face stimuli was at least twice as big as the amplitude of the negative potential elicited by flowers, and three times as big as the amplitude of the potential elicited by nonsense images during the same time. The significant difference between flowers and scrambled faces could reflect either that the neural activity associated to the structural encoding of any object is distinct from that elicited by visual patterns that are not objects (McCarthy et al., 1997), or that there is specificity for other stimulus categories in the vicinity of the face area and that part of this specificity is picked up by the far-field recording (Ishai, Ungerleider, Martin, Schouten, & Haxby, 1999). The present study was not designed to address this issue. Therefore, we will turn now to the focus of this experiment, that is the comparison between the different face types.

The finding that the N170 elicited by schematic faces is not significantly smaller in amplitude is not delayed relative to natural face photographs, and has a similar scalp distribution, suggests that common neural mechanisms underlie the perception of both natural and schematic faces. Assuming that these mechanisms are the neural analogue to the hypothesized structural encoder for faces, these data also suggest that a face-specific visual mechanism is triggered whenever a stimulus contains sufficient information to generate the concept of a face. In other words, schematically drawn faces are categorized by an early visual mechanism rather than by a higher-level cognitive process. Indeed, it is possible that, once triggered, such a face perception module may take precedence over a more general visual processor. A similar account for the suppression of acoustic perception when phonetic features are detected (Bentin & Mann, 1990; Liberman & Mattingly, 1989) led to the hypothesis that the phonetic module may preempt auditory information that is relevant to phonetic perception (Whalen & Liberman, 1987).

The reduced amplitude of the N170 elicited by detail-rich face sketches relative to that elicited by schematic faces, and its delayed peak latency, might be interpreted *prima facie* as evidence against the above hypothesis. However, post hoc, this result can be incorporated into our theory. Although graphically more detailed, the sketches did not attempt to imitate a natural face (as the paintings), or to emphasize those basic features apparently required for the formation of a face concept (as in the schematic faces). Indeed, it is possible that some of the sketches were difficult to interpret as human faces without an elaborate analytic process. Hence, the generation of a face concept on the basis of such sketches may have been somewhat more time consuming either because they were processed by general visual processors or because they activated the special purpose structural encoder later. In addition, unlike the schematic faces that, although not identical, were more uniform in shape,⁴ the sketches were stylistically more diverse. As a consequence, some of the sketches may have introduced noise or a time jitter in the average ERP, which could explain both the reduced amplitude and the longer latency of the N170 associated with the sketched faces category.⁵ All face categories elicited a significantly larger N170 in the right hemisphere. However, the asymmetry was reduced for sketches and absent for nonface categories. This finding is consistent with neuropsychological (e.g., De Renzi, 1997), neuroimaging (e.g., McCarthy et al., 1997; De Renzi, Perani, Carlesimo, Silveri, & Fazio, 1994; Haxby et al., 1993), and behavioral studies (e.g., Luh, Rueckert, & Levy, 1991) indicating a right hemisphere advantage for face processing. Furthermore, it corroborates the hypothesis that some of the sketches were processed by a general rather than a face-specific visual processor or that, in general, the sketches were less efficient in initiating a face-specific process.

The finding that schematic faces elicit a conspicuous N170 that does not significantly differ from the ERPs elicited by natural faces is intriguing in light of the assumption that the N170 is associated primarily with the processing of face components. In previous studies, we have consistently found that eyes in isolation or eyes combined with other inner-face components elicit a larger N170 than that elicited by full faces (Bentin et al., 1996; Bentin & McCarthy, unpublished). Furthermore, full faces elicit an N170, which is only slightly larger than that elicited by face contours from which all inner components have been erased. On the basis of these data, Bentin et al. (1996) suggested that the N170 is modulated by two neural sources (both part of the structural encoding mechanism): one, possibly located in the posterior upper bank of the occipito-temporal sulcus (OTS) or in the inferior temporal (IT) gyrus, is triggered by physiognomic information appearing in the visual field; the other, probably located in the middle fusiform gyrus, is responsible for holistic face processing

and integrates the face parts into a whole. Assuming that without the framing contour and the face-specific configuration, inner components of schematic faces have no physiognomic value,⁶ the N170 elicited by schematic faces should be modulated primarily by the fusiform source, which is responsible for the holistic processing of a face. Consequently, the N170 should be particularly susceptible to manipulations affecting holistic processing of schematic faces. An initial attempt to explore this possibility is reported in Experiment 2. In addition, the similar latency for the N170s elicited by natural faces and schematic faces would thus be in agreement with the primacy of holistic processing for face stimuli; the lateral component, primarily responsible for the N170 is triggered by the holistic component, whether or not the face parts are recognizable as such in isolation or defined by the global configuration.

EXPERIMENT 2

The face inversion effect—that is, the difficulty to identify familiar faces presented upside down (compared with other familiar objects subject to a similar manipulation)—has been the focus of many studies for more than two decades (reviewed by Valentine, 1988). The difference between the effects of face inversion relative to inversion of other objects is usually attributed to a basic difference between the processing strategies incurred by faces versus objects. According to this hypothesis, recognition of faces relies on the formation of a whole-face representation from which relations between the inner-face parts are derived, that is, holistic processing (Farah, Tanaka, & Drain 1995; Tanaka & Farah, 1993; versions of the holistic hypothesis are reviewed by Moscovitch, Winocur, & Behrmann, 1997). In contrast, the recognition of other, perhaps less stereotypic objects, entails initial identification of basic components, which are then related one to the other to form an unequivocal visual percept (e.g., Biderman, 1987; Marr, 1982; Marr & Nishihara, 1978). The face inversion effect has become the hallmark of normal face recognition in humans and has often been used to distinguish face-specific from general visual processes.

Bentin et al. (1996) compared the N170 elicited by upright and inverted faces and found that the N170 elicited by inverted faces was slightly larger and peaked 10 msec later than that elicited by upright faces; only the difference in latency was significant in that study. Other investigators, however, reported that the enhancement of the N170 amplitude by face inversion is significant as well as the delay in latency. (e.g., Rossion et al., 1999, 2000). A possible interpretation of the enhanced N170 amplitude in response to inverted faces is that the scalp-recorded N170 is associated primarily with the perception of face components, a process that might not be impeded by inversion. This hypothesis conforms to the presumed anatomical location of the putative compo-

nents detector (in lateral OTS and IT gyrus) and the orientation of these cell assemblies relative to the recording electrodes on the scalp. In contrast to the N170, the intracranially recorded N200 (on the ventral plane of the temporal lobe) is modulated primarily by the holistic, face integration mechanism located in the posterior lateral parts of the fusiform gyrus (e.g., Allison et al., 1999). Hence, given the orientation of the electrical dipole fields in these two cell assemblies, a manipulation that inhibits holistic perception should enhance the N170 amplitude. This hypothesis also conforms to the finding that the N200 (recorded at the fusiform) elicited by faces is “larger” than that elicited by isolated eyes (McCarthy et al., 1999), whereas the N170 (recorded at the scalp) elicited by faces is “smaller” than that elicited by isolated eyes (Bentin et al., 1996).

According to this model, face perception is specific because the highly stereotypic organization of faces permits categorizing the visual stimulus before a full description of its features is provided by the analytic mechanism. Therefore, most of the time, faces activate the holistic perception mechanism instantly. Indeed, only such a model has the capacity to explain the results of Experiment 1 in which a visual stimulus that can be perceived as a face only if processed as a whole (i.e., a highly schematic face) elicited a similar N170 to that elicited by photographs of natural faces.

Based on the above conceptualization, we hypothesize that if the detection of schematic faces is based primarily on the holistic face configuration, compromising their gestalt by presenting the stimuli upside down should reduce the N170. In contrast, turning natural faces upside down should enhance the amplitude of the N170 since the face component processing module would still be activated. To test this hypothesis, we presented a new group of participants with natural and schematic faces, in upright and inverted orientations.

Results

Figure 3 presents the N170s elicited by photographs of natural faces and schematic faces presented upright or upside down. Replicating the findings of Rossion et al. (1999), the N170 elicited by inverted natural faces was larger than that elicited by upright faces. In contrast, the N170 elicited by inverted schematic faces was reduced in amplitude relative to that elicited by upright schematic faces. Inversion delayed the peak latency of the N170 for schematic as well as for natural faces.

The reliability of these observations was statistically analyzed by a within-subject ANOVA using repeated measures (see data in Table 2A).

The factors were stimulus type (natural faces, schematic faces), stimulus orientation (upright, inverted), site (P7/8, mastoids, PO7/8, IM1/2), and hemisphere (left, right). As a main effect, stimulus orientation was not

Table 2. (A) Amplitudes (in μV) and (B) Latencies (in msec) of the N170 Elicited by Upright and Inverted Natural and Schematic Faces at the Posterior Temporal Scalp Sites

	<i>Natural Faces</i>				<i>Schematic Faces</i>			
	<i>Right Hemisphere</i>		<i>Left Hemisphere</i>		<i>Right Hemisphere</i>		<i>Left Hemisphere</i>	
	<i>Upright</i>	<i>Inverted</i>	<i>Upright</i>	<i>Inverted</i>	<i>Upright</i>	<i>Inverted</i>	<i>Upright</i>	<i>Inverted</i>
<i>A</i>								
Mastoid	-6.4	-8.6	-5.8	-8.2	-5.4	-5.5	-5.1	-3.5
P8/7	-4.8	-6.8	-4.5	-5.4	-4.1	-3.0	-3.2	-3.3
PO8/7	-6.8	-8.8	-4.8	-5.0	-5.6	-5.6	-5.0	-2.3
IM2/1	-5.3	-8.1	-4.9	-7.3	-5.0	-2.9	-6.5	-3.9
Mean	-5.8	-8.1	-5.0	-6.5	-5.0	-4.25	-4.9	-3.25
<i>B</i>								
Mastoid	149	162	151	165	159	164	158	165
P8/7	151	163	154	166	158	164	162	166
PO8/7	151	160	152	164	151	161	159	166
IM2/1	150	162	147	161	161	164	154	162
Mean	150	162	151	164	157	163	158	165

weighted more than its reduction resulting from the inversion of schematic faces. However, the second-order interaction between orientation, hemisphere, and stimulus type was not significant [$F(1,15) < 1.0$]. Finally, a series of second-order interactions involving the effect of site suggested that all the effects were differently conspicuous at different sites. As revealed in Table 2, by and large, it seems that whereas natural faces induced similar inversion effects across the analyzed sites, the inversion effect for schematic faces varied at different sites. We had no prior assumptions about the distribution of the effects within the posterior temporal scalp regions. Furthermore, the spatial resolution of scalp-recorded ERPs is too low to permit reliable interpretation of site effects within a relative small region, unless the same distribution is replicated; therefore, a detailed analysis of this distribution of effects was not attempted post hoc.

ANOVA of the N170 peak latencies showed a significant main effect of stimulus orientation [$F(1,15) = 161.5$, $p < .001$], which significantly interacted with stimulus type [$F(1,15) = 6.2$, $p < .025$]. Planned comparisons showed that for both natural and schematic faces, inversion delayed the N170 significantly [$t(15) = 10.821$, $p < .001$ and $t(15) = 3.626$, $p < .01$, respectively]. However, as evident in Table 2B, this difference was larger for natural than for schematic faces. The latencies were similar over the left and right hemispheres [$F(1,15) = 1.9$, $p = .185$]. The main effect of site was significant [$F(3,45) = 4.7$, $p < .01$], but because the differences between sites were all smaller

than the sampling rate, this effect was not investigated further. No other significant effects emerged in this analysis.

Discussion

The results of the Experiment 2 validated our prediction that inversion should reduce the amplitude of the N170 elicited by schematic faces but have little effect, or enhance the N170 elicited by natural faces. The natural face inversion effect observed in the present study replicates previous results reported by Rossion et al. (1999). These authors suggested that the enhancement in the N170 amplitude reflects the enhanced difficulty of the holistic processor to encode the inverted face. Indeed, the delay in the N170 peak latency elicited by inverted relative to upright faces is congruent with such a view. However, if this were the case, we should have found similar effects of inversion on the N170 amplitude elicited by schematic and by natural faces. Furthermore, the assumption that more difficult face encoding induces an enhancement of the N170 amplitude is also at odds with fMRI data showing that face inversion reduces the activity in the fusiform face areas relative to the activity elicited by upright faces (Kanwisher, Tong, & Nakayama, 1998). Indeed, finding opposite effects of inversion for schematic and natural faces can more likely be accommodated by our hypothesis that the larger N170 amplitude in response to inverted natural faces reflects its stronger associa-

tion with a face component analyzer in the lateral temporal gyri than with the holistic face processor in the fusiform.

Our model proposes that although schematic and natural faces probably activate similar neural networks in the extrastriate visual pathway, these mechanisms are not identical. Natural faces probably activate two perceptual modules, one specialized to detect and process physiognomic features in the visual field and the other specialized in the holistic processing of faces or forming a holistic face representation. Unlike natural faces, the components of schematic faces are not perceived as carrying physiognomic information out of the schematic face gestalt. Therefore, in contrast to natural faces, schematic faces may trigger the holistic processor but they do not trigger directly the analysis of the components. The latter mechanism may be activated only subsequent to the formation of the holistic face representation or, at the very least, following this process “in cascade” (McClelland, 1979). Consequently, adversely affecting the perception of the face gestalt (by inverting the stimuli) inhibited both holistic perception and the perception of individual components, and this inhibition reduced the amplitude of the N170. Such a pattern is also in agreement with the myriad of studies supporting the primacy of holistic processing of face stimuli.

GENERAL DISCUSSION

In the present study, we investigated several characteristics of the extrastriate structural face-encoding mechanisms by comparing the N170 elicited by faces differing in their naturalistic aspect, and by comparing the face inversion effects on the N170 elicited by photographs of natural faces and by schematically drawn faces. To the extent that the N170 is indeed associated with structural encoding, we found that this process distinguishes well between faces and nonfaces stimuli, but not among different face types that render the concept of a face clearly. However, we also found evidence that schematic faces and natural faces are not processed identically: Whereas inversion of natural faces enhanced the amplitude of the N170, inversion of schematic faces reduced its amplitude. For both face types, the latency of the N170 peak was significantly delayed by inversion. This pattern of results suggests that the face-specific structural encoder can be triggered by a variety of stimuli, if they include a canonical face configuration. However, the process of encoding the face information and forming a structural representation is probably different when the physiognomic value of the stimuli depends upon holistic configuration, as opposed to when the individual components can be associated with faces even when presented outside the face context.

Apart from the neurofunctional specificity issue, an enduring question in face perception research is whether

the formation of a mental representation of the face entails extraction of a holistic configuration directly from the visual input, or whether it is feature-based; that is, whether for faces, as for other complex visual stimuli, structural encoding is based on an initial analysis of parts (for a review of this debate see, e.g., Farah, 1990; Farah et al., 1995). In principle, because faces form a very homogeneous and stereotypical category, they lend themselves to holistic perception. Holistic processing is advantageous in face recognition because face components are fairly similar across individuals and, therefore, the distinction among faces is usually based on the interrelationships between the different face components and their overall configuration (e.g., Tanaka & Farah, 1993; Young, Hallowell, & Hay, 1987).⁷ Moreover, our memory for faces is usually based on nondecomposable holistic images, which, at the interface between perception and memory, are most efficiently addressed by integrated face representations.⁸ However, if the well-learned gestalt is altered and holistic perception is not possible, it is probably replaced by a more analytic, component-based process of structural encoding. This analytic process may inhibit, or at least delay the perception of the spatial relationships between face components and, therefore, may impede identification. Indeed, a major source of evidence supporting the holistic view is that faces are more difficult to recognize when presented upside down (i.e., compromising their familiar gestalt) than when presented in upright orientation (see reviews in Farah et al., 1995; Rhodes et al., 1993; Valentine, 1988).

Having to cope with different types of face input, the visual system must be able to encode the face configuration directly from the visual display, as well as by integrating face components encoded individually. These two processes may concur, triggered independently by relevant physiognomic input, and performed by different neural circuits to form a complex structural encoding mechanism in the extrastriate cortex. Furthermore, the two systems may interact, supporting or inhibiting each other. Hence, a regular face presented in upright orientation may activate the holistic perception process, which, in turn, may inhibit superfluous analysis of individual components. By contrast, corrupting the face gestalt, by presenting faces in an unusual orientation, may trigger (and facilitate) face component analysis, which should not depend on a particular configuration. In such a case, the integration of the components into a whole by the holistic processor will follow their independent analysis, assisted by the feature-detection system. This hypothesis may explain, for instance, why within the context of a regular face, individual face parts are distinguished faster than in isolation but, if the configuration of the face is altered, the parts are distinguished as fast as when presented in isolation (Tanaka & Farah, 1993; Sergent, 1984).⁹ The ability to process face information both holistically and piecewise

was implemented in a neural model initially based on scalp recordings (Bentin et al., 1996), and more recently corrected, improved, and elaborated on the basis intracranial recordings (McCarthy et al., 1999; for a detailed description and a review of supportive functional MRI data, see McCarthy, 1999). This model suggests that the neural mechanism dedicated to structural encoding of faces is a complex system composed of at least two separate subsystems. One, probably located in the lateral part of the middle fusiform gyrus, is responsible for processing the configuration of the face as a whole, and provides the integrated face representation that is necessary for face recognition and identification. The other, possibly distributed over posterior regions in the IT and middle temporal (MT) gyri and in the OTS, is responsible for detecting physiognomic information regardless of the spatial configuration of the items that carry it. The fusiform gyrus is farther away from the recording sites at the scalp than where the temporal gyri are, and the orientation of its cells forms an electrical dipole source that is almost perpendicular to the lateral aspects of the scalp. Hence, although both sources contribute to the modulation of the face-specific N170, the contribution of the posterior temporal face component processor should be considerably more conspicuous.

According to this hypothesis, the opposite effects of inversion on the N170 elicited by natural and schematic faces might be explained assuming that the holistic encoding system in the fusiform gyrus indeed inhibits the component analysis in the lateral temporal lobe face sites. Hence, if the face stimulus can trigger both the holistic and the component analysis (as it is true for natural faces), an experimental manipulation that interferes with holistic processing (like the face inversion procedure in the present experiment) should reduce the amount of inhibition of component analysis and, consequently, the amplitude of the N170 would be enhanced. However, if the “components” do not convey physiognomic information so that their analysis by a face-specific mechanism “depends” on the previous recognition of the face gestalt (as it is true for schematic faces), interfering with holistic perception cannot facilitate the activity of the component analysis and, therefore, the modulation of the N170 would only reflect the additional difficulty to encode the face, i.e., reduced amplitude and delayed latency.

The observed effects of the different factors on the N170 peak latency are nicely aligned with the above model. Although stimulus inversion delayed the N170 for both natural and schematic faces, this delay was twice as big (12 msec) for former stimuli than for the latter (6 msec). This interaction suggests that the inversion had a bigger effect on natural than on schematic faces. This may be because regardless of whether they are upright or inverted, the physiognomic value of schematic faces is based on their gestalt. Hence, inversion does inhibit the

processing of the gestalt but does not induce a change in processing strategy. On the other hand, natural faces are analyzed holistically if they are upright, but inversion requires a change in that strategy, encouraging the analysis of components. The significantly faster response to upright natural than to upright schematic faces in Experiment 2 suggests that the extraction of the face gestalt is easier if the face is more natural.

There is ample evidence indicating that the right hemisphere is specialized in global or holistic processing of objects and particularly faces, whereas the left hemisphere is specialized in part-based processing (e.g., Robertson & Delis, 1986; for a review, see Bradshaw & Nettleton, 1983; for specific reference to face perception see, e.g., Bruce, 1998; Rhodes et al., 1993; Corballis, 1988; Bradshaw & Sherlock, 1982; Gilbert & Bakan, 1973). Consequently, interfering with holistic processing should have had stronger effects in the right than in the left hemisphere. Our results confirmed this prediction only partially: The pronounced interhemispheric asymmetry with larger N170 amplitudes at the right than at the left hemisphere sites for all stimulus types suggests a predominant holistic process for all face categories. Interestingly, this asymmetry was reduced (in Experiment 1) for sketches relative to the other stimulus types. The relatively enhanced involvement of the left hemisphere in processing the sketches is in agreement with our suggestion that the face-specific structural encoding system was activated less efficiently by sketches than by other face types.¹⁰ Hence, the same factor that may have constrained the processing of schematic faces may also have constrained the processing of sketches, but in the opposite direction. Whereas schematic faces enhanced holistic processing by providing a clear face gestalt while obscuring the physiognomic value of face components, the sketches reduced holistic processing by blurring or veiling the face gestalt. Also, in Experiment 2, the N170 was larger over right than over left hemisphere sites. However, because the interaction between the inversion and hemisphere effects in Experiment 2 was not modulated by stimulus type (i.e., the three-way interaction was not significant), our interpretation of the interhemispheric asymmetry (particularly in Experiment 2) must be limited.

In conclusion, the present scalp-recorded N170 data converge with the previously published pattern of the intracranially recorded N200 data (for a comprehensive review see Allison et al., 1999; McCarthy et al., 1999; Puce et al., 1999), supporting a multifunctional organization of the structural encoding mechanism for faces. Apparently, distinct neural circuits are responsible for the encoding and initial analysis of face components and for holistic perception of a full-face configuration. The former is probably located in the lateral posterior temporal lobe, whereas the latter is located in the middle fusiform gyrus. Face perception usually entails predominantly holistic processing, which explains the right

hemisphere dominance in performance as well as in electrophysiological and neuroimaging data. However, both mechanisms can be triggered independently by relevant information.

METHODS

Participants

Right-handed participants were recruited among volunteer students from the Hebrew University. They received either experimental credit hours or payment for their participation. Thirty participants (13 males and 17 females) were tested in the first experiment and 16 participants (8 males and 8 females) were tested in the second experiment. Participants' ages ranged from 19 to 37 years (mean = 23.3 years). They all had normal or corrected-to-normal vision.

Stimuli

Stimuli were scanned using a desk scanner, processed with commercially available graphic software and con-

verted to 16 color grayscale 248 by 248 pixel images. There were eight stimulus categories, four face types (photographs of natural faces, paintings, face sketches, and schematic faces), two nonsense patterns (formed by scrambling the natural faces and the schematic faces), flowers, and butterflies (Figure 4). Paintings and sketches of faces were scanned from art books. Paintings were typically realistic Renaissance portraits, and face sketches did not portray details and usually lacked shading and texture. Schematic face parts were drawn by hand, scanned into the computer, and later combined into different schematic faces. Face photographs and paintings contained many shades of gray, and the sketches and the schematic faces were basically black on white. All face stimuli were front views.

Scrambled faces were created by scrambling the phases of the two-dimensional Fourier transform and then smoothing the uneven edges created by this process. The original and the scrambled stimuli were equally luminous following this procedure. Each stimulus category comprised 75 stimuli, except for the butterfly category which consisted of 30 stimuli. In Experiment 2, only natural faces, schematic faces, and flowers were

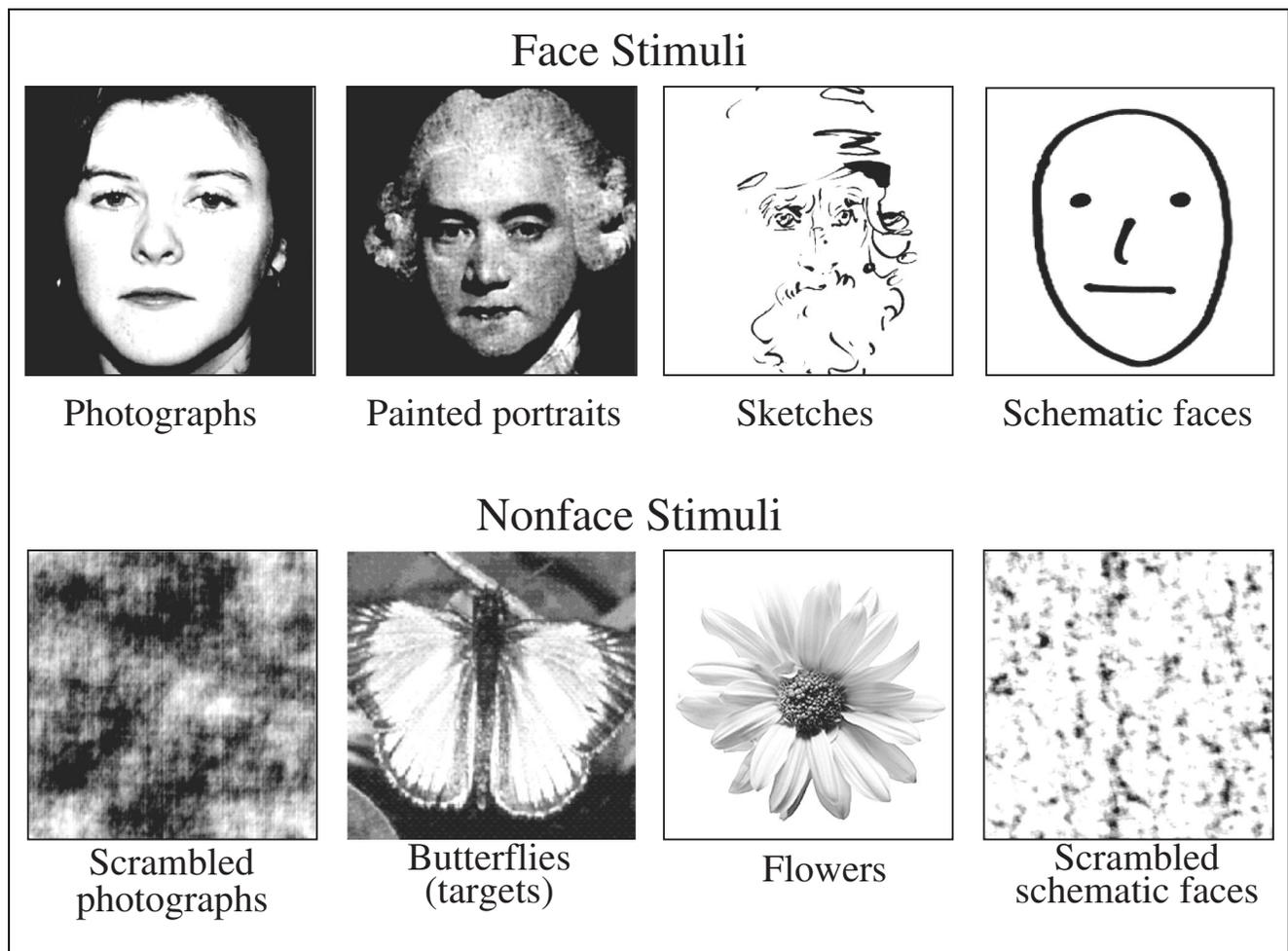


Figure 4. Example of the eight stimulus types used in the present study.

used. The faces were presented in both upright and upside down orientation. Upright stimuli were the same in Experiments 1 and 2.

All stimuli were presented for 350 msec with an average interstimulus interval (ISI) of 1.2 sec and viewed from a distance of 110 ± 10 cm subtending a viewing angle of approximately 6° . Stimulus timing was controlled by the MEL software system (Psychology Software Tools, Pittsburgh, PA).

Task

Participants were instructed to keep a mental count of stimuli belonging to one category: butterflies in the first experiment and flowers in the second. Hence, all of the experiment-relevant categories were, from the participants' point of view, nontargets. The targets were presented in approximately 5% (Experiment 1) and 6% (Experiment 2) of the trials. During each break and at the end of the experiment, participants reported the number of targets counted. A few participants who did not count correctly were excluded.

EEG Recording and Data Analysis

Continuous EEG was recorded from 48 scalp locations using tin electrodes attached to elastic electrode caps (Electrode Caps International, Ohio). A schematic illustration of the electrode locations and names is presented in Figure 1. A ground electrode was located on the forehead and a reference electrode was located on the tip of the nose. Two additional electrodes located at the outer canthus and the infraorbital region of the right eye measured vertical and horizontal electrooculograms (EOGs).

The EEG was sampled at 250 Hz by a DAP2400 Data Acquisition Processor (Microstar Laboratories), amplified 20,000-fold using a battery-operated HU-50/72 Isolated Bioelectric Amplifier (SA Instrumentation, San Diego) with a frequency bandpass of 0.1–30 Hz and stored on disk for off-line averaging. Electrode impedance was below 5 k Ω . A digital filter of 0.8–16 Hz was applied before analysis.

Codes synchronized to stimulus delivery were used to selectively average epochs associated with different stimulus categories. Epochs contaminated with saccades or eye blinks were excluded using root-mean-square EOG-values as criteria. Typically, epochs associated with 55–70 (never less than 50) out of 75 stimuli in each category were averaged.

The mean and peak amplitudes and peak latencies of specific ERP components within bounding intervals were computed for each category for each participant. Significance of the differences between these measures was tested using a repeated-measures ANOVA. Where appropriate, the more conservative Greenhouse–Geisser ad-

justed *df* values were used. ANOVAs were followed by univariate *F* contrasts.

Acknowledgments

This study was supported by grants from the US–Israel Binational Science Foundation and Israeli Foundations Trustees to S. Bentin. We thank Amir Katz for his help in preparing the stimuli and running Experiment 1, as well as Leon Deouell, Michael Fink, Ayelet Landau, Yulia Goland, David Carmel, Gil Raviv, Amir Raz, and Dan Tzur for their support. We are also very grateful to Jerry Levinson, Alex List, and Lynn Robertson for their comments on the manuscript.

Reprint requests should be sent to Shlomo Bentin, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel 91905, or via e-mail: msbentin@mssc.huji.ac.il.

Notes

1. Present address: University of California, Berkeley.
2. There was no reliable difference between the ERP elicited by scrambled natural faces and scrambled schematic faces. Therefore, these two categories were collapsed into one.
3. These were the sites at which the N170 was most conspicuous.
4. So were, in fact, all face categories, except for the sketches.
5. Note that this result also contradicts the view that a more difficult encoding process elicit N170 of a higher amplitude.
6. This assumption was, indeed, supported by a current study (Bentin, Mecklinger, Bosch, Sagiv, & von Cramon, 1999).
7. Of course, some faces have very distinctive components, which can be used as markers for recognition. Such markers are frequently exaggerated and used as distinctive features in caricatures (Rhodes, 1987).
8. Distinctive features (such as unusual components, birthmarks, etc.) may be stored in semantic memory as part of a person's face entry. Such information may have visual qualities that can be addressed directly during visual matching to facilitate identification.
9. Note that this account for the "face superiority effect" is identical to a common explanation of the "word superiority effect" in the identification of single letters. Indeed, our assumption about the coexistence of "whole-face" and "face-part" processors resonates to the "dual route" model of visual word processing. It is quite possible that identical principles govern all the content-specific perceptual modules in the visual system.
10. Recall that, relative to other face types, the reduction of the N170 amplitude elicited by sketches at right hemisphere sites was greater than at left hemisphere sites.

REFERENCES

- Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., & Spencer, D. D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology*, *71*, 821–825.
- Allison, T., McCarthy, G., Nobre, A. C., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex*, *5*, 544–554.
- Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception I.

- Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, *9*, 416–430.
- Bentin, S., Allison, T., Puce, A., Perez, A., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*, 551–565.
- Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, *17*, 35–54.
- Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: Evidence from developmental prosopagnosia. *NeuroReport*, *10*(4), 823–827.
- Bentin, S., & Mann, V. A. (1990). Masking and stimulus intensity effects on duplex perception: A confirmation of the dissociation between speech and non-speech modes. *Journal of Acoustical Society of America*, *88*, 64–74.
- Bentin, S., Mecklinger, A., Bosch, V., Sagiv, N., & von Cramon, D. Y. (1999). In the eyes of the beholder: Context-induced activation of face-specific perceptual mechanisms. Sixth Annual Meeting of the Cognitive Neuroscience Society, April 11–13, Washington, DC.
- Biderman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review*, *94*, 115–147.
- Bradshaw, J. L., & Nettleton, N. C. (1983). *Human cerebral asymmetry*. Englewood Cliffs, NJ: Prentice-Hall.
- Bradshaw, J. L., & Sherlock, D. (1982). Bugs and faces in the two visual fields: Task order, difficulty, practice, and analytic/holistic dichotomy. *Cortex*, *18*, 211–226.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Bruce, V., & Young, A. (1998). *In the eyes of the beholder. The science of face perception*. Oxford: Oxford University Press.
- Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychological Review*, *95*, 115–123.
- De Renzi, E. (1997). Prosopagnosia. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (pp. 245–256). New York: McGraw-Hill.
- De Renzi, E., Perani D., Carlesimo, G. A., Silveri, M. C., & Fazio, F. (1994). Prosopagnosia can be associated with damage confined to the right hemisphere. An MRI and PET study and a review of the literature. *Neuropsychologia*, *32*, 893–902.
- Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, *3*, 1–8.
- Farah, M. J. (1990). *Visual agnosia*. Cambridge: MIT Press.
- Farah, M. J. (1996). Is face recognition ‘special’? Evidence from neuropsychology. *Behavioral Brain Research*, *76*, 181–189.
- 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.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*, 482–498.
- George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research*, *4*, 65–76.
- Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355–362.
- Haxby, J. V., Grady, C. L., Horwitz, B., Salerno, J., Ungerleider, L. G., Mishkin, M., & Schapiro, M. B. (1993). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. In B. Gulyas, D. Ottoson, & P. E. Roland (Eds.), *Functional organization of the human visual cortex* (pp. 329–340). Oxford: Pergamon.
- Ishai, A., Ungerleider, L. G., Martin, A., Schouten, J. L., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences, U.S.A.*, *96*, 9379–9384.
- Jeffreys, D. A., & Tukmachi, E. S. A. (1992). The vertex-positive scalp potential evoked by faces and by objects. *Experimental Brain Research*, *91*, 340–350.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.
- Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*, 1–11.
- Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science*, *243*, 489–494.
- Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representation in the primate. *Cerebral Cortex*, *5*, 270–288.
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, *5*, 552–563.
- Logothetis, N. K., & Scheinberg, D. L. (1996). Visual object recognition. *Annual Reviews of Neuroscience*, *19*, 577–621.
- Luh, K. E., Rueckert, L. M., & Levy, J. (1991). Perceptual asymmetries for free viewing of several types of chimeric stimuli. *Brain and Cognition*, *16*, 83–103.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Marr, D., & Nishihara, H. K. (1978). Representation and recognition of 3-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, *200*, 269–294.
- McCarthy, G. (1999). Physiological studies of face processing in humans. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (chap. 28, pp. 393–409). Cambridge: MIT Press.
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception: II. Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*, 605–610.
- McClelland, J. L. (1979). On the time-relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*, 555–604.
- Moses, Y., Adini, Y., & Ullman, S. (1994). Face recognition: The problem of compensating for illumination changes. *Proceedings of the European Conference on Computer Vision*, 286–296.
- Moses, Y., Edelman, S., & Ullman, S. (1993). *Generalization to novel images in upright and inverted faces*. The Weizmann Institute of Science, Technical Report CS93-14.
- Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*, *74*, 1192–1199.
- Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception III. Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, *9*, 445–458.
- Rhodes, G. (1995). Face recognition and configural coding. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition*. London: Rutledge.
- Rhodes, G., Brake, S., & Atkinson, A. (1993). What’s lost in inverted faces? *Cognition*, *47*, 25–57.
- Robertson, L. C., & Delis, D. C. (1986). ‘Part-whole’ processing in unilateral brain-damaged patients: Dysfunction of hierarchical organization. *Neuropsychologia*, *24*, 363–370.
- Rossion, B., Delvenne, J. -F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guerit, J. -M. (1999). Spa-

- tio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology*, *50*, 173–189.
- Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., & Crommenlink, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of specific processes in the human brain. *NeuroReport*, *11*, 69–74.
- Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology*, *75*, 221–242.
- Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. *Brain*, *115*, 15–36.
- Tanaka, J. W., & Farah, M. J. (1991). Second-order relational properties and the inversion-effect: Testing a theory of face perception. *Perception and Psychophysics*, *50*, 367–372.
- Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.
- Ullman, S. (1996). *High-level vision: Object recognition and visual cognition*. Cambridge: MIT Press.
- Valentine, T. (1988). Upside down faces: A review of the effect of inversion upon face recognition. *British Journal of Experimental Psychology*, *87*, 116–124.
- Whalen, D. H., & Liberman, A. M. (1987). Speech perception takes precedence over nonspeech perception. *Science*, *237*, 169–171.
- Young, A. W., Hellawell, D. J., & Hay, D. C. (1987). Configural information in face perception. *Perception*, *16*, 747–759.