6

Expert Face Processing: Specialization and Constraints

Adrian Schwaninger

Claus-Christian Carbon

Helmut Leder

6.1 Introduction

Face processing in adults is the product of innate mechanisms, and is also based on years of experience. There is no doubt that face processing is a human skill at which most adults are real experts. In the present chapter we review theories and hypotheses concerning adults' face processing skills, as well as what information and processes these are based on. Moreover, we discuss how the high specialization is attained at the cost of being susceptible to specific conditions.

Expertise, according to the American Heritage Dictionary is given when a person shows a high degree of skill in or knowledge of a certain subject. This definition implies that an expert is a high-grade *specialist*. Expertise does not have to be accessible in an explicit way, because an expert does not have to know all the facts of his expertise. The skill humans show in identifying faces is astonishing. According to Bahrick, Bahrick, and Wittlinger (1975) adults are able to recognize familiar faces with an accuracy of 90 per cent or more, even when some of those faces have not been seen for fifty years. Moreover, faces are a class of objects which encourage a special kind of categorization. According to the logic of Roger Brown's seminal paper "How shall a thing be called" (Brown, 1958), the level of the object name reflects the entry point of the recognition process. When asked to name pictures of faces spontaneously, humans produce the concrete names of the persons shown.

Classifying objects at this kind of subordinate level is typical of experts (Tanaka & Taylor, 1991). Expertise can not only be recognized by the frequency of subordinate-level classifications but also by the speed of word generation (Tanaka, 2001a): Adults identified faces as fast at the subordinate level (the name of the person) as at the basic level (e.g., "human"). This is clear evidence for a level of expertise.

To understand the development of face processing from childhood to adulthood better, we review the characteristics of information processing used by adults. First, we consider different types of pictorial information contained in faces. Then we review the holistic hypothesis as well as the schema hypothesis. This is followed by a discussion of important characteristics of adult face recognition, namely the sensitivity to configural information and the specialization in upright faces. Subsequently, the component configural hypothesis is discussed. Finally, we present a model for familiar and unfamiliar face recognition which allows the integration of several important aspects of a fully developed face processing system.

6.2 Information Contained in Faces

Faces are complex three-dimensional surfaces of the front side of the human head. Psychophysical studies using computer graphics have distinguished surface-based shape information from superficial properties such as color and texture (e.g., Hill, Schyns, & Akamatsu, 1997; Troje & Bülthoff, 1996).

Another commonly used distinction is based more on phenomenology. The term *component* information (or componential, piecemeal, featural information) has been used to refer to separable local elements, which are perceived as distinct parts of the whole such as the eyes, mouth, nose or chin (Carey & Diamond, 1977; Sergent, 1984). Components describe the basic primitives in faces, and the number of dimensions on which all components can differ provides the basis for all human faces being unique. A second type of information has been referred to as *configural* or *relational*. According to Bruce (1988), the term configural information refers to the "spatial interrelationship of facial features" (p. 38), i.e., features which come about from spatial arrangements, such as eye-distances, nose-mouth-distance. Distinctiveness correlates positively with the recognizability of faces, and Leder and Bruce (1998) revealed that component as well as configural information contribute to the distinctiveness of faces. Configural information was defined further by Diamond and Carey (1986). They used the term *first-order relational information* for the basic arrangement of the parts and *second-order relational information* to refer to specific metric relations between features.

The term *holistic* has been used to describe representations that store a face as an unparsed perceptual whole without specifying the parts explicitly. It has been operationalized in whole-to-part-superiorities (see section 6.3.1) and refers to properties and features when the face is processed as a Gestalt and not parsed into components (Farah, Tanaka, & Drain, 1995; Tanaka & Farah, 1993). A simple two-dimensional analogy for a holistic face representation would be a bitmap that only specifies the color values of points without providing any information about which

points belong to the mouth or the eyes. Although the bitmap contains eyes and a mouth, it does not represent them explicitly[1].

These different types of information contained in faces are related to hypotheses about adult face processing which are discussed next.

6.3 Mechanisms of Face Processing in Adults

In order to explain the mechanisms used in adult face processing, several hypotheses have been proposed. According to the holistic hypothesis, adults process faces as unparsed perceptual wholes. The schema hypothesis assumes that the ability to process faces improves over many years and is attained at the expense of flexibility. This specialization could be related to adults' high sensitivity to configural information. Since faces are usually seen upright, it is not surprising that orientation is a critical variable for a face processing system that develops from years of experience. According to the component configural hypothesis, the processing of configural information is much more impaired by changes of orientation than the processing of component information. Why this might be the case is explained by the integrative model we propose after discussing each of these hypotheses in more detail.

6.3.1 Holistic Hypothesis

According to the holistic hypothesis, upright faces are stored as unparsed perceptual wholes in which individual parts (components) are not explicitly represented (Farah et al., 1995; Tanaka & Farah, 1993). Several empirical findings have been interpreted in favor of this view. For example, Tanaka and Farah (1993) reasoned that if face recognition relies on parsed representations, then a component (e.g., a nose) presented in isolation should be easy to recognize. In contrast, if faces are represented as unparsed perceptual wholes (i.e., holistically) then a part of a face presented in isolation should be much more difficult to recognize. In their experiments, participants were trained to recognize upright faces, each of which had a different pair of eyes, nose, and mouth. In the test phase, images of faces were presented in pairs. Each pair of faces differed only in the shape of one part of the face. In one test condition, two facial parts were presented in isolation. The subjects had to judge which of the two parts belonged to a face familiar from the training phase. In the whole face condition, the parts were embedded in the facial context. For example, one face contained the original nose and the other contained a different nose. The participants had to judge which of them was the face familiar to them from the training phase. Parts presented in isolation were more difficult to identify than whole faces. In contrast, when participants were trained to recognize inverted faces,

---

[1] Note that this definition is different from the concept of holistic processing, which is understood in terms of overall similarity relations (see Chapter 4).

scrambled faces, and houses no advantage of presenting the parts in their context was found. The authors interpreted this result in favor of the holistic hypothesis and proposed that face recognition relies mainly on holistic representations while the recognition of objects is based much more on part-based representations. Whereas encoding and matching parts are assumed to be relatively orientation-invariant (Biederman, 1987), holistic processing is thought to be very sensitive to orientation (see also Biederman & Kalocsai, 1997; Farah et al., 1995).

The results of a study conducted by Tanaka and Sengco (1997) provide further support for the holistic hypothesis, although their concept of holistic is slightly different. Instead of assuming that faces are processed as unparsed perceptual wholes, the authors reasoned that if *both component and configural information are combined* into a single holistic representation, changes in configural information should affect the recognition of facial parts (component information). This was precisely what was found in their first experiment: After training with upright faces, the subjects recognized components (eyes, nose and mouth) better in the unaltered facial context than in the context of a face in which the configural information had been changed by manipulating the distance between the eyes. If holistic processing is hampered by inversion and if face recognition relies much more on holistic representations than object recognition does, then a similar configural manipulation should have no effect on the recognition of parts of inverted faces or objects such as houses. This indeed was the case. The authors showed that configural manipulations did not affect the recognition of isolated parts when faces were presented upside-down nor did they do so when upright houses were used in the training and test conditions. (For faces, the alteration of configural information was accomplished by increasing the distance between the eyes, and for houses by manipulating the distance between the windows.) Thus, altering the configural information only affects the recognition of parts in the case of upright faces. This finding favors the view that in normal (upright) face processing the component and configural information is combined into a single holistic representation and that this holistic processing is disrupted by inversion. In paragraph 6.3.3 we review further evidence for the importance of configural information in face processing.

Another line of evidence for this view is derived from a study carried out by Rhodes, Brake, and Atkinson (1993). These authors used (coarse) digitized versions of full-face photos in a recognition memory paradigm. Configural alterations, which were induced by altering the internal spacing of the eyes and mouth, were more difficult to recognize when faces were inverted. Interestingly, when the eyes or mouth were replaced with those of another face, effects of inversion were even more detrimental to recognition performance! Rhodes et al. (1993) concluded that either the component changes also affected the configural information or that the assumption that component processing is relatively unaffected by inversion is incorrect. The authors reasoned that if the replacement of components also resulted in a configural change and this caused the decrease in performance for inverted faces, then this effect of inversion should disappear when the components are presented alone. The results of their Experiment 2 favored this interpretation. In line with the results of Tanaka and Sengco (1997), the findings of Rhodes et al. (1993) are consistent with the view that in normal (upright) face processing component and

configural information is combined into a single holistic face representation and that this holistic processing is impaired by inversion. Note that this concept of holistic processing differs slightly from the original definition of Tanaka and Farah (1993) and Farah et al. (1995). In the original view, holistic processing just means that parts are not represented explicitly. In contrast, holistic processing according to the results of Tanaka and Sengco (1997) and Rhodes et al. (1993) would imply that component and configural information are first encoded separately and then integrated into a holistic representation.

According to Farah et al. (1995) the holistic hypothesis also predicts that effects of inversion can be eliminated if participants are induced to represent faces in terms of their parts. Indeed, these authors found that inversion had the expected negative effect on the recognition of faces that were studied normally, while this impairment disappeared when faces were studied as parts (head outline, eyes, nose, and mouth presented simultaneously in different boxes). However, while the authors admit that it is possible to represent faces in terms of their components, they stress that performance is impaired by inversion because faces are usually represented holistically, i.e., parts are not represented explicitly.

An alternative definition of holistic processing of faces was tested by Macho and Leder (1998). Holistic processing could be achieved by an interactive feature processing in which the processing of one feature depends in general on the quality of another feature. In a similarity decision task using faces which systematically varied on two or three dimensions to target faces, they did not find evidence for this kind of interactive processing.

6.3.2 Schema Hypothesis

Goldstein and Chance (1980) have suggested another hypothesis. According to their view, the ability to process faces (i.e., the face schema) improves with exposure to them. These authors suggest that this improvement is attained at the expense of flexibility. Therefore, because faces are usually seen upright, it follows that recognition performance should improve with age, but performance with unusual stimuli such as inverted faces should decline through development. Their predictions have been supported by studies that investigated the development of face recognition (for reviews see Carey, 1992; Ellis, 1992; Johnston & Ellis, 1995). A study by Diamond and Carey (1986) provides another line of evidence in favor of the schema hypothesis. These authors used faces and dog profiles as stimuli. They found that the performance of novices was affected by inversion when tested with human faces but not when dog profiles had to be recognized. In contrast, there was an effect of inversion on dog experts'(dog show judges and breeders with an average of 31 years experience with dogs' appearance) recognition of dog profiles which was comparable to the observed effect of inversion on their recognition of human faces! This result was also found when bird and dog experts were shown bird and dog pictures, and

their N170-ERP[2] component was compared: Approximately 164 ms after presentation, objects of expertise (dogs for dog experts; birds for bird experts) can be dissociated from objects from lower expertise categories (Tanaka, 2001b). Thus, based on the schema hypothesis, one would assume that this vast amount of object exposure has resulted in an expert-specific schema that is orientation sensitive because all the exemplars have usually been encountered in the upright position.

Goldstein and Chance did not elaborate on how a schema is used. Nevertheless, the linking element between the results discussed in the previous paragraph might be the processing of configural information in faces: The use of this special class of information could be an essential element of a holistic representation as proposed by Tanaka and Sengco (1997) and might also develop with age as well as the face schema.

6.3.3 Sensitivity to Configuration

Adult face recognition is characterized by a high sensitivity to configural information. For example, Haig (1984) showed for unfamiliar faces that configural alterations produced by changing the distance between facial features are sometimes detected at the visual acuity threshold level. Hosie, Ellis, and Haig (1988) found similar results using familiar faces. Kemp, McManus, and Pigott (1990) used two-tone images and found that the high sensitivity to configural information is reduced in negative or inverted images. While these studies were primarily concerned with the perceptual level, Bruce, Doyle, Dench, and Burton (1991) revealed a specialization for processing configural information at the level of memory processes. When tested, participants had to decide whether faces and houses were identical to the ones presented in a previous block or whether they had been altered configurally. Although the alterations were smaller for faces than for houses, participants were more sensitive in detecting them. Similar to the result of Kemp et al. (1990), this effect diminished when the stimuli were inverted. Leder and Bruce (2000) tested directly whether individual configural elements are represented in memory explicitly. They used a set of 8 faces, each of which differed only in a distinctive local configural feature such as a lowered mouth or a smaller eye-distance. In the test phase, they presented the whole face or the distinctive features in isolation or embedded into an empty head shape. Participants were surprisingly efficient at recognizing faces from the isolated configural elements. Moreover, all the experiments in Leder and Bruce (2000) revealed that the processing of configural information was particularly disrupted by inversion. The authors conclude that it is the reliance on configuration that is essential for adult's expertise at processing upright faces.

Thus, based on the review of recent studies, better processing of configural information seems to be applicable for adults rather than children. This is in accordance with findings that the limits of face processing are often accompanied by

---

[2] The N170 is a posterior negativity of the event-related potential (ERP) which reflects an early stage of face processing

a disruption of configural rather than other sorts of information. In the next paragraph we describe three effects which are known to be particularly disruptive to adult face processing.

6.3.4 Testing for Limits: The Advantage of Being Upright

The remarkable ability of recognizing faces reliably is highly dependent on orientation. We have already shown how the holistic hypothesis and the use of configural information by adults suggest that orientation is a critical variable. Moreover, to process facial information reliably, a large amount of expertise is required (for a review see Carey, 1992; Chapter 4). Through years of practice, the face recognition system becomes more specialized but at the same time more limited to processing the upright orientation (schema hypothesis). In the following section we review three effects that illustrate this specialization in upright faces: the face inversion effect, the Thatcher illusion and the face composite illusion.

In order to investigate whether inversion particularly affects the recognition of faces, Yin (1969) used a forced-choice recognition paradigm with pictures of human faces, airplanes, houses, and stick figures of men in motion as stimuli. In one condition the stimuli were learnt and tested in the upright orientation. Upright faces were recognized better than all the other upright stimuli but were stronger affected by inversion. In another condition the stimuli were learnt in the upright orientation and then tested in the inverted orientation. Generally, when the stimuli had to be recognized in the upside-down position, error rates increased for all stimuli. The interesting finding was that this increase was disproportionately high for faces when compared with the other objects. Whilst faces were recognized best in the upright test condition, performance for inverted faces dropped below the recognition levels of the other object classes. This finding, namely that upside-down faces are disproportionately more difficult to recognize than other inverted objects, has been referred to as the *face inversion effect*. Subsequent replications of Yin's study have refined the initial methodology by comparing faces with stimuli that were equivalent in terms of familiarity, complexity, and psychosexual importance (e.g., Ellis, 1975; Goldstein & Chance, 1981; Scapinello & Yarmey, 1970). Valentine (1988) presented a comprehensive summary of studies investigating the face inversion effect. The review of recent results on holistic and configural processing suggests that the disruption of configural information explains most of the effects of the inversion of faces (Leder & Bruce, 2000).

Another impressive demonstration for the orientation-sensitive nature of face processing comes from a study carried out by Thompson (1980). In a photograph of Margaret Thatcher, he rotated the eyes and mouth within the facial context, which resulted in a grotesque facial expression (see Figure 1 for a demonstration). Interestingly, this strange expression is not perceived when the face is turned upside-down, but is immediately apparent when the face is turned upright. This effect has been referred to as the *Thatcher illusion*. It is clear that this manipulation of the orientation of components alters the form of the eyes and mouth to the point of grotesqueness.

Figure 1. Thatcher illusion. Both inverted pictures look more or less "normal". But when turned upright, the thatcherized version is seen to be highly grotesque. Try it!

Inverting the eyes within the facial context clearly changes the spatial relationship of the parts. Indeed, this alteration has been considered by some authors to produce a change in the configural information (e.g., Bartlett & Searcy, 1993; Diamond & Carey, 1986; Stevenage, 1995).

Young, Hellawell, and Hay (1987) discovered another interesting effect (see Figure 2 for an illustration). They created composite faces by combining the top and bottom half of different faces. If the two halves were aligned and presented upright, a new face resembling each of the two originals seemed to emerge. This made it very difficult to identify the persons from either half. If the top and bottom halves were misaligned horizontally, then the two halves did not fuse spontaneously to create a new face, and the constituent halves remained identifiable. However, when these stimuli were inverted, the constituent halves of the aligned and misaligned displays were equally identifiable. Furthermore, the subjects were significantly faster at naming the constituent halves in inverted composites than in upright composites.

Figure 2. Aligned and misaligned halves of different identities (here two of the authors). When upright (as above), a new identity seems to emerge from the aligned composites (left), which makes it more difficult to extract the original identities. This does not occur for the misaligned composite face (right). When viewed upside-down, the original identities can be extracted easily from both pictures.

Young et al. (1987) have argued that it is the new configuration in the composite face, which makes the identification of the parts difficult. Thus again we have evidence that an effect specific for upright faces might be due to the use of configural information in upright faces and the disruption of this in upside-down faces.

Concerning the developmental course, Cashon and Cohen (2001) showed that 7-month-old infants process composites from outer and inner features as one face. This may be taken as evidence for a kind of configural processing, which is in accordance with Tanaka, Kay, Grinnell, Stansfield, and Szechter (1998) who found that 6-year-olds showed the same whole-to-part superiority effects as adults. Carey and Diamond (1994) also found that adult-like composite effects emerge at the age of 6 while configural processing (indicated by inversion effects) develops continually until adulthood. Recently, Mondloch, Le Grand, and Maurer (2002) showed that configural processing develops later than featural or component processing and that it may still develop after the age of ten (see also Chapters 4 and 5).

6.3.5 Component Configural Hypothesis

While numerous studies have been presented which stress the importance of configural processing, it is not yet clear how different features are combined to form a representation of faces in memory. In the present paragraph we discuss a hypothesis in which two modes of processing are assumed: the component configural hypothesis. According to this hypothesis, component and configural information is processed separately, and configural processing is much more affected by changes of orientation than the processing of components. There is a large amount of evidence in favor of this view. The first demonstration of a differential effect of inversion on the processing of component and configural information was provided by Sergent (1984). She used pairs of faces where either the eyes or facial contour (change of component information) or the internal spacing of components (change of configural information) were mismatched. A multidimensional scaling technique for the analysis of dissimilarity judgments, and regression analyses on reaction times revealed that configural and component information were used for upright faces. In contrast, there was no evidence that subjects made use of configural information when faces were inverted. It should be noted, however, that Sergent (1984) used schematic faces which could make it difficult to generalize this result to the processing of real faces. However, similar results were found by Searcy and Bartlett (1996), who used color photographs of faces in which configural changes had been induced by moving the eyes and mouth up or down, and manipulation of the component information had been achieved by changing the color of the pupils and teeth or by shortening and elongating the teeth. In line with Sergent's (1984) results, a grotesqueness-rating task and a simultaneous paired-comparison task provided further evidence for the view that inversion is particularly disruptive to the processing of configural information. Leder and Bruce (1998) manipulated the distinctiveness of either components or configural features directly and showed how both make upright faces easier to recognize. When faces were presented upside-

down, the effects of distinctiveness based on configural features vanished in nearly all conditions.

Another demonstration of the differential effects of orientation on the processing of component and configural information was provided by Schwaninger and Mast (1999). They used a sequential same-different matching task and found that the detection of component changes (eyes and mouth replaced) was relatively invariant to planar rotations. In contrast, rotation had a detrimental effect upon the detection of configural changes that were induced by increasing the distance between the eyes and the eyes and mouth (Figure 3). Interestingly, the effect of rotation on configural processing was nonlinear; most errors were found at intermediate angles of rotation between upright and inverted orientations, i.e., at 90° – 120°. Similarly, Murray, Yong, and Rhodes (2000) found a discontinuity in the function relating bizarreness to a rotation of between 90° and 120° which was found for Thatcher faces and faces in which configural changes were induced by changing the relative position of the eyes and mouth. The bizarreness ratings of unaltered or component-distorted faces (teeth blackened and eyes whitened) showed only a linear trend. Leder and Bruce (2000, Experiment 5) compared directly whether configurations are also accessible when, at the same time, components vary from face to face: the isolated configurations, though composed of components which they shared with other faces, were recognized and showed inversion effects. To show directly that configural information is processed differentially in upright as compared to inverted faces, Leder, Candrian, Huber, and Bruce (2001) used a sequential comparison task. Participants saw two faces sequentially which differed in interocular eye-distance only. The task was to decide for each pair of faces which face had the larger interocular eye-distance. The judgments were more accurate when the faces were presented upright, and the decrement in accuracy in the inverted condition was independent of the size of the surrounding context (e.g., whether the nose or the mouth and nose were added).

One possible caveat of the studies that investigated the processing of component and configural information by replacing or altering facial features is that this type of manipulation often changes the holistic aspects of the face and is difficult to carry out selectively.
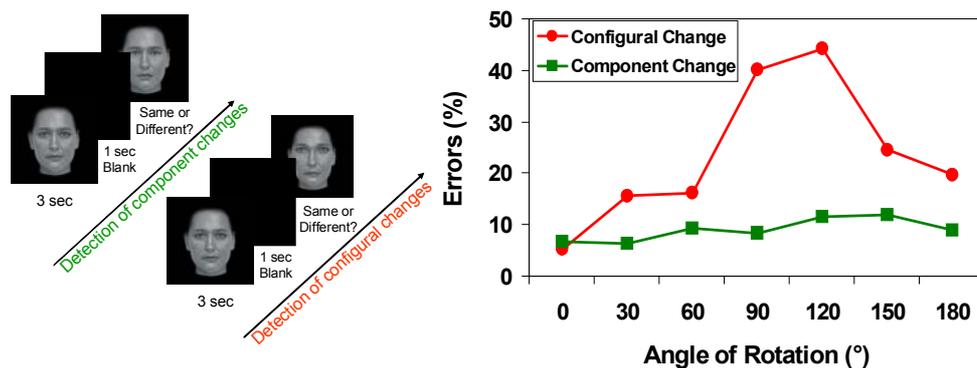

Figure 3. Study by Schwaninger and Mast (1999). Left: The detection of component and configural changes was tested using a sequential same-different matching task in separate experiments. Right: Whereas the identification of component changes was almost unaffected by rotation, the detection of configural changes was strongly impaired.

For example, replacing the nose (component change) can change the distance between the contours of the nose and the mouth and thus alter the configural information (Leder & Bruce, 1998; 2000). The same applies to configural changes when they are carried out by altering the relative position of the components. For example, moving the eyes apart (configural change) can lead to an increase in size of the bridge of the nose, i.e., a component change (see Leder et al., 2001).

Problems like these can be avoided by using scrambling and blurring procedures to reduce configural and component information separately (e.g., Collishaw & Hole, 2000; Davidoff & Donnelly, 1990; Sergent, 1985). Recently, Schwaninger, Lobmaier, and Collishaw (2002) used scrambling and blurring techniques in an old-new recognition paradigm. Their experiments extend previous research by ensuring that scrambling and blurring effectively eliminate configural and component information separately. Furthermore, in contrast to previous studies, Schwaninger et al. (2002) used the same faces in separate experiments on unfamiliar and familiar face recognition to avoid potential confounds with familiarity (Figure 4).

In Experiment 1, unfamiliar face recognition was studied. In the first condition it was shown that previously learnt intact faces could be recognized even when they were scrambled into constituent parts. This result challenges the assumption of purely holistic processing according to Farah et al. (1995) and suggests that facial features or components are encoded and stored explicitly. In a second condition, the blur level was determined that made the scrambled versions impossible to recognize. This blur level was then applied to whole faces in order to create configural versions that by definition did not contain local featural information. These configural versions of previously learnt intact faces could be recognized reliably. This result suggests that separate representations exist for component and configural information. Familiar face recognition was investigated in Experiment 2 by running the same conditions with participants who knew the target faces (all distractor faces were unfamiliar to the participants).
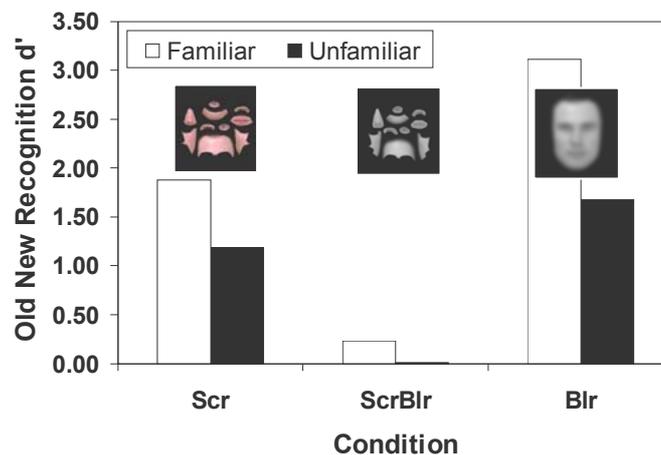
Figure 4. Recognition performance in unfamiliar and familiar face recognition across three different conditions at test. Scr: scrambled, ScrBlr: scrambled and blurred, Blr: blurred. (Adapted from Schwaninger, Lobmaier, & Collishaw, 2002)

Component and configural recognition was better when the faces were familiar, but there was no qualitative shift in processing strategy since there was no interaction between familiarity and condition (Figure 4).

In sum, there is converging evidence in favor of the view that separate representations for component and configural information exist which are relevant for the recognition of familiar and unfamiliar faces. Whereas component information is not very orientation-sensitive, configural information is difficult to recover when faces are rotated.

6.4 An Integrative View of Face Recognition

Everyday object recognition is often a matter of discriminating between quite heterogeneous object *classes* that differ with regard to their global shape, parts and other distinctive features such as color or texture. In contrast, face recognition relies on the discrimination of *exemplars* of a very homogenous category. All faces share the same basic parts in the same basic arrangement. In each face the eyes are above the nose which is located above the mouth. Therefore, reliable face recognition relies on the detection of subtle featural and configural differences, which needs years of experience. Since faces are usually seen upright, this learning must become more and more restricted to the upright orientation. A strong dependency on orientation is the consequence for objects that are usually perceived in one specific orientation. Since effects of rotation and inversion are much more detrimental for faces than for basic level object recognition, a certain type of information must be more relevant for faces. According to certain authors, expert face recognition is characterized by holistic processing (e.g., Biederman & Kalocsai, 1997; Farah et al., 1995; Tanaka & Farah, 1993). Farah et al. (1995) answer the question "Why is face recognition so orientation sensitive?" in the following way: "Face perception is holistic and the perception of holistically represented complex patterns is orientation sensitive." (p. 633). According to Rock (1973, 1974, 1988), rotated faces overtax an orientation normalization mechanism, which makes it impossible to match them against stored upright memory representations. Rotated faces can only be processed by their components, and configural information is hard to recover. This would explain why effects of rotation are much smaller for component as opposed to configural changes (Leder & Bruce, 1998, 2000; Schwaninger & Mast, 1999). At the same time, these results challenge a purely holistic view of face processing which assumes that explicit representations of facial parts do not exist. The recent results of Schwaninger et al. (2002) offer further evidence against such a purely holistic view. They revealed that facial components and configural information are encoded and stored explicitly, both in unfamiliar and familiar face recognition, when faces are upright.

In order to integrate the different hypotheses outlined in this chapter, we propose the model depicted in Figure 5. All pictorial aspects of a face are contained in the pictorial metric input representation which is presumably correlated with activation in primary visual areas. Based on years of expertise, neural networks are trained to extract specific information in order to activate component and configural

representations in the ventral visual stream. The output of these representations converges towards the same identification units. These units are holistic in the sense that they integrate component and configural information. Note that this concept of holistic differs from the original definition of Tanaka and Farah (1993) and Farah et al. (1995). In their view, holistic means that parts are not represented explicitly. In contrast, according to our model, holistic processing implies that component and configural information are encoded separately first and then integrated into a holistic representation. Our concept of holistic is fully compatible with the results from Schwaninger et al. (2002) and Leder et al. (2001) who showed that featural and configural information is encoded explicitly. Moreover, our integrative definition of holistic is consistent with the results of Tanaka and Sengco (1997) and Rhodes et al. (1993) which imply that in normal (upright) face processing, component and configural information is combined into a single holistic face representation. Finally, our concept of holistic can be related to holistic processing in terms of overall similarity relations (see Chapter 4). A holistic similarity decision would be based on a linear or nonlinear integration of component and configural information, a prerequisite of a developed face processing system. An analytical similarity decision would mean that only component information is used to judge the similarity of faces.

Adult face recognition is characterized by the processing of configural information and by the fact that faces are quite hard to recognize when they are rotated substantially from the upright position. In the model this can be explained in the following way: When faces are rotated, the pictorial information in the input representation is changed remarkably. As a consequence, the component and configural representations which have been learnt based on exposure to upright faces, cannot be activated well enough to allow reliable recognition.
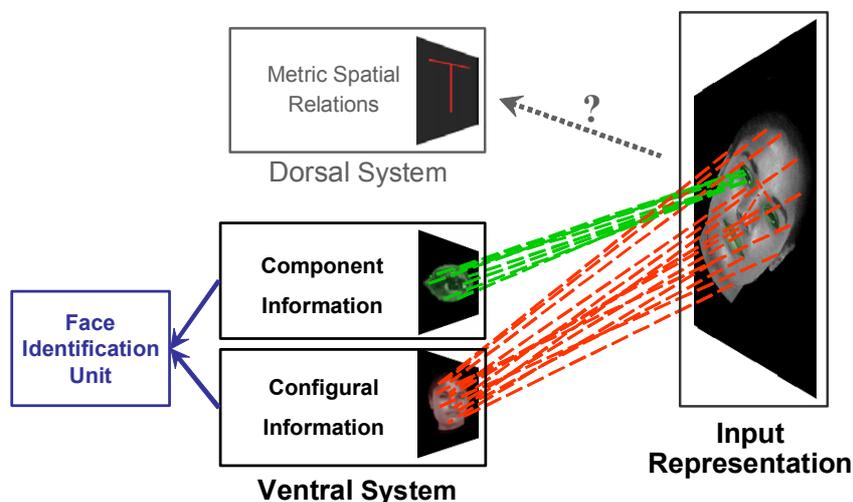
Figure 5. Integrative model of face processing. Facial information is encoded in a metric input representation that contains all the features we perceive in faces. Information of local features and relations between them is extracted in order to activate component and configural representations in the ventral stream. The outputs of these representations converge towards the same face identification units. Whether dorsal processing is relevant for processing metric spatial relations in faces such as the eye-mouth or the inter-eye distance remains to be investigated.

Rotated faces overtax orientation normalization mechanisms so that they have to be processed by their components (Rock, 1973, 1974, 1988). As pointed out by Valentine and Bruce (1988), this implies that information about the spatial relationship of components (configural information) is hard to recover. Consequently, the processing of configural information is much more affected by rotation or inversion than the processing of component information. Since face recognition relies heavily on processing configurations, the inversion effect is in disproportion to that of other objects (Yin, 1969). We believe that this is the deeper answer to the question "Why is face recognition so orientation sensitive?"

Our model also offers an explanation for the Thatcher illusion and the composite face illusion. Thatcherizing a face, i.e., inverting the eyes and mouth within an upright face, results in a strange activation pattern of component and configural representations. Consequently, the face looks very bizarre. When a thatcherized face is inverted, the activation of configural representations is strongly impaired due to the limitation in capacity of an orientation normalization mechanism. Consequently, the strange activation pattern of configural representations is reduced and the bizarre perception vanishes. Moreover, in an inverted Thatcher face the components themselves are in the correct orientation which results in a relatively normal activation of component representations. Consequently, inverted Thatcher faces appear relatively normal (Rock, 1988). Finally, the composite face illusion can be explained by similar reasoning. Aligned upright face composites contain new configural information resulting in a new perceived identity. Inverting the aligned composites reduces the availability of configural information and it is easier to access the two different face identification units based on the component information alone.

In short, the model we propose allows the integration of the component configural hypothesis and holistic aspects of face processing. It explains striking perceptual effects such as the Thatcher illusion and the composite face illusion. Most importantly, it provides an integrative basis for understanding special characteristics of adult face recognition such as the specialization in upright faces and the sensitivity to configural information.

Acknowledgments

References

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*, 54-75.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*(3)*,* 281-316.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*(2)*,* 115-147.

Biederman, I., & Kalocsai, P. (1997). Neurocomputational bases of object and face recognition. *Philosophical Transactions of the Royal Society of London, B, 352*, 1203-1219.

Brown, R. (1958). How shall a thing be called? Psychological Review, 65, 14-21.

Bruce, V. (1988). *Recognising faces*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition, 38*, 109-144.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London, 335*, 95-103.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science, 195*, 312-314.

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1*, 253-274.

Cashon, C., & Cohen, L. (2001). Do 7-month-old infants independent features of facial configurations? *Infant and child development, 10*, 83-92.

Collishaw, S. M., & Hole, G. J. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893-910.

Davidoff, J., & Donnelly, N. (1990). Object superiority: A comparison of complete and part probes, *Acta Psychologica, 73* 1990 225–243.

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

Ellis, A. W. (1992). Cognitive mechanisms of face processing. In V. Bruce, A. Cowey, A. W. Ellis, & D. P. Perrett (Eds.), *Processing the facial image* (pp. 113-119). New York, NY: Clarendon Press.

Ellis, H. (1975). Recognizing faces. *British Journal of Psychology*, 66, 409-426.

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

Goldstein, A. G., & Chance, J. E. (1980). Memory for faces and schema theory. *Journal of Psychology, 105*, 47-59.

Goldstein, A. G., & Chance, J. E. (1981). Laboratory studies of face recognition. In G. M. Davies, H. D. Ellis, & J. W. Sheperd (Eds.), *Perceiving and remembering faces*. London, England: Academic Press.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception, 13(5),* 505-12.

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62,* 201-222.

Hosie, J. A., Ellis, H. D., & Haig, N. D. (1988). The effect of feature displacement on the perception of well-known faces. *Perception*, 17, 461-474.

Johnston, R.A., & Ellis, H.D. (1995). The development of face recognition. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition, explorations in face space* (pp. 1-23). London, New York: Routledge.

Kemp, R., McManus, C., & Pigott, T. (1990). Sensitivity to the displacement of facial features in negative and inverted images. *Perception, 19*, 531-543.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology, 51A*, 449-473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology, 53A,* 513-536.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73-83.

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(5),* 1486-1500.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural processing develops more slowly than feature face processing. *Perception, 31*, 553-566.

Murray, J. E., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11*, 498-502.

Rhodes, G., Brake, S., & Atkinson, A.P. (1993). What's lost in inverted faces? *Cognition, 47*, 25-57.

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

Rock, I. (1974). The perception of disoriented figures. *Scientific American, 230*, 78-85.

Rock, I. (1988). On Thompson's inverted-face phenomenon (Research Note). *Perception, 17*, 815-817.

Scapinello, F. F., & Yarmey, A. D. (1970). The role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Science, 21*, 329-331.

Schwaninger, A., & Mast, F. (1999). Why is face recognition so orientation-sensitive? Psychophysical evidence for an integrative model. *Perception (Suppl.), 28*, 116.

Schwaninger, A., Lobmaier, J., & Collishaw, S. M. (2002). Role of featural and configural information in familiar and unfamiliar face recognition. *Lecture Notes in Computer Science, 2525,* 643-650.

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

Sergent, J. (1984). An investigation into component and configurational processes underlying face recognition. *British Journal of Psychology , 75,* 221-242.

Sergent, J. (1985). Influence of task and input factors on hemispheric involvement in face processing. *Journal of Experimental Psychology: Human Perception and Performance, 11(6)*, 846-61.

Stevenage, S. V. (1995). Expertise and the caricature advantage. In T. Valentine (Ed.), *Cognitive and computational aspects of face recognition, explorations in face space* (pp. 24-46). London, England: Routledge.

Tanaka, J. W. (2001a). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General, 130*, 534-543.

Tanaka, J. W. (2001b). A neural basis for expert object recognition. *Psychological Science, 12*, 43-47.

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

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

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23*, 457-482.

Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition, 5*, 479-496.

Thompson, P. (1980). Margaret Thatcher -- A new illusion. *Perception*, *9*, 483-484.

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36(12),* 1761-1771.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471-491.

Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory and Cognition, 16,* 556-566.

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

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