The many faces of configural processing

Daphne Maurer, Richard Le Grand and Catherine J. Mondloch

Adults’ expertise in recognizing faces has been attributed to configural processing. We distinguish three types of configural processing: detecting the first-order relations that define faces (i.e., two eyes above a nose and mouth), holistic processing (glueing the features together into a gestalt), and processing second-order relations (i.e., the spacing among features). We provide evidence for their separability based on behavioral marker tasks, their sensitivity to experimental manipulations, and their patterns of development. We note that inversion affects each type of configural processing, not just sensitivity to second-order relations, and we review evidence on whether configural processing is unique to faces.

Adults are experts at recognizing faces: they can recognize thousands of individuals at a glance, even at a distance, in poor lighting, with a new hairdo, after 10 years of aging, or when the face is seen from a novel viewpoint — unless the person is upside down. Relative to upright faces, recognizing inverted faces is surprisingly poor, with the decrement far larger than it is for shoes or houses. Adults’ expert skill in recognizing faces is attributed to configural processing — processing not just the shapes of individual features but also the relations among them. This ability develops from years of experience in differentiating among upright faces. In fact, since a classic paper by Yin [1], the inversion effect — much poorer accuracy and longer reaction times when faces are upside down — has been taken as diagnostic of configural processing.

The term ‘configural processing’ has been used to refer to any phenomenon that involves perceiving relations among the features of a stimulus such as a face. It is contrasted with ‘featural processing’, which is also called ‘componential processing’, ‘piecemeal processing’, and ‘analytic processing’. Configural processing of faces can be divided into three types: (1) sensitivity to first-order relations — seeing a stimulus as a face because its features are arranged with two eyes above a nose, which is above a mouth; (2) holistic processing — glueing together the features into a gestalt; and (3) sensitivity to second-order relations — perceiving the distances among features. However, there is no consensus about terminology, with some authors restricting the term ‘configural processing’ to one of these types of relational processing and others applying the term indiscriminately to all three types or distinguishing between only two of the three types.

Sensitivity to first-order relations

Adults have a remarkable ability to detect faces based on first-order relations, even in the absence of normal facial features, at least when the stimuli are upright (see Fig. 1) [2,3]. Indeed, faces can play a special role in capturing attention. Newborns orient preferentially towards stimuli that have face-like first-order relations [4,5] (but there are alternative explanations [6]), and patients with visual extinction are more likely to detect a face presented in the neglected hemifield than they are to detect a scrambled face, a name, or a meaningless shape [7].

Detecting face-like first-order relations is facilitated by the fact that all faces share the same basic configuration; as a result, normalized facial representations can be superimposed and the resulting stimulus remains recognizably face-like [8]. Studies using event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI) have identified neural correlates of detecting a face. The event-related negative potential called the N170 has a larger amplitude of the N170 is for faces than for many other stimuli, including hands, houses, and cars [9,10]. fMRI activation in regions of the ventral occipitotemporal cortex, the inferior occipital gyrus, and the lateral fusiform gyrus, i.e., the fusiform face area (FFA), is larger for faces than for a variety of non-face objects, including cars, houses, and furniture [11–13]. Although isolated eyes can evoke the N170 (e.g., [9], but see [12]), these neural correlates seem to be related to perceiving a face rather than to stimulus characteristics per se. For example, when adults view ambiguous stimuli, fMRI activity in the FFA is higher when the background encourages perception of the stimulus as a face than when it encourages perception of a vase [14] (see also [15,16]).

Although there is no consensus on what type of processing these neural correlates reflect (see [9,14,17]), they are affected much more by manipulations that influence sensitivity to first-order relations than by manipulations that affect sensitivity to second-order relations. This difference provides evidence for the separability of these two aspects of configural processing. For example, when gray-scale photographs of faces are inverted, adults have difficulty recognizing the identity of the face based on the spacing of features [18,19] but continue to detect first-order relations (i.e., perceive the stimulus as a face). Inverting gray-scale faces causes little [3] (but see [60]), or no [11,20], change in the magnitude of the FFA response (although it increases activity in object regions [11,20]) and a delay, and sometimes an increase, in the amplitude of the N170 [9,10,21,22].
By contrast, inverting Mooney faces disrupts the ability to detect a face (see Fig. 1b) and produces a significant drop in fMRI activation in the FFA [3]. As would be expected if these neural correlates reflect perception of the first-order relations of a face, they are unaffected by the subject’s familiarity with the particular face or by its repetition [23–25] (but see [26] for evidence of fMRI effects in a nearby region of bilateral fusiform cortex and [27] for evidence of an effect of typicality on the N170), and at least some of them can be elicited by animal faces [28,29].

Holistic processing
When adults detect the first-order relations of a face, they tend to process the stimulus as a gestalt, making it harder to process individual features. The most convincing demonstration is the ‘composite face effect’. Subjects are slower and less accurate in recognizing the top half of one face presented in a composite with the bottom half of another face when the composite is upright and fused than when the composite is inverted or the two halves are offset laterally – manipulations that disrupt holistic processing (Fig. 2a) [30,31]. This phenomenon demonstrates that when upright faces are processed, the internal features are so strongly integrated that it becomes difficult to parse the face into isolated features, at least at short exposures that prevent feature-by-feature comparisons [31]. The composite face effect occurs even with upright faces that are presented as negatives [32], even though differences among individuals in the spacing of features are hard to detect in negatives, a result suggesting separate effects of negation on holistic processing and on sensitivity to second-order relations.

Holistic processing of faces has also been demonstrated by showing that subjects are about 10% more accurate in recognizing the identity of a feature (e.g. Larry’s nose) when it is presented in the context of the whole face (e.g. Larry’s face with Larry’s nose versus Bob’s nose) rather than as an isolated feature (Larry’s nose versus Bob’s nose) (the ‘part–whole recognition effect’) [33,34]. No such benefit occurs for scrambled faces or houses. The benefit is reduced if the spacing of other features is altered in the original face and, like the composite face effect, is lost if the face is inverted. It is also reduced by simultaneous processing of upright flanker faces but not inverted or fractured flanker faces [35] – as would be expected if interference occurs only when the flanker task engages holistic processing.

The strength of holistic processing of upright faces is evident in several additional phenomena.

Fig. 1. Sensitivity to first-order relations. (a) An Archimbaldo painting (The Vegetable Gardener, inverted). (b) An inverted two-tone Mooney face. These images do not appear face-like when they are upside-down. When viewed upright, they appear face-like because the features are arranged to form the first-order relations of a face. Those first-order relations are detected even when the features are formed from vegetables (a) and when they are constructed from only patches of intense light and shadow and require closure (b). Archimbaldo painting provided courtesy of Museo Civico ‘ala Ponzoné’, Cremona. Mooney face provided courtesy of Scania de Schonen.

Fig. 2. Manipulations that demonstrate holistic processing. (a) The composite face effect. (b) Al Gore/Bill Clinton composite. (c) Overlapped transparent faces. Because upright faces engage holistic processing, it is difficult to extract information about individual features such as the fact that in (a) the top halves of both faces are the same, or that in (b) the internal features of both faces are the same. Holistic processing makes it easy to see two overlapped images as two alternating whole faces when they are nearly upright (c, top). Inverting these stimuli improves adults’ ability to compare individual features (a,b) and to see two sets of features simultaneously in overlapped faces (c, bottom). Gore/Clinton composite adapted with permission from Ref. [37]. Overlapped faces reproduced with permission from Ref. [39].
Even with simple circles containing three curved lines, adults are 50\% slower to detect a deviation in the curvature of one line if the lines are arranged to look like a smiling or a frowning face rather than arranged arbitrarily or as an inverted face [36]. Holistic processing seems to occur not only among the internal features but also between the internal features and external contour, making it difficult to recognize that the internal features of two faces are the same when they are presented in different external contours (Fig. 2b) [37] (see also [30], Exp. 4). The spacing of the internal features also affects the perceived shape of a surrounding contour, creating illusions of stretching and rounding, with stronger effects if the face is upright rather than inverted and if the correct features define the first-order relations [38]. Finally, it is the strength and seeming automaticity of holistic processing that allows adults to disambiguate two overlapping transparent faces into separate whole faces that alternate – as long as the faces are nearly upright (tilted 45° right or left) [39] (Fig. 2c).

**Sensitivity to second-order relations**

Because all faces share the same first-order relations, recognition of individual faces requires the encoding of information about subtle variations in the shape or spacing of the features. Second-order relations refer to the spatial distances among internal features [8] (e.g. the distance between the eyes). Adults can detect variations in these distances as small as one minute of visual angle, a value close to the limits of acuity [40].

To tap processing of second-order relations, a set of faces can be created that differ from one another only in the spacing of individual features (Fig. 3a) [18,19,41–45]. This manipulation has minimal effect, if any, on information about local features, provided that the spatial variation is not so extreme as to create a new facial feature (e.g. a broad nose between widely spaced eyes). To tap featural processing, a set of faces can be created that differ from one another in local information by changing the shape (Fig. 3b) [18,19,42], color [44], or luminance [44,45] of the features. Such manipulations have little or no effect on second-order relations provided that the size and location of individual features are kept similar across faces. Inverting such stimuli reduces accuracy and increases reaction times when adults discriminate faces that differ in second-order relations much more than when they discriminate faces that differ in featural information [18,19,41,44], especially for less salient regions within the face [45]. Similarly, adults’ ratings of the distinctiveness of faces with distortions in second-order relations (exaggerated spacing between the eyes) drop significantly following inversion, whereas their ratings of faces with featural distortions (e.g. darkened eyebrows) do not change [43] (see also Box 1). Taken together, these findings indicate that separate mechanisms are involved in second-order relational versus featural processing of individual faces. The difficulty in processing second-order relations in inverted faces is likely to occur at the level of perceptual encoding because it is of similar magnitude when faces are presented simultaneously as when they have to be remembered for up to 10 seconds [18].

Another technique for isolating second-order relational processing is to blur the stimuli to remove fine-detailed information about facial features [46,47]. Adults are able to recognize the identity of blurred faces with reasonable accuracy [47,48] but are severely impaired if the faces are simultaneously blurred and inverted – presumably because blurring removes featural information and inversion disrupts sensitivity to second-order relations [49]. Adding heavy random noise to face stimuli also eliminates useful featural information. Following training with upright versions of such stimuli, adults demonstrate categorical perception in making judgments about similarity and likeness. They do not show categorical perception if trained with isolated features or inverted faces, presumably because second-order relational information was not available [50]. Taken together, these results explain why inverted faces can be recognized at levels exceeding chance when featural information has not been removed by a manipulation such as blurring or superimposing noise.

Negation does not disrupt holistic processing but it impairs the ability to distinguish between faces that differ in the spacing among features to the same extent as inversion disrupts discrimination of positive faces [51] (but see Exp. 2 in [51]). Negation also makes faces more difficult to recognize because of

---

Fig. 3. Separability of featural versus second-order relational processing. (a) Faces that differ in the spacing among features. (b) Faces that differ in the shape of individual features (eyes and mouth). When viewed upright, adults readily distinguish among the faces within each set. Inverting these stimuli severely impairs the ability to distinguish faces from (a) that differ in second-order relations, but has little effect on the ability to distinguish faces from (b) that differ in featural information [19].
Is configural processing unique to faces?

Although many researchers argue that configural processing is unique to faces, others suggest that it is used with other categories of objects, particularly if, like faces, the objects are homogeneous and the viewer has developed expertise in distinguishing individual members at the subordinate level. Most of the evidence is based on the inversion effect, which, as noted previously, is an inadequate diagnostic. For example, dog judges and breeders, but not novices, are less accurate in recognizing which of two dogs they saw previously when the photographs are inverted – at least when the test sets include breeds for which the dog handlers are expert [8]; see [55] for similar results for handwriting. This result could reflect greater orientation-specific recognition of features, greater holistic processing and/or greater skill in using second-order relations to individuate dog bodies.

The best evidence for configural processing of non-face objects comes from training studies with photographs of ‘greebles’ (Fig. 4). After extensive training, adults show some evidence for all three types of configural processing. They demonstrate a face-like N170 (a marker of sensitivity to first-order relations) for upright greebles [56], much as dog experts and bird experts do when shown stimuli from their category of expertise [57]. As with faces, the N170 is delayed for inverted greebles [56]. After training they also show some evidence of holistic processing [58,59]: a slowing of response or decrease in accuracy when an individual’s part or top half has to be recognized in a configuration made novel by moving other parts or fusing it with the bottom half of another greeble. Their accuracy is also impaired more than that of novices by negation, as would be expected.
Although during infancy there is evidence for featural processing and all three types of configural processing (e.g. [a–d]), they later mature at different rates. By age 6 years (the youngest age tested), the magnitude of the composite face effect [e] and the part–whole recognition effect [f] – both measures of holistic processing – are the same as in adults. At this age, featural processing is almost adultlike, but second-order relational processing is only slightly above chance and much worse than in adults [g,h]. Perhaps as a result, children have difficulty matching faces when they differ in point of view, clothing or lighting [i,j]. Data from adolescents treated for bilateral congenital cataracts indicate that early visual deprivation spares the development of featural processing but permanently impairs the later development of sensitivity to second-order relations [k,l]. Similar data on the development of adultlike sensitivity to first-order relations are not available, but studies of prosopagnosia (impairment in face recognition) indicate that it can be spared despite severe deficits in other types of face processing [m,n].

References
f Tanaka, J.W. et al. (1998) Face recognition in young children: when the whole is greater than the sum of its parts. Visual Cogn. 5, 479–496
h Mondloch, C. et al. (2002) Configural face processing develops more slowly than featural face processing. Perception 31, 553–566

Questions for future research
• To what extent do the different types of configural face processing reflect distinct versus overlapping neural mechanisms? Do they vary in lateralization?
• To what extent are the different types of configural face processing hierarchical? For example, does holistic processing follow and depend upon detection of first-order facial relations – both developmentally and within each instance of perceiving a face?
• What are the contributions of each type of configural processing to deficits in face perception such as prosopagnosia? Can differences between patients in the pattern of deficits be explained by variability in which type of processing is damaged?
• What spatial frequencies are most informative for the different types of configural face processing? To which spatial frequencies do adults attend?
• Is the application of the different types of configural processing influenced by familiarity?
• Does the other-race effect (better and faster recognition of the age, sex, and identity of own-race faces than other-race faces) arise from better sensitivity to the features that distinguish faces of one’s own race and/or from better sensitivity to their second-order relations?
• What is the developmental trajectory for sensitivity to first-order relations?
• How does holistic processing of internal features resemble holistic processing between the internal and external features in, for example, timing, sequence, development and neural instantiation?
• Are second-order relations detected and/or encoded with reference to an average or prototype?
among the features. However, none of the existing data rules out the possibility that the three types of configural processing operate largely in parallel or that, under some conditions, a seemingly higher level (e.g. sensitivity to second-order relations) can operate expertly in the absence of processing at the other levels (e.g. holistic processing).

References