Developing a Brain
Specialized for Face
Perception: A Converging
Methods Approach

Received 21 February 2000; Accepted 13 November 2001

ABSTRACT: Adults are normally very quick and accurate at recognizing facial identity. This skill has been explained by two opposing views as being due to the existence of an innate cortical “face module” or to the natural consequence of adults’ extensive experience with faces. Neither of these views puts particular importance on studying development of face-processing skills, as in one view the module simply comes on-line and in the other view development is equated with adult learning. In this article, we present evidence from a variety of methodologies to argue for an “interactive specialization” view. In this view, orienting tendencies present early in life ensure faces are a frequent visual input to higher cortical areas in the ventral visual pathway. In this way, cortical specialization for face processing is an emergent product of the interaction of factors both intrinsic and extrinsic to the developing child. © 2002 Wiley Periodicals, Inc. Dev Psychobiol 40: 200–212, 2002. DOI 10.1002/dev.10027

Keywords: face processing; infancy; neuroimaging; cortical specialization; prosopagnosia

INTRODUCTION

Our sophisticated ability to perceive and analyze information from the faces of our fellow humans underlies much of our social behavior. In just a glance, we can recognize a single face from among the hundreds stored in our memories as well as detect its age, sex, emotional state, and direction of attention. These abilities provide the foundations for successful social group collaboration and the societies that have led to our success as a species. Given the extreme importance of faces, it is not surprising that some investigators have proposed that there is an innate “social brain” with pathways and circuits genetically prespecified for processing social information (e.g., Baron-Cohen et al., 1999). Reports that certain regions of the adult cortex are selectively activated by faces (reviewed in Grelotti, Gauthier, & Schultz, this issue) and that newborn infants preferentially orient to faces (e.g., Johnson, Dziurawiec, Ellis, & Morton, 1991) have been interpreted as evidence in support of this view. Other investigators have pointed out that, over the course of our lives, we have extensive and continuous exposure to faces. They have argued that this intensive visual “training” with faces results in a high level of expertise, but that this process is no different from acquiring expertise in discriminating and recognizing individual examples from other visual

Correspondence to: M. de Haan

© 2002 Wiley Periodicals, Inc.
categories (e.g., Diamond & Carey, 1986). Reports that “face-specific” brain tissue can be activated by other highly trained object categories (e.g., by birds in expert bird watchers; Gauthier, Skudlarski, Gore, & Anderson, 2000) have been taken as evidence in support of this view. In this article, we present evidence from developmental, animal, electrophysiological, neuroimaging, neuropsychological, and computational studies to argue for a third “interactive specialization” view (Johnson, 2000). In this view, cortical specialization for face processing is an emergent product of the interaction of factors both intrinsic and extrinsic to the developing child.

THE TWO-PROCESS THEORY
OF DEVELOPMENT OF FACE PROCESSING

The framework that has motivated research in our lab over the past decade is based on the two-process theory of the development of face recognition initially advanced by Johnson and Morton (1991; Morton & Johnson 1991) and updated by Johnson and de Haan (2001). In the original theory (Johnson & Morton, 1991), two distinct brain systems were proposed to underlie development of face processing in infancy: (a) “Conspec,” a system operating from birth that functions to bias the newborn to orient toward faces and (b) “Conlern,” a system sensitive to the effects of experience through passive exposure. In this model, Conspec is mediated by primitive, possibly largely subcortical, circuits whereas Conlern is mediated by developing cortical circuits in the ventral visual pathway. The purpose of Conspec is to bias the input to the still-plastic cortical circuits underlying Conlern, providing the first step towards the eventual emergence of the specialized circuits for face processing observed in adults.

In the initial formulation of the theory (Johnson & Morton, 1991), Conlern begins to influence behavior at 6 to 8 weeks of age and allows emergence of the ability to recognize the identity of individual faces. However, subsequent studies demonstrated that even newborn infants show evidence of recognizing the facial identity (Bushnell, Sai, & Mullin, 1989; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995). Taking these findings into account, Johnson and de Haan (2001) revised the original theory and proposed that, prior to the specialization of cortical circuits for faces, these stimuli can be processed in the same general-purpose way as other visual patterns. This provides some basis for the ability to recognize a number of individual faces from early in life. This limited ability is then supplement-
ed with the emergence of Conlern at 6 to 8 weeks of age, when more adultlike processing of faces begins to emerge. Our discussion focuses mainly on development of this system when infants are first learning about faces over the first year of life (for review of development during childhood, see de Haan, 2001).

METHODS

Comparative Studies

Face recognition is an important aspect of life not only for humans but also for many other species. Both face and species recognition have been frequently studied in animals, but only rarely from a developmental perspective. This approach is useful both because research with animals allows experimental manipulations not possible with humans and because a comparative approach provides insight into the similarities or differences in adaptation by different species to the common requirement of early recognition of conspecifics.

One area in which early recognition abilities have been studied is filial imprinting in precocial birds. Imprinting is the process by which young, precocial birds, such as domestic chicks, recognize and develop a social attachment for the first conspicuous object that they see after hatching. In the laboratory, newly hatched domestic chicks will imprint onto a variety of objects such as moving, colored balls and cylinders. Even after only a few hours of exposure to such a stimulus, chicks will come to prefer it to any other object. Thus, in the absence of a mother hen, the learning is relatively unconstrained: Virtually any conspicuous moving object larger than a matchbox will do.

Studies using autoradiographic, biochemical, lesion, and electrophysiological techniques have established that a particular localized region of the chick forebrain, known as the Intermediate and Medial part of the Hyperstriatum Ventrale (IMHV), is crucially involved in visual imprinting (for recent reviews, see Horn, 1998). For example, bilateral lesions to IMHV placed before or after training severely impair preference for the trained object in subsequent choice tests, but do not affect several other types of visual or learning tasks (Johnson & Horn, 1986, 1987; McCabe, Cipolla-Neto, Horn, & Bateson, 1982). Further, similarly sized lesions placed elsewhere in the chick forebrain do not result in significant impairments of imprinting preference (Johnson & Horn, 1987; McCabe et al., 1982).

Although the avian forebrain lacks the layered structure of mammalian cortex (but see Karten &
Shimizu, 1989), the relation of the forebrain to subcortical structures is similar, following a basic higher vertebrate brain design (Ebbesson 1980; Nauta & Karten, 1970). The evidence from a variety of vertebrate species supports the suggestion that the IMHV is a site of plasticity, and not the location of in-built, automatic types of behavior (Ewert, 1987; MacPhail, 1982). Figure 1 illustrates the location of IMHV within the chick brain. The area occupies about 5% of total forebrain volume. Its main inputs come from visual projection areas (hyperstriatum accessorium and the optic tectum), and some of its projections go to regions thought to be involved in motor control (such as the archistriatum). Thus, the area is well placed to integrate visual inputs and motor outputs. While regions of the avian forebrain may have analogous functions to the mammalian cerebral cortex (Horn, 1985), cytoarchitectonic studies of IMHV have revealed that it is much simpler in structure. In contrast to the six-layered structure with many cell types found in the mammalian cerebral cortex, there is no clear laminar structure of IMHV, and only four distinctive types of cells have been classified to date (Tombol, Csilag, & Stewart, 1988). A connectionist model based on characteristics of the cytoarchitectonics of IMHV suggests that the microcircuitry found in IMHV makes it an especially powerful mechanism for acquiring certain types of representations, such as translation invariant object-based representations (O’Reilly & Johnson, 1994).

In the laboratory, a wide range of objects, such as red boxes and blue balls, are as effective for imprinting as more naturalistic stimuli such as a moving stuffed hen. In the wild, however, precocial birds such as chicks invariably imprint on their mother and not on other passing objects. What mechanisms ensure that the plasticity in the chick brain is constrained to encode information about conspecifics and not about the characteristics of other objects present in its early visual environment?

An answer to this question arose from a series of experiments in which striking, stimulus-dependent effects of IMHV lesions were found (Horn & McCabe, 1984). While groups of chicks trained on an artificial stimulus such as a red box are severely impaired by IMHV lesions placed either before or after training, groups of chicks trained on a stuffed hen are only mildly impaired. Other neurophysiological manipulations also demonstrate differences between the hen-trained and box-trained birds. For example, administration of the noradrenergic neurotoxin DSP4 results in a severe impairment of preference in birds trained on the red box, but only a mild impairment in birds trained on the stuffed hen (Davies, Horn, & McCabe, 1985). In contrast, levels of plasma testosterone correlate with preference for the stuffed hen, but not preference for the red box (Bolhuis, McCabe, & Horn, 1986).

These results support an earlier suggestion (Hinde, 1962) that naturalistic objects such as hens may be more effective at eliciting attention in young chicks than are other objects. To investigate this possibility further, Johnson and Horn (1988) conducted a series of experiments in which completely dark-reared chicks were presented with a choice between an intact stuffed hen and a variety of test stimuli created from cutting up and jumbling the pelt of a stuffed hen. The results of these experiments forced the conclusion that chicks have a spontaneous tendency to attend toward characteristics of the hen’s head and neck region. While this untrained preference seems to be specific to the correct arrangement of features of the face/head, it is not specific to the species. For example, the head of a duck is as attractive as that of a hen (Johnson & Horn, 1988).

**FIGURE 1** Outline drawing of the chick brain. The brain is viewed from the side and is about 2 cm from front to back (left). IMHV = the Intermediate and Medial part of the Hyperstriatum Ventrale.
These findings and others led to the proposal that there are two independent neural systems that underlie filial preference in the chick (Horn, 1985; Johnson, Bolhuis, & Horn, 1985). The first system is a specific predisposition for the young chick to orient toward objects resembling conspecifics. This system appears to be specifically tuned to the correct spatial arrangement of elements of the head and neck region (Johnson & Horn, 1988), but not to the color or size (see Johnson & Bolhuis, 1991). While the stimulus configuration triggering the predisposition is not species- or genus-specific, it is sufficient to pick out the mother hen from other objects the chick is likely to see in the first few days after hatching. The second system acquires information about the objects to which the young chick attends, and is subserved by the brain region IMHV. In the natural environment, the first system ensures that the second system acquires information about the particular individual mother hen close by. Biochemical, electrophysiological, and lesion evidence all support the conclusion that the two systems have largely independent neural substrates (for review, see Horn, 1998). For example, selective lesions to IMHV impair preferences acquired through exposure to an object, but do not impair the predisposition (Johnson & Horn, 1986).

**Visual Tracking**

The experiments described above demonstrate that the development of recognition of conspecifics in the chick appears to involve two neuroanatomically distinct processes. It was this discovery that led to the formulation of the two-process theory of development of human face recognition previously described. The proposition that there may be a mechanism present in human neonates that causes them to orient to faces was controversial, as it was widely believed at the time that infants could not recognize faces until they were at least a few months of age. Indeed, the results of most experiments supported this view (reviewed in Johnson & Morton, 1991), with the exception of one study (Goren, Sarty, & Wu, 1975) reporting that newborn human infants around 10 min old turn their heads and eyes further to track a moving facelike pattern than a “scrambled” face pattern. Due to the controversial nature of this result, Johnson et al. (1991) attempted to replicate it, with some changes to improve the methodology (see also Maurer & Young, 1983). One change was to expand the set of stimuli to include (a) a pattern composed of the configuration of high-contrast areas that compose a face, but without the details of the features of a face. This pattern might be effective if the newborn’s tracking is triggered by the appropriate spatial arrangement of high-contrast “blobs” (somewhat similar to a defocused image of a face) and (b) a pattern that had the same features, but in the wrong arrangement. The results of this and other experiments (see Johnson & Morton, 1991; Simion, Valenza, Umiltà, & Dalla Barba, 1998; Valenza, Simion, Cassia, & Umiltà, 1996) all replicated the original findings and extended them to indicate that there is no consistent difference in response to a detailed schematic face compared to a facelike arrangement of three blobs. These results indicate that newborns’ preferential tracking is not specific to the fine details of facial features, but is specific to the arrangement of elements that compose a face (rather than simply a unidimensional psychophysical variable such as spatial frequency; Morton, 1993; Morton, Johnson, & Maurer, 1990).

Controversy remains about the basis of the newborn’s bias. For example, Simion, Macchi Cassia, Turati, and Valenza (2001) recently demonstrated with simple geometric patterns that newborns preferentially orient to patterns with a higher density of elements in the upper visual field, and Acerra, Burnod, & de Schonen (in press) used neural-network simulations to argue that the newborn’s bias to orient to faces is due to the details of receptive field properties of cells in V1. Such studies raise the possibility that the preferential orienting to facelike stimuli is a consequence of less specific constraints, or gestalt principles, in visual processing. A combination of less specific biases could have the same behavioral consequence as a more specific “face” bias in ensuring that, in a natural environment involving faces and other objects, newborns orient toward faces. However, these “sensory” explanations are not necessarily, as is often implied, simpler explanations of newborns’ preferential tracking of facelike patterns. First, a description of the preferential tracking of the Conspec pattern on the basis of a combination of known sensory preferences in the newborns could still be relatively complex [e.g., patterns with a high density of elements in the upper visual field (Simion et al., 2001), a strong vertical component (Farroni, Valenza, Simion, & Umiltà, 2000), curved outer contour, etc.; see Johnson & Morton, 1991, p. 135, Box 6.1]. A laboratory environment with psychophysically prepared geometric visual patterns (e.g., the upright and inverted “T” shapes used by Simion et al., 2001) can potentially tell us more about the representations underlying behavior, but not about the ethology of the infant in its natural environment. Second, the crude representation of a face in newborns assumed by Johnson and Morton (1991) may not, in fact, be any more complex to construct than the apparently less
specific biases proposed. Recent neural-network simulations have shown that sensitivity to the basic spatial arrangement of high-contrast areas that compose a face could arise as a result of spontaneous (prenatal) activity in certain neural networks (Bednar & Miikkulainen, 2000). Whatever the exact description of the basis of the newborn’s orienting tendencies, none contradict our proposal that these reflect more primitive biases that could provide input to developing cortical areas rather than the existence of an innate cortical “face module.”

Some investigators have in fact interpreted newborns’ sensitivity to faces as support for the existence of innate, face-specific cortical areas. However, there are several indirect sources of evidence that suggest newborns’ preferential orienting toward faces is mediated primarily by subcortical circuitry. First, the preferential tracking of faces declines sharply between 4 and 6 weeks after birth (Johnson et al., 1991, Experiment 3). This is similar to the time course of the decline of other newborn reflexlike behaviors that is thought to occur due to inhibition by developing cortical circuits (e.g., Muir, Clifton, & Clarkson, 1989). Thus, the influence of the Conspec process may normally decline at 4 to 6 weeks of age due to inhibition by the emerging cortical Conlern mechanisms. A second line of evidence in favor of a subcortical basis of newborns’ preferential tracking is the nature of the task itself. Newborns demonstrate a preference for facelike stimuli in this task, but not in standard visual-preference tasks with static stimuli present in the central visual field (e.g., Maurer & Barerra, 1981). One explanation for these discrepant findings is that the tracking task is more sensitive to subcortical visual mechanisms. In young infants, the peripheral visual field feeds more directly into the subcortical retino-tectal pathway while the central visual field feeds mainly into the cortical geniculo-striate pathway. In the tracking task used with newborns, the stimulus is continually moving out of the central visual field and toward the periphery (stimulating the subcortical pathway) while in the visual preference tasks used, the stimuli typically remain in the central visual field (stimulating primarily the cortical pathway). Further support for this is a study in which newborns were tested monocularly so that orienting to faces in the temporal and nasal visual fields could be compared (Simion et al., 1998). Since, in newborns, vision in the nasal visual field is mediated primarily by the geniculo-striate pathway while vision in the temporal visual field is mediated primarily by the retino-tectal pathway, newborns should preferentially orient to faces only in the temporal visual field if this behavior is subcortically mediated. The results confirmed this prediction. A final piece of evidence in favor of subcortical mediation of newborns’ preferential orienting to faces is that adult patients with hemi-neglect show an advantage for faces in capturing attention in their neglected fields (Vuilleumier, 2000).

While none of these lines of evidence bring us any closer to specifying in detail the neural substrates of newborns’ preferential tracking of faces, a number of different possibilities remain including the superior colliculus, pulvinar, deeper layers of the cortex, and any combination of these. Probably only direct methods, such as neuroimaging, will provide a more definitive answer to this question.

Visual Attention Studies

The results of studies with newborns suggest that a preference for orienting to facelike patterns within a natural environment is present from the earliest moments after birth. At the same time, newborn infants are able to recognize individual facial identity: They look longer at the mother’s face than at the face of a stranger, even with static presentations and even when extra cues such as voice and smell are controlled (Bushnell et al., 1989; Pascalis et al., 1995). Despite these impressive abilities, the way in which newborns process faces is very different from the way adults do. At 6 to 8 weeks of age, striking changes occur which may reflect the beginnings of cortical specialization for face processing. In the two-process theory, these changes are viewed as reflecting the functional emergence of the cortical “Conlern” mechanism.

One such change is in the mental representation of facial identity. According to one view (Valentine, 1991), adults encode faces in a multidimensional space. The dimensions of this “face space” are not prespecified, but are thought to be learned to allow optimal discrimination of faces in the individual’s environment. Face representations are assumed to be normally distributed about the origin of the space such that typical faces are clustered together around the center, and more distinctive faces are located further out in more sparsely populated regions. Valentine proposed that face space might take one of two forms (Figure 2). In the norm-based model of face space, faces are encoded as vectors relative to a central prototype, the weighted average of all faces encountered. The alternative exemplar-based model assumes that faces are represented as points, rather than vectors, and that the origin of the space is simply the point of greatest exemplar density.

Few studies have addressed the issue of how face space might develop. However, it seems reasonable
to make certain assumptions. First, face space will contain fewer entries during development than in adulthood, as fewer faces have been encountered. Second, infants and children are likely to encode faces along fewer dimensions than adults because they do not need many to distinguish the smaller number of faces in their environment (Johnston & Ellis, 1995; Pedelty, Levine, & Shevell, 1985).

Infants appear to begin to form a face space based on the faces they see only at around 3 months of age. For example, we tested 1-month-old and 3-month-old infants’ abilities to recognize an individual face and an “average face” prototype following familiarization to four individual faces (de Haan, Johnson, Maurer, & Perrett, 2001). Consistent with the results of prior studies (e.g., Pascalis, de Haan, Nelson, & de Schonen, 1998; Pascalis & de Schonen, 1994), both 1- and 3-month-olds were able to recognize the individual face. However, only the 3-month-olds showed evidence of recognizing, and thus having mentally computed, the average face prototype. Results of further experiments showed that the 1-month-olds’ failure to recognize the prototype still occurred when the recognition test was made easier and could not be accounted for by preexisting preferences among the faces (de Haan et al., 2001). These results suggest that only by about 3 months of age do infants relate information from one face to another and begin to form a category of “face” based on the faces they see.
Studies of adults’ learning of perceptual categories indicate that recognition of individual exemplars is mediated by different substrates than categorization of exemplars and recognition of the prototype (Knowlton & Squire, 1993; Squire & Knowlton, 1995). The former is mediated by medial temporal lobe structures (Knowlton & Squire, 1993; Squire & Knowlton, 1995) while the latter involves visual cortex (Aizenstein et al., 2000; Reber, Stark, & Squire, 1998a, 1998b). In adults, both of these regions are believed to be part of the network involved in storing knowledge about object categories (Martin & Chao, 2001). Infants’ early abilities to recognize individual faces may rely on medial temporal lobe structures (e.g., the hippocampal-based “preexplicit memory;” Nelson, 1995) while their later developing abilities to form mental prototypes and use these to facilitate encoding of new faces may rely on further development of areas of the ventral visual cortical pathway.

Studies with older children demonstrate that there is a prolonged course of development of mental representation of faces. For example, with increasing experience of faces, children become better at discriminating between faces of the type that comprise the majority of their environmental input, and do not improve, and may even get worse, at discriminating between members of other categories of faces. For example, infants and children show a reduced or absent “other-race effect” (e.g., Chance, Turner, & Goldstein, 1982) and a reduced face inversion effect (e.g., Carey & Diamond, 1977). In face-space terms it appears that the dimensions of face space change to become increasingly attuned to making fine discriminations between upright, same-race faces with growing expertise with these stimuli. The cost of this increasing expertise is that, with development, we may become less able to discriminate between faces of types to which we have little exposure, such as monkey faces or inverted faces (e.g., see Werker & Tees, 1999 for a discussion of this phenomenon with perceptual narrowing in phoneme discrimination and Nelson, 2001 for further discussion of this view on the development of face processing).

Event-Related Potentials

Studying the development of processing of stimuli such as inverted faces or animal faces can provide insight into how face space is being altered during development. For example, adults are slower and make more errors in processing inverted than upright faces while they often are equally good at processing other objects in both orientations (e.g., Yin, 1969). This disruptive effect of inversion on face processing also is observable in event-related potentials (ERPs) recorded while adults passively view faces (for method description of ERPs, see Johnson et al., 2001, or see Taylor & Baldeweg, 2002 in a parallel special issue of Developmental Science on imaging methods). In adults, there is a “face-sensitive” negative component in the ERP, called the N170, that peaks at around 170 ms after stimulus onset and is most prominent over posterior temporal electrodes (e.g., Bentin, Allison, Puce, Perez, & McCarthy, 1996). The N170 is thought to be linked to “…late stages of structural encoding, where representations of global face configurations are generated in order to be utilised by subsequent face recognition processes” (Eimer, 2000). This component also is observed in children as young as 4 years of age, though its amplitude is smaller and latency longer at this age (Taylor, Edmonds, McCarthy, & Allison, 2001; Taylor, McCarthy, Saliba, & Degiovanni, 1999). The N170 is affected by stimulus inversion: It is of larger amplitude and longer latency to inverted compared to upright faces (Bentin et al., 1996; de Haan, Pascalis, & Johnson, 2002; Eimer, 2000; George, Evans, Fiori, Davidoff, & Renault, 1996; Rebai, Poiroux, Bernard, & Lalonde, 2001; Taylor et al., 2001). In contrast, there is no difference in amplitude or latency of the N170 elicited by upright compared to inverted animal faces (de Haan et al., 2002) or upright compared to inverted objects (Rebai et al., 2001; Rossion et al., 2000). These results suggest that the N170 elicited by the human face is not simply a reaction to the basic configuration of eyes–nose–mouth (since this is also present in animal faces), but is tuned more specifically to characteristics of the upright, human face.

In infants as young as 6 months, an ERP component is elicited by faces that is of similar morphology to the adult N170 but has an amplitude and latency more similar to that observed in young children (de Haan et al., 2002). However, unlike in adults, in 6-month-olds the N170 is not affected by inversion. This is not because infants of this age cannot detect the difference between upright and inverted faces, as they show modulation of the amplitude of a longer-latency ERP component (P400) by inversion. In contrast, 12-month-olds, like adults, show a larger N170 for inverted than upright human faces, but no difference in the N170 to inverted and upright monkey faces (Halit, de Haan, & Johnson, 2000). These results are consistent with the idea that the infants’ representation of the face initially may be relatively broad and only gradually become more specific to the upright human face. This process likely extends beyond
infancy, as there are developmental changes in the characteristics of the N170 throughout childhood (Taylor et al., 2001; Taylor et al., 1999).

**Neuropsychological Studies**

Results from the studies described previously suggest that some adultlike characteristics of face processing are present by 12 months of age. One possible explanation is that, rather than emerging through an experience-dependent process as we have proposed, these characteristics reflect a preprogrammed and anatomically localized cortical module for face processing that emerges over the first year. We will evaluate this possibility in light of recent neuropsychological studies.

**Williams Syndrome.** One piece of evidence that has been interpreted in support of this view is the impressive face-recognition abilities in individuals with Williams Syndrome. Williams Syndrome is characterized by mild to moderate mental retardation together with an uneven cognitive profile, with poor visuo–spatial skills but relatively intact verbal and face-processing skills (Donnai & Karmiloff-Smith, 2000). The seemingly preserved face-processing abilities, despite severe impairments in visuo–spatial processing, have been taken as evidence for the existence of a face-processing module that can be selectively impaired or spared. However, the results of recent studies demonstrated that the relatively good performance of people with Williams Syndrome on face-recognition tasks is mediated by atypical neural and cognitive processes. For example, typical adults show a gamma burst in EEG recorded during the viewing of faces, a response linked to perceptual integration of features; in contrast, this response is absent in individuals with Williams Syndrome (Grice et al., 2001). These results challenge the idea that face processing is normally mediated by a specific module.

**Developmental Prosopagnosia.** In adults, damage to occipito–temporal cortex can result in prosopagnosia, a selective impairment in face processing. Recently, several cases of selective deficits in face processing attributed to damage sustained in infancy have been reported (e.g., Bentin, Deouell, & Soroker, 1999; Farah, Rabinowitz, Quinn, & Liu, 2000; Jones & Tranel, 2001; Nunn, Postma, & Pearson, 2001). These results could be interpreted as evidence for the existence of a domain-specific and anatomically localized face-processing module formed early in life.

However, the results of other studies do not support this view. In one study of face-processing abilities in 5- to 14 year-olds who had experienced perinatal unilateral lesions, the effects were fairly mild: Fewer than half the children showed impaired performance relative to controls on tests of face or object-identity recognition (Mancini et al., 1998). Furthermore, (a) face-processing deficits were no more common than object-processing deficits following a right-hemisphere lesion, (b) face-processing deficits were no more common after right-sided than left-sided damage, and (c) a face-processing deficit never occurred in the absence of an object-processing deficit. This general pattern is similar to that reported in other studies (Ballantyne & Trauner, 1999; Mancini, de Schonen, Deruelle, & Massoulier, 1994), and suggests that the infant face-processing system is more widely distributed and/or more plastic following damage than is the adult system.

While the existence of even one “pure” case of developmental prosopagnosia might be thought to be sufficient proof of a domain-specific face-processing module localized early in life, in practice interpretation of such results is not entirely straightforward. First, if there is an anatomically localized, preprogrammed face module, then it would be expected that the location of damage causing prosopagnosia would be the same in developmental and adult-onset cases. However, in several developmental cases, MRI revealed no observable structural damage (Nunn et al., 2001) or diffuse rather than focal abnormalities (Laeng & Caviness, 2001). This raises the possibilities that the nature of the initial damage differed in these cases compared to adult-onset cases or that reorganization in response to injury differs in the developmental compared to the adult-onset cases. A second factor to consider is that, if the modular argument is correct, one would expect the nature of the face-processing deficit to be similar in developmental and adult-onset cases. While this possibility has not been extensively investigated, one recent comparison of two such cases demonstrated that the processes underlying impairments in face processing may differ for developmental compared to adult-onset cases (de Gelder & Rouw, 2000). A third factor to consider when interpreting cases of developmental prosopagnosia is that perceptual impairments that may not greatly affect adults’ processing of faces could have a big impact on the development of face-processing skills. For example, a recent study showed that visual input in the first months of life is critical to normal development of face processing later in life (Le Grand, Mondloch, Maurer, & Brent, 2001). Thus, early lesions that affect general visual function could, by altering the quality of the perceptual input, have a lasting impact on the developing face-recognition
system. Finally, the existence of cases with early sustained brain injuries and later difficulties in face processing is not necessarily evidence that a face-processing module initially existed and was damaged. For example, in the majority of normal adults, the left side of the brain plays a dominant role in language, and damage to the left perisylvian areas results in severe difficulties with language known as dysphasia or aphasia. In contrast, similar damage occurring in the first 5 years of life has much less devastating consequences for language, but can later result in unexpected deficits in visuospatial processing (see Vargha-Khadem & Mishkin, 1997). In the same way, it is possible that deficits in face processing could be caused by early acquired lesions in areas not normally associated with face processing in adults. Thus, for many reasons the existence of cases with developmental prosopagnosia cannot necessarily be considered strong evidence in support of the existence of an early-developing “face module.”

CONCLUSIONS

Infants arrive in the world with a preference to look at faces. We have argued that this early bias to orient to faces is not the result of an innate cortical face module, but rather reflects more primitive neural circuitry, involving mainly subcortical structures. Although the perceptual basis of this preference is still debated, no view disputes our proposal that it ensures a face-biased input to developing, modifiable cortical circuits. In this sense, the infant’s brain can be said to contribute to its own subsequent specialization. Grelotti et al. (this issue) present a view similar to ours in that they also propose that more “primitive” circuits of the brain function to ensure faces are a frequent input to higher cortical areas in the ventral visual pathway. Grelotti et al. focus on the amygdala and argue that it provides a “social interest” in faces whereas the two-process view we have presented focuses on structures involved in early visual processing and argues that they provide a “visual interest” in faces. We consider that both of these factors may be important, as there may be more than one source of constraint on the experiences that influence the development of face processing. Visuospatial mechanisms (necessary to provide appropriate input from the visual system for face recognition to occur) and social needs (endogenous drives for communication and interaction with other people) may combine to ensure appropriate specialization (Elgar & Campbell, 2001).

In this article, we also have alluded to the processes of specialization within the cortex, suggesting that with development the cortical circuits involved in face processing become both increasingly localized and increasingly specialized. These processes at the neural level are reflected in the emergence of several adult-like characteristics of face processing over the first year of life. In our view, the learning occurring during this period is critical to the normal development of face processing. In contrast to the view that face recognition is just one example of expertise that can be acquired by a more general-purpose “expert visual processing” system, we propose that face recognition is special in that the timing of certain visual inputs during development is important for normal development of the system to occur. By our view, the regions in the ventral occipito–temporal cortex have the potential to become specialized for face recognition, but it is only through appropriate experiences with faces that specialization for this function emerges. Other regions may initially be activated by faces, but in typical development eventually become specialized for different functions.

Thus, this account does not entail an innate module, but neither is specialization for faces wholly shaped by experience. Instead, we suggest that a neural system “prepared” for exposure to particular input develops rapidly and becomes increasingly specialized with ongoing exposure to faces. Currently, there is little empirical evidence to distinguish between developmental learning of face recognition and learning of visual expertise for other object categories as adults. However, there are some indications that the processes may differ. First, in the expertise view it is argued that adults become experts in processing categories; certain characteristics of processing that are often erroneously thought to be unique to faces (e.g., inversion effects, configural encoding, etc.) emerge. If we apply the same view to development, we would expect that these expert characteristics of processing would emerge when children became expert (adult-like) at processing faces. However, young infants, who are not considered experts in face recognition, show many of the characteristics of expert processing (e.g., configural encoding, inversion effects, right hemisphere bias; see discussion earlier and de Haan & Nelson, 1997, 1999; Deruelle & de Schonen, 1991, 1995; de Schonen & Mathivet, 1990). Thus, the early appearance of these characteristics of face processing may be one difference (or be indicative of differences) between the process of development and the process of adult learning. In contrast to normally developing infants, individuals with Williams Syndrome do appear to be relative experts at face processing at the behavioral level, yet their processing of faces does not show many of the typical characteristics of experts.
(e.g., they show reduced or absent inversion effect, are biased to processing featural information, etc). In the same way, while the development of face processing may on the surface resemble learning of other visual categories, the processes involved may not be the same.

Comparison of studies of the emergence of the effect of inversion on the N170 in response to faces during normal development (de Haan et al., 2002) and in response to nonface objects following training in adults (Roisin, Gauthier, Goffaux, Tarr, & Crommelinck, 2000) provides additional evidence on this question. There are some similarities between the infants and learning adults: In both cases, inversion initially has no effect on the N170, but an inversion effect on its amplitude emerges following “training” (natural experience with faces for infants and explicit training in recognizing nonface objects for adults); inversion leads to an increase in amplitude of the N170. However, the fact that changes in latency do not show parallel effects in developing infants and learning adults suggests that the processes involved in development of face recognition and adult perceptual learning may not be identical.

A final point to consider is that there is evidence that, unlike adult learning which can occur at any time, timing of experience is critical for normal development of face processing. This is demonstrated by recent results showing that individuals who were deprived of pattern visual input in the first months of life later show impairments in encoding configural information from faces despite normal encoding of featural information in faces and configural information in geometric patterns (Le Grand et al., 2001).

In summary, we have presented evidence to support an interactive specialization account of the development of face processing. This view does not entail an innate cortical module for face processing, but neither does it equate the process of development with the process of adult perceptual learning. Instead, we argue that primitive, probably subcortical, biases in infants’ visual orienting cause faces to be a frequent input to developing cortical systems. This process is one of the initial steps in what is likely a prolonged development of the ventral occipto–temporal cortical pathways involved in face processing in adults.

REFERENCES


