The Development and Neural Bases of Face Recognition

Charles A. Nelson*
Institute of Child Development, University of Minnesota, Minneapolis, MN, USA

Evidence from fields as diverse as cognitive, evolutionary, and developmental psychology, as well as cognitive neuroscience, has increasingly pointed to the ‘special’ nature of face recognition. A critical examination of the literature supports the view that faces begin to be seen as a separate class of objects within the first 6 months of life. Not surprisingly, the neural systems that underlie face recognition also come on line during this period of time. Less clear, however, are the mechanisms whereby these events occur. It seems likely that face recognition reflects an experience-expectant process, whereby exposure to faces during a sensitive period of development likely leads to perceptual and cortical specialization. However, it is unknown what the role of experience is in maintaining this ability, and how long this sensitive period lasts. After reviewing three related models that attempt to account for the way the ability to recognize faces develops, a number of suggestions are offered for testing the hypothesis that face recognition depends on experience for acquisition, and for evaluating the role of experience in maintaining this ability. Copyright © 2001 John Wiley & Sons, Ltd.

Key words: brain; development; face; object; recognition

Interest in face recognition has played prominently in various scientific disciplines for much of this century, and even parts of the last (Darwin, 1872). Cognitive psychologists have been interested in this phenomenon because there is evidence that faces are somehow perceived differently than other patterned objects, and thus, may represent a ‘special’ class of stimuli. Cognitive neuroscientists are interested in face recognition because there is evidence that this ability is subserved by discrete neural circuits, and thus, represents a specialized brain function. Developmental psychologists have long been interested in face recognition because faces provide an early channel of communication (prior to the onset of language) between infant and caretaker. Finally, evolutionary psychologists and ethologists have been interested in face recognition because it appears to be a ‘special’ ability that has been selected for through evolutionary pressures, and conserved across species. Despite the wide-ranging and prolonged interest in this topic, however, it is still unclear how face recognition

* Correspondence to: Institute of Child Development, University of Minnesota, 51 East River Road, Minneapolis, MN 55455, USA. E-mail: canelson@tc.umn.edu

Copyright © 2001 John Wiley & Sons, Ltd.
becomes specialized, and what drives the development of the neural systems that support this ability. It is in addressing these two issues wherein the focus of this paper lies.

IS FACE RECOGNITION A SPECIAL CASE OF OBJECT PERCEPTION?

There is evidence from many quarters to suggest that there is something special about the recognition of faces versus non-faces, although there is less agreement as to how the term ‘special’ should be defined. For example, some investigators have used this term to suggest that the perceptual processes involved in recognizing faces differ from those involved in recognizing objects, or that the neural substrate involved in face recognition is different than that which underlies the recognition of non-face stimuli, or that the neural computations involved in face recognition differ from those engaged in the perception of non-facial stimuli (for discussion of these perspectives, see Farah et al., 1999). Each of these perspectives is briefly illustrated in the sections below.

STUDIES OF INTACT AND NEUROLOGICALLY IMPAIRED ADULTS

Experimental studies performed with neurologically normal samples have suggested that faces are perceived as a special class of stimuli, distinct from other patterned objects. For example, Farah, among others, has suggested that face recognition differs from object recognition in that the former involves representing a face as a single, complex whole, whereas the latter typically involves decomposition into constituent elements (e.g. Farah et al., 1998; for review, see Farah et al., 1999). In support of this hypothesis are demonstrations that, in neuropsychologically normal adults, recognition performance is worse when faces are presented upside down than when objects are presented upside down (Valentine, 1988; Farah et al., 1995b).

A second line of evidence that faces are accorded special status comes from individuals with prosopagnosia. Prosopagnosia presents with a rather specific impairment in the ability to recognize familiar faces, and is frequently accompanied by damage to the ventral occipitotemporal and temporal cortex. An illustrative example of prosopagnosia can be found in patient L.H. (see Farah et al., 1995a,c; for review see Farah, 1996). L.H. is very impaired in recognizing familiar faces, although his general object perception is intact. Conversely, Moscovitch et al. (1997) have reported on patient C.K., who has a different pattern of brain damage. As a result, C.K. has intact face processing, but impaired object processing, thereby providing additional evidence that faces and objects can be dissociated at the perceptual level. (For representative examples of this literature, see Damasio et al., 1982; DeRenzi, 1986; McNeil and Warrington, 1993; Farah et al., 1995c, 1999). Finally, even more specific deficits have been observed in patients with highly focal lesions. For example, Adolphs et al. (1994, 1995) reported that patient S.M. (who incurred bilateral calcification of the amygdala) was impaired in her ability to judge facial expressions, particularly negative expressions (e.g. fear).

Research with patients has increasingly been complemented by neuroimaging studies with neurologically intact adults, and with single unit work with
monkeys (which provides even greater spatial localization than human neuroimaging). With regard to the former, Kanwisher et al. (1998) (using fMRI, a procedure whereby changes in the level of oxygen are measured in distinct brain regions) have reported increased activation in the fusiform gyrus to faces in general, although less activation is observed if the faces are presented upside down (Kanwisher et al., 1997). This same group has reported greater activation in this region to faces than to human or animal heads (Kanwisher et al., 1999). Also using fMRI, Gauthier et al. (1999) reported that as naïve subjects acquired expertise in recognizing artificial stimuli (e.g. ‘greebles’), the middle fusiform gyrus in the right hemisphere was recruited and showed a pattern of activation that was indistinguishable from that elicited by faces. Similarly, under passive viewing, activation in this area was greater in a single subject with expertise in viewing greeble faces versus individuals lacking such expertise. Overall, these results suggest that the fusiform ‘face area’ becomes specialized with experience, a point I shall return to in a subsequent section of this paper.

In terms of facial emotion, Whalen et al. (1998) reported increased activation in the amygdala to fearful faces, but decreased activation to happy faces. Finally, Morris et al. (1996) used positron emission tomography (PET) to examine the neural responses of subjects presented with happy and fearful faces. The authors observed that neuronal activity in the left amygdala was significantly greater to fearful as opposed to happy faces.

Collectively, from the neuroimaging studies, it appears that regions in and around the fusiform gyrus appear to play a role in face recognition, whereas the amygdala plays a particularly important role in the recognition of facial expressions.

Turning to the work with non-human primates, face-responsive cells have been found in several areas of the temporal cortex, particularly the temporal polysensory area (the superior TPO), areas TEa and TEM of the inferior temporal (IT) cortex (e.g. Rolls and Baylis, 1986; Baylis et al., 1987; Yamane et al., 1988), and along the ventral bank of the superior temporal sulcus (STS) (e.g. Perrett et al., 1982, 1984, 1985; Perrett and Mistlin, 1990). Moreover, using a combination of optical imaging and single unit recordings, Wang et al. (1996) have reported regional clustering of cells in the anterior IT cortex of monkeys during facial viewing. Finally, neurons responsive to faces have also been observed in the amygdala (Rolls, 1984; Leonard et al., 1985). Within the amygdala, some nuclei have been found to be responsive to individual faces, whereas others respond to individual expressions (Aggleton et al., 1980; Nahm et al., 1991). Importantly, both IT and the STS project to the lateral nucleus of the amygdala (Aggleton et al., 1980; Amaral et al., 1992), and thus, it may be that different regions of the IT cortex (e.g. IT, STS) and the medial portions of the temporal cortex (e.g. amygdala) work together to process faces at many different levels.

For a variety of reasons, it is difficult to compare the single unit work with monkeys to the functional imaging work with humans (e.g. differences in species, differences in the spatial resolution of single unit and optical recordings relative to fMRI and PET). Nevertheless, the general findings from both programmes of research collectively suggest that regions within the IT cortex play an important role in face recognition. Moreover, the role of the amygdala in recognizing the social/affective significance of faces seems particularly clear.

Although an unambiguous picture is emerging that faces may be accorded special status by the brain, it remains unclear upon what basis face specialization develops. From an evolutionary perspective, recognizing faces (and
particularly facial expressions) would be adaptive, and thus, selected for through evolution. This, in turn, should result in face recognition (a) being present in non-human primates to a comparable degree as in humans, (b) appearing early in life, and (c) being subserved by specialized neural tissue. The data reviewed above support the last of these assertions, and the remaining two topics are discussed next.

**BEHAVIOURAL STUDIES WITH MONKEYS**

Both human and non-human primates use their faces to produce a range of social signals; more importantly, non-human primates may depend more on this medium for communication than do adult humans, given the absence of oral language. Thus, it is not surprising that monkeys are adroit in both face and emotion recognition (for discussion, see Boysen and Bernston, 1989; Phelps and Roberts, 1994; Parr et al., 1998; Pascalis and Bachevalier, 1998). To cite but one example, Pascalis and Bachevalier (1998) tested both human adults and mature Rhesus monkeys on their ability to discriminate a range of human and monkey faces and objects. Both groups did equally well in recognizing objects. However, humans did better than monkeys in recognizing human faces, whereas the converse was true for monkeys. This species-specific effect suggests that experience in viewing faces played an important role in the findings.

In a recent review of this literature, Pascalis et al. (1999) concluded that many species of monkey are able to distinguish members of their own species from that of another species, and to recognize individual members of their own species. Moreover, these authors concluded that monkeys employ similar perceptual processes when studying faces, as do human adults; for example, in both species, the internal parts of the face are more salient than the external parts of the face. Collectively, then, the hypothesis that face recognition represents an important adaptive function that has been conserved across species appears to be supported.

**STUDIES OF DEVELOPMENT**

**Behavioural Studies in Newborns**

The study of development has played an important historical role in understanding the ‘special’ nature of faces, given early speculation that this ability was innate (e.g. Bühler, 1933; Bowlby, 1969). The logic behind this argument is that it is adaptive for the young infant to be able to recognize potential caretakers and/or emotional signals carried by the face (prior to the onset of language). Accordingly, experience with faces should be unnecessary for this ability to develop. Although there were, in fact, early reports that newborns (who would, by default, lack experience with faces) preferred to look at faces over other patterned stimuli (e.g. Fantz, 1963; Stechler, 1964; Goren et al., 1975), there were other reports that failed to find such preferences (e.g. Hershenson, 1965; Thomas, 1965). Over the years, however, the pendulum has swung back to supporting the view that newborns are capable of recognizing faces (e.g. Maurer and Young, 1983; Johnson et al., 1991; Pascalis et al., 1995; Valenza et al., 1996; Simion et al., 1998; for review, see Johnson and Morton, 1991; Morton and Johnson, 1991; Nelson and Ludemann, 1989), and indeed, are able
Face Recognition

to discriminate mother from stranger (Pascalis and de Schonen, 1994). Curiously, in some studies, this preference for face-like stimuli gradually wanes towards the second month, and then returns shortly thereafter (e.g. see Morton and Johnson, 1991).

As impressive as these findings are, the ability to recognize faces earlier than 1–2 months of age is somewhat fragile (e.g. the dependent measure and the behavioural state of the infant can contribute to whether one obtains positive or negative findings), and is not nearly as robust as it will be after 2 months; moreover, the mechanisms controlling face recognition may be different in the newborn than in the older infant (e.g. relative to older infants, newborns possess poor visual acuity, contrast sensitivity, and cannot resolve the high spatial frequencies that make up the fine details of faces; see de Schonen and Mathivet, 1990; Simion et al., 1998). In addition, most newborn studies have used highly schematized stimuli, in which the resemblance to a real face is loosely based on having eye sockets and openings for mouth and nose (e.g. Morton and Johnson, 1991; Simion et al., 1998), thereby questioning the extent to which these stimuli serve as a proxy for real faces. Moreover, in most studies, faces have not typically been contrasted to objects, and thus, it is not clear whether faces were, in fact, perceived as a distinct class of stimuli.

Behavioural Studies in ‘Older’ Infants

Suffice it to say there is rapid development in the sophistication with which infants respond to faces (for reviews, see Nelson and Ludemann, 1989; Johnson, 1997; Simion et al., 1998). For example, Fagan (1972) demonstrated that, beginning at around 4 months of age, infants’ recognition of upright faces is superior to the recognition of upside down faces, suggesting that, at this age, infants have developed a face ‘schema’, and thus, have begun to view faces as a special class of stimuli. Between 3 and 7 months, their ability to distinguish mother from stranger becomes more robust (e.g. Maurer and Salapatek, 1976), and they begin to categorize faces by gender (e.g. Cohen and Strauss, 1979), and by facial expressions (e.g. happy versus fearful; see Ludemann and Nelson, 1988).

Collectively, the recognition of faces appears to develop rapidly over the course of the first 6 or so months of life, and is far more robust than it is in the newborn period. In the sections that follow, I discuss the neural mechanisms that may underlie these impressive developments.

Neuropsychological Studies

de Schonen and colleagues (e.g. de Schonen et al., 1986; de Schonen and Mathivet, 1989, 1990) have elegantly demonstrated that infants ranging in age from 4 to 9 months show a right hemisphere (left visual field; LVF) bias towards processing faces, similar to what is observed in the adult (e.g. Lecuyer et al., 1978; Proudfoot, 1983; Rhodes, 1993). For example, de Schonen and Mathivet (1989) have reported that infants recognize a face faster if the face is initially presented in the LVF as opposed to the right visual field (RVF).

Neuroimaging-Type Studies

de Haan and Nelson (1997) used event-related potentials (ERPs) to examine the recognition of familiar faces in 6-month-old infants by contrasting mothers’ faces to strangers’ faces under a variety of conditions. The authors reported no ERP differences when two strangers’ faces were used, but clear ERP differences when mother was paired with stranger; importantly, the specific ERP
components that distinguished the two faces (i.e. negative component (NC)); positive slow wave (PSW)) differed depending on whether the stranger was similar or dissimilar looking to the mother. In both cases, greater ERP activity was observed at the right versus left temporal scalp, consistent with the right hemisphere bias for face processing observed by de Schonen and colleagues.

In a follow-up study, these same authors (de Haan and Nelson, 1999) examined the recognition of familiar and novel faces and objects in 6-month-old infants. The authors observed ERP differences that distinguished familiar from novel events (faces and objects), and importantly, faces versus objects. For example, a P440 component was observed over occipital scalp that had a shorter latency to faces than to objects, while the PSW was larger for novel than familiar stimuli, independent of whether the stimulus was a toy or a face. The NC was generally larger for familiar than for novel stimuli; for faces, however, the difference was seen at midline and right temporal scalp sites, while for objects, it was more widespread and bilaterally distributed over the surface of the scalp.

Finally, in a study of emotion recognition (specifically, happiness, fear and anger), Nelson and de Haan (1996a) reported that the amplitude of the PSW was greater to happy than to fearful faces, whereas the amplitude of the NC was greater to fearful than to happy faces. No ERP differences were observed when fear and anger were contrasted. The NC results of Experiment 1 are consistent with the behavioural findings that infants attend more to fear than to being happy (e.g. Nelson et al., 1979). The results of the Experiment 2 are consistent with the behavioural findings that infants have difficulty discriminating between two negative emotions (see Nelson and de Haan, 1996b for discussion).

Recently, Mazoyer et al. (1999) have reported a very unusual study, in which 8 neurologically compromised 2-month-old infants were studied using PET. Infants were presented with faces or flashing red and green diodes. Although a rather broad array of areas were activated to faces that seemingly had little to do with face processing per se (e.g. Broca’s area; medial superior frontal gyrus), activation was observed in the right fusiform gyrus, consistent with adult reports (see earlier sections). Activation of the fusiform gyrus would also be consistent with behavioural reports of infant preferences for face-like stimuli, although activation in other cortical regions is inconsistent with adult findings.

**Face Recognition in Infants and Children with Brain Damage**

Only a handful of studies has been conducted with children suffering from damage to the areas of the brain implicated in face recognition (for review, see Bentin et al., 1999). For example, Ellis and Young (1988) reported on a prosopagnosic 9-year-old girl (K.D.) who contracted meningitis at the age of 14 months. There was no evidence that K.D.’s disability showed signs of remitting over the first 9 years of her life; moreover, she showed no signs of being able to be taught to recognize faces. Similarly, Mancini et al. (1994) reported on a series of six patients, ranging in age from 7 to 11 years, three of whom sustained damage to the left hemisphere and three to the right (in five cases, the damage occurred prenatally, whereas in one, it occurred at 10 months of age). As is often the case with infants, the damage was diffuse and not restricted to any one area. Children varied enormously in their face and speech perception abilities (none were formally prosopagnosic). However, for those children who were impaired in face processing, the fact that they had had years of experience with faces raised the spectre of a certain lack of plasticity in the face processing ‘system’, as though there was no neural reorganization that permitted this ability to be
subserved by another region of the brain. A similar conclusion has been drawn from an even larger series of patients (n = 11) by this same group (Mancini et al., 1994).

More recently, Farah et al. (2000) reported on the case of a 16-year-old boy who developed meningitis at 1 day of age. At age 6, C.T. revealed bilateral occipital and occipitotemporal lesions. The authors report that ‘Adam’ shows no overt object agnosia when real objects are used, but does have difficulty with photographs of objects (particularly line drawings). Adam is, however, impaired in recognizing faces, including those he encounters in everyday life (and is, not surprisingly, even more impaired when photographs of faces are used). From this single case Farah and colleagues concluded that

... the distinction between face and object recognition, and the anatomical localization of face recognition, are explicitly specified in the genome. Whatever role environmental factors play in the normal unfolding of separate face and object recognition systems, the distinction between faces and non-face objects and their separate brain localizations does not require experience with stimuli from these different categories.

SUMMARY

The Mancini et al. and Farah et al. studies suggest that damage to the regions of the brain that subserve face recognition in the pre- or immediate postnatal period lead to long-term impairments in face recognition; the patient described by Ellis and Young (1988) shows a similar profile, although in this instance the damage was done after ‘K.D.’ experienced faces for a full 14 months of life. Although it is tempting to concur with the nativist conclusion reached by Farah et al., caution must be exercised before doing so. First, in none of these cases was the damage restricted to the regions of the brain that are typically associated with face recognition; indeed, the pathophysiology of the injuries (e.g. prenatal stroke) or illness (e.g. meningitis) that befell these patients rarely targets, in young infants, the inferotemporal cortex. On the assumption that these children were all exposed to a normal visual world following their injuries (including exposure to faces), it is possible that regions of the brain up or downstream from the ‘face’ regions ultimately led to the long-term impairment observed by the investigators; that is, the face region received faulty input or the output from the face region to higher cortical centres was corrupted. Second, it must be kept in mind that none of these patients performed normally on tasks involving the recognition of objects; they all possessed some degree of residual deficit. True, their face recognition abilities (particularly in the case of ‘K.D.’ and ‘Adam’) were poorer than their general object recognition abilities, thus leaving open the interpretation that the face ‘system’ was, in fact, compromised relative to the object system. Nevertheless, because the neural systems that subserve face and object recognition overlap, it is hard to apportion variance cleanly. This, in turn, suggests that the brain damage that occurred affected a more general purpose, higher level visual system than a specific face system (it may be possible to test this hypothesis if detailed structural information about the brains of these patients was available). Third, it is not clear if it was the inability (owing to brain damage) to perceive and interpret correctly a normal visual world (including exposure to faces) at a critical period of development that led to the observed deficits, or the failure to maintain normal visual access to such a world well beyond infancy and childhood that led to the observed deficits (i.e. in the
case of K.D., the system may have been set correctly, but not maintained, whereas in the case of Adam, both acquisition and maintenance were compromised). Finally, one must address the paradox of arguing that face recognition represents an innate ability (or at least an experience-independent ability), and thus, one that confers some evolutionary advantage . . . yet, if this is the case, why would something so seemingly important for survival show so little plasticity? On the whole, then, as compelling as these cases are (particularly ‘K.D.’ and ‘Adam’), a definitive answer to the question of whether face recognition represents an experience-independent ability is lacking.

SUMMARY AND DISCUSSION OF DEVELOPMENTAL FINDINGS

First, it is clear that face recognition develops rapidly in the first months of life and, under a number of conditions, is observed as early as a few hours after birth. Second, there is support for a right hemisphere (and possibly right temporal lobe) bias in processing faces. Third, there is evidence that faces are processed differently than objects (even when familiarity is controlled for). Fourth, we know that processing facial expressions develops on a similar time frame as processing other aspects of faces. Finally, there are some data to suggest that early damage to the regions of the brain that would normally subserve face recognition results in a long-term impairment, suggesting a lack of plasticity in this system. This, in turn, has led some authors to suggest that face recognition may not depend on experience, and thus, represents an innate ability (e.g. Farah et al., 2000).

How are we to account for the development of a face processing ‘system’ in light of the data reviewed herein? A number of authors have attempted to do so by positing ‘neoconstructive’ (see Karmiloff-Smith, 1998) accounts. These models, and an alternative, are discussed in the following section.

MODELS OF FACE RECOGNITION

*de Schonen Model*

de Schonen and Mathivet (1989) have proposed that the right hemisphere plays a particularly important role in face recognition, for two reasons. First, the right hemisphere is better suited than the left in processing configural information because such information consists predominantly of low spatial frequencies, which is ideal, given the limited visual abilities of infants in the first 6 months of life. Second, the right hemisphere develops before the left hemisphere. As infants’ contrast sensitivity function improves, experience with faces further drives the development of the right hemisphere face ‘system’, which gradually leads to greater and greater neural specialization. However, with time, the left hemisphere (which can also process faces, albeit less well) also benefits from this experience and, in so doing, it becomes possible to account both for the LVF bias for processing faces, as well as the more recent imaging studies pointed to bilateral activation of regions like the fusiform gyrus.

*Johnson Model*

Johnson and colleagues (e.g. Johnson and Morton, 1991; Morton and Johnson, 1991; Johnson, 1997) have proposed that infants less than 2-months-old will track face-like objects because of a subcortical visuomotor mechanism (possibly
involving the superior colliculus) referred to as CONSPEC. CONSPEC is, essentially, an innate structure that facilitates attention to faces (or at least face-like objects), possibly because this subcortical region responds preferentially to movement and to objects in the periphery (hence, the finding that newborns preferentially track face-like stimuli). The influence of this mechanism begins to wane over the first 2 months and be replaced by a second system, CONLERN, which is a cortical system that benefits from experience with faces.

In a recent review of this model, Johnson (1997) suggests that the CONLERN mechanism is essentially ‘set’ by the CONSPEC mechanism, that is, biases it towards faces. Once set, the cortical pathways involved in CONLERN are activated by experience with faces, thereby accounting for the remarkable progress in face recognition observed after 2–3 months of age.

Critique of de Schonen and Johnson Models

The de Schonen and Johnson models are similar in many respects, particularly in suggesting that development after the first few months depends heavily on experience. Both models also assume that the newborn brain possesses neural tissue that predisposes it towards faces. de Schonen, for example, speculates that the right hemisphere is biased towards processing low spatial frequencies, exactly the kind of information involved in recognizing differences in facial identity (but, perhaps, not the kind involved in making more subtle discriminations, such as distinguishing the facial expression of fear from surprise). Johnson’s model assumes that there is a built-in mechanism (possibly involving the superior colliculus) that biases infants to attend to face-like stimuli. Following these initial biases, both models argue strongly for the role of experience in influencing face recognition. Unspecified in both models is (a) whether there is a critical period for when exposure to faces must occur in order to stimulate the development of the face recognition system (by inference, the Mancini et al. and Farah data would suggest the first 6–12 months of life), and (b) for how long experience must occur in order to maintain the ability to recognize faces.

Both of these models do well in explaining the child patient data reviewed earlier. In the case of de Schonen, general visual deficits may have led to faulty input to the right hemisphere, which in turn, led to more specific deficits in recognizing faces. Similarly, Johnson’s model would predict that the CONSPEC and/or CONLERN mechanisms may have been perturbed by brain damage (although, as careful neuroimaging was not done with any of the patients, it is hard to know whether the midbrain tectum was damaged; as this structure lies in the visual pathway, we shall assume it might have been). As a result, this or these mechanisms likely did not receive the normative input that was expected. (This should not be true in the case of K.D., in whom brain damage did not occur until 14 months of age, thereby giving both systems ample experience with faces).

Despite the considerable strengths of both models, a number of questions remain unanswered. First, the de Schonen model fails to make clear why the right hemisphere is superior to the left in processing low frequency information (an observation that is belied by the symmetrical development of all other aspects of the visual system). Second, although there are behavioural data to suggest that the right hemisphere leads the left in development, there are, in fact, few anatomical or physiological data to support this claim (see Nelson and Bloom, 1997). Given the imprecision with which behavioural probes can be used to infer brain development, this claim requires further proof before being
accepted outright. Finally, there are many aspects of the visual world besides faces that contain low spatial frequency information, and thus, it is not clear why the right hemisphere would be positively biased towards faces per se.

The CONLERN element of Johnson’s model does a splendid job of accounting for many of the findings with older infants. However, the CONSPEC element is more problematic, for several reasons. First, faces are far from the only stimulus that moves or that exists in the periphery, and thus, it is not clear why CONSPEC would be positively biased towards faces per se. Second, to the best of this author’s knowledge, there is no evidence that the visual motor pathway has feature detectors that would cause faces to be attractive (relative to any other patterned stimulus). Finally, although many authors have engaged in heroic feats to ensure that newborns were tested as soon as possible after birth, the assumption that CONSPEC is an experience-independent process again assumes early and immediate exposure to faces (i.e. in the first hours of life) is insufficient to stimulate development; this remains to be tested.

AN ALTERNATIVE ACCOUNT

Rather than assume an ‘innate’ system like CONSPEC, or a right hemisphere bias towards face-like stimuli, it may be useful to consider instead an alternative explanation by drawing an analogy to speech perception. For example, like faces, even very young infants are able to discriminate a range of speech sounds (for review, see Aslin et al., 1998) and even recognize their mother’s voice (perhaps similar to the reports of newborns recognizing their mother’s face—DeCasper and Fifer, 1980). Second, non-human species, such as monkeys and chinchillas, are able to categorize human speech, comparable, perhaps, to how monkeys are able to discriminate human faces from one another (Kuhl and Padden, 1983). Third, in both human infants and in birds, experience appears to play a prominent role in recruiting cortical areas that specialize in speech perception, which, in turn, leads to increased perceptual proficiency (Doupe and Kuhl, 1999). Like faces, then, there appears to be the potential for cortical specialization for speech perception that is dependent on experience. Finally, and perhaps most importantly, there appears to be a narrowing of the perceptual window with increased exposure to speech. For example, prior to approximately 6–8 months of age, infants throughout the world are able to discriminate between phonemes from languages other than the one in which they are being reared. However, as infants approach 1 year, the perceptual window begins to narrow, and they behave more like adults—that is, they are best at discriminating those sounds that are native to their own language (Werker and Tess, 1984; Kuhl et al., 1992; Kuhl, 1993; Cheour et al., 1998). There are several examples of a similar phenomenon with faces. First, our laboratory published a preliminary report (Nelson, 1993) that suggested that human infants were superior to adults in discriminating monkey faces; presumably prolonged exposure to human faces in the adult led to a perceptual narrowing of the faces that could be easily discriminated, whereas, in the infant, the perceptual window was broadly tuned so as to support a range of perceptual discriminations. Second, as reviewed earlier, adults have difficulty discriminating inverted faces, a phenomenon that appears to emerge by about 4 months of age (Fagan, 1972). Third, there is the well-known ‘other race’ effect, in which adults, more so than children, find it easier to recognize faces from their own race (see Chance et al., 1982; O’Toole Deffenbacher et al., 1994). Fourth, there is the finding that expertise recognizing
greebles leads to recruitment of the fusiform gyrus (Gauthier et al., 1999). Fifth, as discussed previously, both monkeys and human adults are better at recognizing faces from their own species (Pascalis and Bachevalier, 1998). Finally, maltreated children perform more poorly on emotion recognition tasks than do non-maltreated children, a finding interpreted as supporting the view that experience with facial emotion facilitates the development of the ability to recognize emotion (Camras et al., 1983, 1988).

Let us assume, then, that like language, selection pressures have led to the genetic specification for neural tissue that has the potential to become specialized for face recognition (Darwin, 1872). However, there must be input into this system in order to set it initially, and to set it correctly (whether there is a critical or sensitive period for doing so is unknown). Over time, different portions of the temporal lobe are ‘captured’ by these early experiences, with the result that the face recognition system becomes fine-tuned. This argument is quite similar to that put forth by Morton and Johnson’s CONLERN proposal (Morton and Johnson, 1991), in which infants older than 2 months of age benefit from experience viewing faces.

What is assumed here is simply that evolutionary pressures have led to a cortex that is flexible and open to learning during development; thus, there are domain-relevant mechanisms that become domain-specific (see Karmiloff-Smith, 1998). In terms of the development of face recognition, then, all one need posit is that (a) regions within the IT cortex have the potential to become specialized for face recognition (much as our brains possess tissue that can become specialized for speech and language), (b) the face perception apparatus becomes tuned with exposure to faces, which in turn, leads to increased specification and parcellation of neural tissue, and (c) this specification includes the many types of information conveyed by faces, including recognizing gender, age and emotional information. The developmental literature reviewed herein suggests that the time frame for the development of general face perception abilities occurs rapidly within the first months of life, with fine tuning (e.g. in recognizing negative emotions like fear and anger) taking place later in the first year. Of course, a similar time frame applies to the recognition of objects (see Kellman and Banks, 1998).

CRITIQUE OF THE ALTERNATIVE ACCOUNT

There are a number of difficulties with the alternative model. First, as is the case with language (in which there appears to be dedicated neural tissue), there is the dilemma of accounting for why a specific region of the brain (in this case, regions in the inferotemporal cortex) is targeted to subserve face recognition. Thus, if the assumption is correct that experience drives development, how does experience recruit a specific brain region? Second, as is the case with the other models discussed, the alternative account fails in its precision to state whether there is a critical or sensitive period for the acquisition of a face ‘module’ (as most experience-expectant processes require, such as the development of ocular dominance columns), and in addition, for how long experience must be present to maintain the system until it has reached a steady state. Third, it is unclear exactly what kind of experience is necessary to drive the system. For example, if infants were raised with two-dimensional line drawings of faces for the first 6 months of life, would they then automatically generalize their face ‘schema’ to real faces? The question here, of course, is whether the system is capable of
bootstrapping from minimal information, or whether the input required must be fairly precise (i.e. exposure to normal faces).

CONCLUSIONS

The evidence from many quarters reviewed herein suggests that the development of face recognition is an experience-expectant process. This term refers to the development of skills and abilities that are common to all members of the species, and that depend on exposure to certain experiences occurring over a particular period of time (see Greenough and Black, 1992). Thus, here we assume that it is adaptive for the species to be exposed to faces (e.g. because such exposure confers an advantage in recognizing conspecifics, emotional messages etc.). Unfortunately, the extant literature fails us in specifying exactly what kinds of experiences are necessary, when these experiences need to occur, and lastly, for how long they need to occur. What kinds of studies are needed to address these questions? Selective rearing studies, of course, would assist in this goal, as might studies of human and monkey infants born prematurely. In addition, by studying infants born with specific damage to the IT cortex (or in the case of animals, performing neonatal lesions), it may be possible to observe plasticity in the face recognition system, and in so doing, infer the importance of experience.

Overall, the bulk of the evidence suggests that the ability to recognize faces is one that is learned. Through exposure to faces, tissue in the inferotemporal cortex becomes specialized for face recognition, and, in theory, continued exposure to faces maintains this tissue until it becomes dedicated to face recognition (which is not to say this same tissue cannot be used for other related purposes, such as recognizing stimuli like greebles). Why the inferotemporal cortex has been targeted to play this role is unclear, although a similar dilemma exists with regard to other specialized abilities, such as speech and language. Perhaps such tissue has been selected for through evolutionary pressures, or that the properties of the neurons and synapses that comprise this tissue are particularly tuned to this particular task. Regardless, it appears that this specialization occurs rapidly, within the first months of life. With greater experience with faces, a form of perceptual learning takes place that further develops this tissue. The cost of such specialization, of course, is that if this tissue is damaged, rather specific deficits can result (e.g. prosopagnosia). Moreover, early specialization may also led to a lack of developmental plasticity, an observation that may account for the failure of both children and adults to show recovery of function following neural injury (or, perhaps, a lack of sparing). Why such a seemingly important ability lacks plasticity—if true—remains to be determined.

ACKNOWLEDGEMENTS

The writing of this paper was made possible, in part, by grants from the NIH (NS23389) and the John D. and Catherine T. MacArthur Foundation and the J.S. McDonnell Foundation, through their support of the research network on Early Experience and Brain Development. I wish to thank David Amaral, Allison Doupe, Michelle de Haan, Sara Webb, and Chris Stevens for editorial feedback.
Notes

1. Greebles represent an artificial class of stimuli that bear some resemblance to toy animals. As described by Gauthier et al. (1999, p. 367) All Greebles have four protruding parts organized in approximately the same spatial configuration on a vertically oriented central part. The set is organized orthogonally along two categorical dimensions, such that each Greeble is a member of one of two ‘genders’ and one of five ‘families’. There are five central part shapes, each defining one of the five families. The gender difference is defined by the orientation of the parts relative to the central part, either all pointing upward or downward. Although some of the parts are very similar to each other, every individual part is unique within the set.

(For an illustration of Greebles, see figure 3, p. 356 of Gauthier et al., 1999).

2. This perspective seems more reasonable than suggesting that face recognition requires the expression of specific genes. Specifically, given the limited number of genes each member of our species possesses (perhaps 30000), it would seem an inefficient use of genetic material to specify an ability that will likely develop due to experience (see Greenough and Black, 1992).

REFERENCES


Phelps MT, Roberts WA. 1994. Memory for pictures of upright and inverted faces in humans (Homo sapiens), squirrel monkey (Saimiri sciureus), and pigeons (Columba livia). Journal of Comparative Psychology 108: 114–125.


Rolls ET, Baylis GC. 1986. Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. Experimental Brain Research 65: 38–48.


