The Origins of Face Processing in Humans

Phylogeny and Ontogeny

Olivier Pascalis and David J. Kelly

1University of Sheffield, Sheffield, United Kingdom, and 2University of Glasgow, Glasgow, United Kingdom

ABSTRACT—Faces are crucial for nonverbal communication in humans and related species. From the first moments of life, newborn infants prefer to look at human faces over almost any other form of stimuli. Since this finding was first observed, there has been much debate regarding the "special" nature of face processing. Researchers have put forward numerous developmental models that attempt to account for this early preference and subsequent maturation of the face processing system. In this article, we review these models and their supporting evidence drawing on literature from developmental, evolutionary, and comparative psychology. We conclude that converging data from these fields strongly suggests that face processing is conducted by a dedicated and complex neural system, is not human specific, and is unlikely to have emerged recently in evolutionary history.

Indisputably, humans are a social species. From the very first moments of life, newborn infants demand near constant maternal care and attention during their waking hours. This very special form of socialization represents the advent of our induction into life as a social organism. Throughout ontogeny, we socially interact with perhaps thousands of other individuals. Our ability to perform this function successfully hinges on our capacity to reliably convey our internal emotional state to conspecifics and understand theirs in return. Central to this communicative process is the human face.

The faces of our human conspecifics are ubiquitous within the visual environment. Indeed, face processing researchers often claim that humans spend more time looking at human faces than any other category of complex visual stimuli. To the best of our knowledge, this is an assertive observation rather than a statement of fact. However, few people would challenge the legitimacy of such a notion, and it is certainly one that we readily endorse. As a consequence of this immeasurable and unparalleled exposure, adults seemingly possess an extraordinary level of competence in the domain of face processing. It is this proficiency that lies at the heart of a debate that is endemic within the face processing literature: Does face processing hold a unique status within the brain?

In this review, we will argue that the answer to this question is an emphatic "yes" and that the evidence to support this conclusion is overwhelming. In particular, we will draw on developmental, evolutionary, and comparative psychology to demonstrate that face processing in humans and related species is performed by a distributed cognitive system that is unlikely to have evolved recently in evolutionary history and is pivotal for our ability to function as a social species.

EVIDENCE FROM DEVELOPMENTAL PSYCHOLOGY

Despite much contention within the developmental face-processing literature, there is at least a certain level of consensus. It is established that newborn human infants demonstrate a visual preference for both real and schematic human faces over almost any other category of stimulus (Fantz, 1964; Goren, Sarty, & Wu, 1975; Johnson, Dziuarawiec, Ellis, & Morton, 1991; Valenza, Simion, Macchi Cassia, & Umilà, 1996). Although this may not provide conclusive evidence for an innate face processing module, it is suggestive of a specialized cognitive system operating from very early in life. Despite the apparent generality of the newborn face-processing system, there are two notable specific face abilities found early in life. First, from 2–3 days of age, newborns prefer to view their mother's face when paired with a stranger's face (e.g., Bushnell, Sai, & Mullin, 1989; Pascalis, de Schonen, Morton, Denuelle, & Fabre-Grenet, 1995), and second, newborns show a preference for attractive faces when paired with less-attractive faces (Slater et al., 1998). Both of these preferences can be abolished with some ease. In the...
case of the mother’s face, masking the external features (outer face contour and hairline) is sufficient to remove the preference. For attractive faces, the preference can be removed by inverting the stimuli.

In addition to face preference, human infants also demonstrate evidence of categorization—the propensity to group faces into distinct morphological sets. However, newborns do not show sensitivity to the subtle differences of morphology that are clearly perceptible to adults, and they instead show a broad preference for “faces” in general. However, face-processing capabilities develop quickly from birth, with infants exposed to more and more faces in their environment. De Haan, Johnson, Maurer, and Perrett (2001) demonstrated that infants begin to show evidence of face prototype formation at 3 months of age. A prototype is an average face that has been created by encoding several individual faces. This is further supported by the findings that 3-month-old infants show a visual preference for faces that match the gender of their primary caregiver (Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002) and for faces of their own ethnic group (Bar-Haim, Ziv, Lamy, & Hodes., 2006; Kelly, Liu, et al., 2007; Kelly et al., 2005). Collectively, these findings demonstrate the importance of experience with faces from the infant’s visual environment and how they impact the development of the face processing system.

MODELS OF DEVELOPMENT

CONSPEC/CONLERN

To account for newborns’ face preference at birth, Morton and Johnson (1991) proposed that a subcortical visuomotor system termed CONSPEC exists at birth and serves to orient the newborn toward facelike patterns. According to their account, CONSPEC contains a set of innately specified principles that are responsible for directing the attention of the newborn to facelike patterns present in the visual environment. The primary function of CONSPEC is to ensure that facial input is maximized during the first 2 months of life before a second system termed CONLERN comes “online.” CONLERN is located cortically and develops as a result of the cortex’s exposure to faces, which, as described above, is ensured by CONSPEC.

Evidence for CONSPEC/CONLERN

Principal support for the CONSPEC tier of Morton and Johnson’s (1991) model comes from reports that newborn infants preferentially look at schematic faces over scrambled faces and patterns (Goren et al., 1975; Johnson et al., 1991). In these experiments, researchers assessed newborns’ relative interest for three types of stimuli: a face, a scrambled face and a blank image. Researchers mounted stimuli on paddles and steadily moved them across the newborns’ visual field while recording eye and head movements. Under these testing conditions, both eye and head movements were found to be greater for schematic face stimuli, thus demonstrating that newborns show more interest in facelike patterns relative to other stimuli. The finding that newborns only display a face preference when stimuli are presented in the temporal visual field (which feeds to subcortical pathways) is consistent with the subcortical location hypothesis. In contrast, when stimuli are presented in the nasal visual field, which feeds to cortical pathways, a face preference is not elicited (Simion, Valenza, Umlàtto, & Della Barba, 1995). Researchers also found preferential looking to schematic faces when using a visual preference task in which the images are presented statically (Valenza et al., 1996). This finding does not necessarily undermine the notion of CONSPEC, but it does demonstrate that movement is not required to engage the system as originally postulated by Morton and Johnson (1991).

As introduced above, the proposed role of CONSPEC is to guide newborns’ interest toward relevant stimuli in the visual environment and to ensure that cortical areas responsible for higher order face processing (e.g., individuation) develop adequately. However, this limited functionality of CONSPEC renders it unable to account for a range of face abilities reported in newborns. For example, newborns prefer to view their mother’s face when paired with a stranger’s face (Bushnell et al., 1989; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis et al., 1995), show a preference for attractive faces when paired with less attractive faces (Slater et al., 2000, Slater, Quinn, Hayes, & Brown, 2000; Slater et al., 1998), show a preference for a female adult face with open eyes when paired with the same face with closed eyes (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000), and look longer at a face with a direct gaze when paired with a face with an averted gaze (Farroni, Csibra, Ahluwalia, 2000), and look longer at a face with a direct gaze when paired with a face with an averted gaze (Farroni, Csibra, Ahluwalia, 2000). Furthermore, newborns are able to recognize an unfamiliar face even following a delay of 2 min (Pascalis & de Schonen, 1994) or change in profile orientation between habituation and test (Tutari, Bull, & Simion, 2007).

In response to such findings, Johnson (2005) has updated his model by drawing on functional imaging, neuropsychological, and electrophysiological studies to argue that converging evidence favors the existence of a rapid, low-spatial frequency (LSF) face detector incorporating the superior colliculus, pulvinar, and amygdala (de Gelder, Frissen, Barton, & Hadjikhani, 2003). Johnson advocates that the alternative view to a subcortical processor, namely a cortically located complex face processor, is unnecessary. Instead, he proposes that the various face effects delineated above can be accounted for by an LSF face detector. However, in dismissing the complex face-processing argument, Johnson only draws reference to two of the findings described above: eyes present or not present and attractive face preference. Although classifying faces with eyes present or not present does not require high spatial frequency (HSF) information and could plausibly be conducted by a subcortical mechanism, accounting for the attractiveness preference in this manner is more problematic. In addition, none of the other “face effects” described above are mentioned, but it is...
questionable whether recognition of invariant face features across a change in viewpoint can be achieved using LSF information alone. However, it has recently been shown that newborns’ recognition of static faces (no change across viewpoint) is based on visual information contained in spatial frequencies between 0–0.5 cycles per degree (c/°; de Heering et al., 2008). Furthermore, newborns were unable to recognize a face when only the information contained within 0.5–1.0 c/° was available. This is a clear demonstration of an LSF bias in newborns, and it highlights the fact that this information is more important for the newborn than for adults, who use a combination of LSF and HSF information for identity recognition (Sowden & Schyns, 2006).

So what can we conclude about CONSPEC and the role it plays in the development of the face-processing system in early ontogeny? It is unequivocal that newborn infants are equipped to not only detect faces, but also to conduct rudimentary forms of higher order processing, such as recognition, for which adults use both LSF and HSF information but which may be achieved by newborns using only LSF information. Nevertheless, it is far from certain that newborns do not use HSF information for certain face-processing tasks. Indeed, it seems highly likely that they must use HSF information, as fine-grained information contained around the eyes required to perceive the white sclera to judge gaze direction is only discernable at 0.5–1.0 c/°. It appears that as we discover more about the newborns’ face-processing system, it increasingly resembles an immature adult system, marginalizing the need for a system such as CONSPEC.

Still, a rapid LSF face detector cannot be dismissed, as there is compelling evidence to suggest that such a mechanism does indeed exist. In adults, evidence from both electroencephalograph and magnetoecephalograph techniques have revealed face-selective components occurring prior to V1 activation (Bailey, Braeutigam, Jousmaki, & Swithenby, 2005; Braeutigam, Bailey, & Swithenby, 2001), indicating subcortical processing. In addition, prosopagnosic patients who are unable to recognize the identity of even highly familiar faces have little difficulty detecting faces (e.g., Rossion et al., 2003), signifying a probable dissociation between the pathways responsible for carrying information required to perform these tasks.

In conclusion, we think there are two plausible explanations for the face-tracking observations reported by Johnson and others. CONSPEC may exist at birth, but akin to other reflexes observed in infancy, such as the grasping reflex, it disappears by 2 months of age. These evolutionary ancient reflexes become suppressed at this time, as they are now unnecessary for survival in humans. CONSPEC may cease functioning at 2 months of age for precisely the same reason. If this interpretation is correct, there exists an intriguing possibility that prosopagnosic patients may no longer be suppressing CONSPEC and may display the face-tracking behavior observed in newborns in the same way that patients with a lesion of the supplementary motor area demonstrate recovery of the grasping reflex seen in young infants. It is thought that this reflex reemerges due to the removal of cortical suppression. It is more likely, however, that CONSPEC does not vanish at 2 months of age but is rather a rudimentary LSF face detector that endures throughout life to perform this important function. It is dissociable from the cortical face system, although feedback between the systems is likely (Johnson, 2005), arguably making them one integrated system.

### Gestational Proprioceptive Feedback (GPF)

An alternative or possibly complementary hypothesis to CONSPEC/CONLERN, proposed by Quinn and Slater (2003), is that the face-processing system could be innately provided by evolutionary adaptation, and the newborn’s initial face representation could be formed through proprioceptive feedback provided by facial movements could contribute to the formation of a face representation at birth. According to this hypothesis, a two-tier model is unnecessary. Instead, the face representation formed in utero elicits the attraction to matching configurations, and subsequently, the face processing system develops postnatally through visual experience with faces.

### Evidence for GPF

A range of prenatal abilities are necessary for Quinn and Slater’s (2003) hypothesis to become tenable. First, certain motor abilities are necessary for any form of proprioceptive learning to take place. The fetus must be capable of arm and hand movement so physical contact with their face can be experienced. A second form of proprioceptive feedback can come directly from the face via mouth opening (e.g., yawning), eye movement, and the flexing of the developing facial muscles. Fortunately, the remarkable images produced by 4-D ultrasound scans have revealed that these types of behavior are commonplace as early as the 11th week of pregnancy. At this point, we can observe the movement of the arms and legs, followed by finger dexterity at 14 weeks and movement of facial muscles at 16 weeks.

At 18 weeks, the fetus is fully engaged in regular hand to face contact (e.g., eye rubbing), as well as yawning, sucking, hiccuping, and the production of different facial expressions. Finally, by 26 weeks, the fetus’ eyes have developed more fully, and blinking behavior is exhibited. These abilities continue to become fine-tuned during the remaining gestational period, resulting in many weeks of opportunity for a face representation to develop.

Second, in addition to these capabilities, the fetus must also be capable of learning, so proprioceptive information can be utilized to form a meaningful representation. It is well known, for example, that newborns display a preference for amniotic fluid odor and their mother’s voice (see Lecanuet & Schaal, 1996, for a review). In addition, the work of DeCasper and others has demonstrated that newborns will show a preference for a familiar
passage that has been read aloud by their mother during the final 6 weeks of pregnancy (DeCasper & Spence, 1986) and that 5-month-old infants display a preference for carrot juice if it was ingested by the mother toward the end of pregnancy (Mennella, Jagnow, & Beauchamp, 2001). When considered together, it seems unequivocal that learning during gestation routinely takes place.

Finally, crossmodal recognition is required for the face proprioceptive representation to become meaningful. The newborn must be capable of matching a face (or schematic face) presented visually with the representation that has been built-up in utero. Until recently, such a suggestion would have been highly controversial, but Streri and colleagues have provided evidence of crossmodal recognition in 3-day-old infants. In their studies, they tactually habituated newborns to a shape (prism or sphere) and then visually presented them with familiar and novel shapes. They observed a significant preference for the novel shape, suggesting that “tactual-to-visual” recognition is achievable during the first days of life (Streri & Gentaz, 2003).

A recent study by Fatma Sai (2005) has yielded further evidence for the existence of crossmodal matching abilities in newborns. Sai elegantly demonstrated that newborns only recognize their mother’s face if postnatal exposure to the mother’s voice–face combination was available. It appears that the mother’s face is in fact learned in conjunction with the mother’s voice, which has been heard during gestation. If the infant is denied the auditory input of the mother’s voice after birth, recognition of the mother’s face is delayed. It is important to note that this finding highlights the multimodal aspect of the infant’s face representation from the first week of life. Simion, Regolin, and Bulf (2008) recently showed that newborns display a preference for looking at point-light displays of chickens (biological motion) relative to scrambled, nonbiological motion. This result suggests strongly an innate ability to recognize the movements of other living creatures that can complement the process of facial information.

A potential stumbling block for the GPF hypothesis comes from a report that the face preference observed in newborns may be linked to the contrast difference between the sclera and the iris, which is unique to humans (Kobayashi & Kohshima, 1997). In a series of experiments, Farroni and colleagues assessed preference for several pairs of face and schematic faces stimuli (Farroni et al., 2005). Of particular relevance here is the finding that contrast reversed stimuli (i.e., a black face with white blocks) did not elicit a preference, but the same stimuli did produce greater looking when black dots were added to the centers of the blocks. The authors conclude that the contrast polarity between the eyes and/or face is critical for eliciting a face preference in newborn infants, which clearly cannot be learned through proprioceptive feedback. The findings of Kobayashi and Kohshima (1997) also support this conclusion, making it unlikely that the preference for faces is linked to proprioception.

However, if a face preference is dependent on the contrast between iris and sclera, then we shouldn’t expect to find face preferences in other species. Yet, as described in greater detail later in this article, preferences do exist in a range of primates during infancy, including macaques, gibbons, and chimpanzees (Kuwahata, Adachi, Fujita, Tomonaga, & Matuzawa, 2004; Lutz, Lockard, Gunderson, & Grant, 1998; Myowa-Yamakoshi & Tomonaga, 2001). These preferences are harder to explain in terms of contrast differences between iris and sclera as it is much less perceptible. Therefore, exploiting this contrast difference may just be one way to categorize a human face as such.

PERCEPTUAL NARROWING

To address development in later infancy, Nelson (2001) has drawn a parallel between the developmental trajectories of language and face processing. With regards to language, studies investigating the development of phonemic perception have shown that the ability to discriminate phonemic contrasts that are absent in the infants’ native language (e.g., Werker & Tees, 1984) declines toward the end of the first year of life. For example, English contains a contrast between /t/ and /l/, whereas Japanese does not. English does not, however, contain a distinction between the retroflex /D/ and the dental /d/ that is common in the languages of Southern Asia. Numerous studies (e.g., Lisker & Abramson, 1971), have demonstrated that adults are unable to differentiate between contrasts such as these which fall outside their native language, but infants are able to make discriminations despite having never heard such sounds (e.g., Aslin, Pisoni, Hennessy, & Perey, 1981; also see Werker & Tees, 2005, for a review).

With regards to face processing, Nelson proposes that infants may begin life with a general mechanism dedicated to processing faces that subsequently becomes “tuned” to human faces as a direct consequence of the facial input received within the first months of life (see Scott, Pascalis, & Nelson, 2007, for a recent review). According to this account, the infant begins life with a crude and unspecified face representation that is then subject to modification as a result of the category (i.e., human) of facial input received. This notion is best understood within the framework of the multidimensional face space model described by Valentine (1991). Valentine proposes a norm-based coding model in which faces are encoded as vectors according to their deviation from a prototypical average. As argued by Nelson, at birth, the dimensions of the prototype are considered to be broad and largely unspecified, with ensuing development of the prototype dependent on facial input. The resulting dimensions will differ according to the input received with certain salient, individualizing dimensions carrying more weight than others. Predominant exposure to faces of a specific species, gender, or ethnicity early in life will cause the dimensions of one’s prototype to become tuned toward such faces.
EVIDENCE FOR PERCEPTUAL NARROWING

Nelson’s hypothesis does not address the emergence of the face-processing system but instead focuses on its development between 3 and 12 months of age. According to Nelson’s hypothesis, face recognition should become specific to human faces soon after 3 months. To test this, we investigated the ability of 6- and 9-month-old infants to recognize faces from their own species (human) and those from other species (Rhesus macaque) using a standard infant recognition paradigm. As expected, infants at both ages were able to demonstrate recognition with human faces. However, when tested with the monkey faces, only the 6-month-old group showed evidence of recognition, suggesting that the face system becomes tuned to human faces between 6- and 9-months of age (Pascalis, de Haan, & Nelson, 2002). We then investigated whether the system could be manipulated during this period to maintain the ability to recognize individual monkey faces. To accomplish this, we first replicated our findings with a new group of 6-month-old infants. The parents of the infants then took home a book containing a selection of monkey faces, which they were instructed to show to their infant following a fixed schedule. We found that this brief exposure was sufficient to preserve recognition capabilities with monkey faces when the infants returned for testing at 9 months of age (Pascalis et al., 2005). Recently, we found a very similar pattern of results with faces from other races (Kelly, Quinn, et al., 2007).

A dramatic demonstration of perceptual narrowing and the importance of early visual input and has recently been reported by Sugita (2008). Infant Japanese macaques (Macaca fuscata) were separated from their parents at birth and reared by humans for a period of 6–24 months. During this period, the monkeys had no interaction with other monkeys, and the human caregivers wore masks to ensure that the monkeys were not exposed to faces of any kind. Following the face-deprivation period, monkeys were then either introduced to fellow macaques or humans, and face learning was allowed to commence. When tested for face preference during the deprivation period, monkeys showed a preference for both monkey and human faces over objects, but no preference for either category of face when presented simultaneously. Furthermore, monkeys were able to recognize familiar monkey and human faces, as measured with a standard recognition paradigm, despite no previous exposure to any such faces. When the monkeys were tested a month after the termination of the deprivation period, their behavioral performance had altered dramatically. Monkeys exposed to human faces retained their preference for human faces over objects, but they also preferred to look at human faces over monkey faces. The finding that monkeys exposed to monkey faces preferred monkey faces over objects but showed no preference when human faces and objects were presented simultaneously is also consistent with this pattern of results. Furthermore, when tested for recognition, monkeys were only able to display recognition for faces from the category to which they were exposed.

Surprisingly, the period of deprivation did not have a significant impact on performance in preference or recognition tasks at any stage. Sugita (2008) interprets this as evidence for a face processing system that is protected against a lack of exposure to faces and interference from nonface objects for at least 2 years. Intuitively, one might have anticipated that face deprivation would be harmful for the face recognition system, but instead the system appears to simply remain in an immature state capable of processing faces from at least two categories, despite a complete lack of exposure. A challenge for future research will be to unify this finding with the apparently conflicting evidence from human congenital cataract patients (e.g., Geldart, Mondloch, Maurer, de Schonen, & Brent, 2002; Le Grand, Mondloch, Maurer, & Brent, 2001), which seemingly demonstrated that facial input in the first months of life is critical for development of typical face processing abilities.

In adults, a face selective electrophysiological activity has been observed in event-related potential (ERP) studies. Faces elicit a negative deflection around 170 ms after stimulus onset (N170), which is of larger amplitude and shorter latency than the activity elicited by objects (Bentin, Allison, Puce, Perez, & McCarthy, 1996). It has been reported to be influenced by stimulus inversion, the N170 being of larger amplitude and longer latency for inverted human faces than that found for upright human faces. This effect is particular to human face stimuli and has not been observed for animal faces (de Haan, Pascalis, & Johnson, 2002) or objects (Rossion et al., 2000). De Haan et al. (2002) have found an “infant N170” in 6-month-olds that was elicited by faces (290 ms) and was followed by a positivity at 400 ms. The infant N170 showed similar latency range for both upright and inverted faces, whereas the positivity at 400 ms differs significantly for the two category of stimuli. They also examined the influence of stimulus inversion for monkey faces and found that inversion in adults does not affect the processing of monkey faces, whereas inversion in 6-month-old infants affected the ERPs similarly for processing of human and monkey faces. At around 12 months of age, the adult ERPs pattern is observed for both human and monkey faces (Halit, de Haan, & Johnson, 2003). This pattern of results closely mirrors the behavioral data reported above, thus adding weight to the theory of perceptual narrowing.

Models Summary

It seems that the models of face processing development described here can be seen as complementary rather than conflicting, as they converge to describe a highly specialized process that is functional from the very first moments of life. It appears unequivocal that newborn infants will preferentially look toward facelike stimuli even in the first hours of life and will demonstrate a host of further specific face preferences in the first days of life. Furthermore, the capability of infants to process faces undergoes a period of increasing specialization that occurs
Evidence From Comparative Psychology

Face processing is not human specific, and similarities with other primate species are suggestive of a common origin. In both human and primate societies, face recognition plays a crucial role. The adult face systems of primates share several similarities with the human system: eye scanning, individual recognition, and inversion effects (see Pascalis, Petit, Kim, & Campbell, 1999, for a review). In addition to similarities in the adult system, the developmental trend is also comparable. Faces represent a highly attractive stimulus for infant primates, including, pigtailed macaques (Macaca nemestrina; Lutz et al., 1998), gibbons (Hylobates agilis; Myowa-Yamakoshi & Tomonaga, 2001) and chimpanzees (Pan troglodytes; Kuwahata, & Kiyohara, 2004). Although the developmental time scale differs between species, the same is true for both general brain development and other cognitive functions. As described above, human infants show an ability to recognize primate faces until 9 months of age, demonstrating that the face processing system is becoming tuned to human faces during the first year of life. In primates, 2-month-old pigtailed macaques are already demonstrating a unique preference for their own species (Kim, Gunderson, & Swartz, 1999). Face processing is very specialized from an early age in both humans and primates and may have played an important role in the development of primate societies that are based on strong relationships (Thierry, 1994).

In addition to primates, other social species demonstrate comparable face abilities. For example, sheep show highly competent and specialized face-recognition abilities (Kendrick, da Costa, Leigh, Hinton, & Pierce, 2001). Tibetts (2002) has found that wasps, who primarily rely on chemical signature for communicating with and identifying conspecifics, can also demonstrate hierarchical categorization (e.g., worker) on the basis of facial cues only. Collectively, these results strongly imply that a face-processing system is not human or even primate specific and is unlikely to have emerged recently in evolution.

EVOLUTIONARY PRESSURES

If we are to consider that face-processing capabilities are carried out by an evolved system, it is necessary to reflect on the evolutionary pressures that might have been responsible for bringing this system into existence. More specifically, it is interesting to consider the system in terms of natural and sexual selection: survival and reproduction. To survive within the natural world, it is vital that an organism is capable of detecting conspecifics and heterospecifics, as detailed above. In the case of humans, once we establish that another entity is indeed a fellow Homo sapien, we will need to make judgments concerning familiarity and in-group/out-group membership. In addition, it is important to be able to read an individual's intentions, which may be revealed by assessing that individual's emotive state through analyzing their facial expression. With regards to reproductive success, an organism must make a series of additional judgments when selecting an appropriate mate. When a male encounters a female, he must assess her for attractiveness (fitness) and reproductive potential. When a male encounters another male, he must make judgments about dominance and subsequently behave appropriately.

Some researchers have argued that primates may possess an innate face-processing system that is predisposed to respond to conspecifics (e.g., Sackett, 1966). The argument is supported by a handful of studies demonstrating that monkeys raised without experience of other monkeys from their own species will still prefer to look at faces of conspecifics (Fujita, 1990, 1993). However, this proposal does not fit well with findings from the human face-processing literature (e.g., Pascalis et al., 2002) or with the most recent work on primates (Sugita, 2008). Furthermore, this extreme view is almost certainly unnecessary. A more realistic argument is that humans and primates possess an evolved system for processing faces that became specialized as a consequence of predominant exposure to faces from a single species. According to this interpretation, expert levels of face recognition are dependent on both an evolved mechanism and experience. Therefore, due to the lack of biological relevance of objects (e.g., Greebles; Gauthier & Tarr, 1997) used in studies that have attempted to mimic the development of naturally acquired expertise, even intense exposure from birth would not produce the same “expert” levels of face processing observed in adulthood.
factors (i.e., symmetry and averageness) that drive attractiveness preferences in adults. Unsurprisingly, researchers have typically endeavored to explain attractiveness preferences in adults by searching for a physical characteristic that is common to faces rated as attractive. However, the approach of evolutionary psychologists has been to explain attractiveness in terms of mate choice by attempting to understand how hormones and genetic heterozygosity manifest themselves visibly within the face. Intriguingly, it appears that we are able to unwittingly extract a host of information regarding physical fitness and reproductive viability.

For example, females rate males with greater heterozygosity at key loci in the major histocompatibility complex, which is responsible for producing a strong immune system, as more attractive than males who are more homozygous in these regions (Roberts et al., 2005). Also, although females typically prefer male faces with feminine qualities (Perrett et al., 1998), their preference switches to more masculine faces during the follicular phase of their cycle when the likelihood of conception is highest (Penton-Voak et al., 1999). In terms of female attractiveness ratings, photographs of women taken during the follicular phase are judged as more attractive than images of the same woman taken during the luteal phase (Roberts et al., 2004). This suggests that ovulation in females is visibly manifested in the face and that seemingly imperceptible, subtle differences make females most attractive during the fertile period. Furthermore, ratings of female attractiveness, femininity, and general health are also positively correlated with levels of estrogen and progesterone, which are known to be good predictors of fecundity and reproductive health (Law Smith et al., 2006).

When considered collectively, a strikingly consistent pattern of results has emerged from this literature. Both male and female faces contain covert cues that are integral for distinguishing attractiveness preferences and ultimately maximizing reproductive success. Moreover, in addition to the display of these cues, the brain has also evolved mechanisms that reliably identify and decode these otherwise indiscernible facial variations. A challenge for future research will be to better understand the relationship between the factors that drive attractiveness in adults and the preference for attractive faces reported in newborns.

SUMMARY

In our opinion, it would be wholly surprising if a face-specific processing mechanism did not exist. It is vital for all species to be able to detect conspecifics and to discriminate between ingroup and out-group members. Any organism that is incapable of performing this task would be unable to detect potential threats and would therefore be in near constant danger. Because this ability is so fundamental for survival, it almost certainly would have appeared at a very early stage during evolution. This is supported by the fact that most species appear capable of performing this basic function (Baron-Cohen, 1995). Furthermore, fossils recovered from the Burgess Shale in British Columbia, Canada and from Chengjiang, China clearly illustrate that distinct predators and prey have coexisted for at least 525 million years. However, the ability to detect conspecifics and potential threats does not equate to individuation.

The ability to individuate conspecifics is imperative for any species that lives in a social group. Individuation can be performed in a variety of ways, and it is reasonable to assume that most species will use the most developed of their senses to perform this task. Although primates individuate using their visual faculties, other species may rely on their olfactory (e.g., hamster; Lai, Ramiro, Yu, & Johnston, 2005) or auditory (e.g., birds; Moseley, 1979) capabilities. It could be that a general mechanism to individuate emerged at some point during evolution and is exploited in different ways by various species. Alternatively, such abilities may have evolved independently in different species. Regardless, it is unequivocal that successful interaction for social species hinges on highly developed individuation capabilities.

To summarize, we argue that previously postulated developmental models have tended to concentrate on restricted periods of development. As a consequence, there is no single model that fully incorporates all aspects of the developing face system. From birth to the end of infancy, or even earlier, the face representation is present very early in life, but has not necessarily accounted for how and why a face representation might be found at this time. We suggest that an early face system observed in newborns is a conjunction of evolutionary inheritance, in utero learning, and rapid learning after birth. At birth, the face representation is crude, based on the evolution template and enhanced by the proprioceptive acquisition (e.g., eye movement). This early template then benefits from a rapid learning system that takes advantage of crossmodal abilities as described above.

It is surprising that the evolution of a face-processing system has been largely overlooked within the developmental literature. Individuation is clearly a vital mechanism for any species that must be able to identify family, friends, and foes. Indeed, we should be surprised if such systems were not found in both humans and other species that must perform this task for both survival and socialization purposes. Although it is difficult to directly test for the evolution of a face-processing system, evidence from the literature discussed above has found many similarities between the human and primate face-processing systems. It is our belief that future studies conducted with primates and other species that are more distantly related to humans are likely to uncover further similarities that will enrich our understanding of the evolution of face-processing capabilities in humans.

Face processing is a complex task, and if we are to truly understand its nature and development it will be beneficial to exploit available electrophysiological and neuroimaging techniques. The core of the human neural system for face processing
as revealed by fMRI studies consists of three cortical areas: the inferior occipital gyrus (occipital face area), the middle fusiform gyrus (fusiform face area; FFA) and the superior temporal sulcus. These areas become more activated in humans while viewing human faces rather than other visual objects (Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Asgari, Gore, & McCarthy, 1996). To date, only a handful of studies have attempted to use such methods in children and infants, but evidence for the development of the cortical areas associated with face processing in adults is starting to emerge.

Using positron emission tomography, Tzourio-Mazoyer et al. (2002) found adultlike brain activation in 2-month-old infants when viewing faces. However, faces also activated areas that are typically devoted to language in adults, indicating a possible early link between the visual and auditory systems or even a more distributed network. In terms of evidence from fMRI studies, Gathers, Bhatt, Corbly, Farley, and Joseph (2004) found that the area identified as the FFA is bilaterally activated by faces in children from 9 years of age. A more recent study by Golarai et al. (2007) compared recognition of objects and faces in children, teenagers, and adults. When viewing objects, similar brain activation was found for all three age groups. In contrast, FFA activation was found when the two older groups were presented with faces, but there were difficulties isolating the FFA in the youngest group. Although the findings currently remain equivocal, we can tentatively conclude that the FFA is developing from an early age, but the overall pattern of results support a steady specialization of the face neural network throughout ontogeny.

A further challenge for future research will be to integrate the many aspects of face processing into a complete model. Currently, the majority of research conducted in this field is solely concerned with understanding face recognition. Although this is clearly of interest and importance, it is crucial to acknowledge that when we view a face, we also rapidly make judgments regarding gender, attractiveness, age, and emotion and/or intentions. If we are to understand face processing completely, it will be essential to discover how we perform these judgments and the neural architecture that underlies them.

**CONCLUSION**

Face processing is a complex and multifaceted task, and consequently we should not expect a simple system to underlie it. Convergent evidence from behavioral, electrophysiological, and neuroimaging studies provide strong support for a complex neural architecture that is most likely dedicated to performing this task. In addition, comparative studies indicate that faces are highly important for closely related primate species and other mammals, such as sheep. The studies described in this review indicate that faces may be processed in a similar way in many species by a common system. Presently, attempting to pinpoint when such a system might have emerged would be speculative. However, given the abilities of all species to accurately detect predators and the different forms of individuation capabilities observed, it is unlikely to have occurred recently.

**Acknowledgment**—David J. Kelly was supported by an Economic and Social Research Council postdoctoral fellowship (DTA-026-27-1678).

**REFERENCES**


