
The Social Cognitive Neuroscience of Infancy:

*Illuminating the Early Development of Social Brain Functions*

Mark H Johnson¹
Tobias Grossmann¹
Teresa Farroni¹,²

¹ Centre for Brain and Cognitive Development, School of Psychology, Birkbeck, London, UK
² University of Padua, Italy

For: Advances in Child Development and Behavior, Volume 36.

Contact Author-
Mark H Johnson
Centre for Brain and Cognitive Development
School of Psychology, Birkbeck College
The Henry Wellcome Building
Torrington Square
London WC1E 7HX
UK
I. Introduction

II. Face Processing
   A. The sensory hypothesis
   B. Non-face structural preferences
   C. Newborns have complex face representations
   D. Developing cortical areas for face processing

III. Eye Gaze Processing
   A. Newborns detect direct gaze
   B. Gaze following
   C. Neural processing of direct gaze in infants
   D. Understanding eye-object relations

IV. Perception of Emotions
   A. Processing emotions in the face
   B. Processing emotions in the voice
   C. Integration of emotional information from face and voice

V. Interaction between Face Identity, Eye Gaze and Emotion
   A. Dissociable neural pathways for face processing in adults
   B. Interactions between face processing pathways in adults
   C. Interacting face pathways in infants

VI. Conclusions
I. Introduction

A distinct network of cortical and sub-cortical structures is selectively activated when we are presented with, or need to process, information about fellow humans. This network has been termed the “social brain” (Brothers, 1990; Adolphs, 1999, 2003), and the specific field of research around it has become known as “social (cognitive) neuroscience”. Although hundreds of studies have focused on this network in adults, its origins in infancy remain unclear. Studying the social brain in infancy is important not only for developmental psychologists, but also for those interested in fundamental questions about the adult system. For example, there is a long-standing debate in the adult literature about the degree to which a cortical area known to be activated by faces, the “fusiform face area” (Kanwisher, 1997), is specifically dedicated for processing this type of stimulus (to the exclusion of others). While some argue that the region undertakes (face) category-specific processing (Kanwisher, 2000), others have presented evidence that the region can be recruited by any class of stimulus to which the participant acquires perceptual expertise (Gauthier et al., 1999). Evidence relevant to this debate comes from developmental studies in which the effects of experience can be more easily assessed (Cohen-Kadosh & Johnson, in press). More importantly, accounts of the emergence of specialised functions in the brain that are inspired by developmental considerations can reconcile apparently conflicting views based only on data from adults (Johnson, 2005).

The earliest stage of postnatal development, infancy (0-2 years), is the time of life during which the most rapid changes take place. The dependent newborn seems almost a different species from the active, inquisitive 2-year-old toddler. During this period, human infants are faced with at least two vital and daunting challenges. The first is to understand their physical world and its properties, such as that objects still exist even when they are out of sight, are solid, and tend to fall downwards from tables. The second major challenge is for the infant to gain understanding of the social world, filled with other people, including
parents, siblings, other family members. Learning how to act in the physical world is difficult, but it is generally predictable in the sense that most events are contingent and highly replicable. In contrast, learning about the social world means not only developing the ability to predict the general behaviour and responses of fellow human beings, but also going beyond the sensory input to understanding the intentions, beliefs and desires of others. Although the social world has some predictability, some cues are less reliable than others, and events are rarely exactly the same from one time to the next. Furthermore, relating socially to others has not only profound effects on what infants feel, think, and do, it is also essential for healthy development and for optimal functioning throughout life. Indeed, the ability to learn from other humans is perhaps one of the most important adaptations of our species (Csibra & Gergely, 2006). Therefore, developing an understanding of other people is arguably the most fundamental task that infants face in learning about their world.

The development of the brain circuitries involved in all kinds of cognitive processes depends upon the interaction of two broad factors: nature (our inheritance or genetic factors), and nurture (environmental influences). If we aim to understand how these factors interact to ‘build’ the mature social brain network, it seems of particular importance to look at how the human brain deals with social information during infancy. In this chapter we review several aspects of the emerging social brain in infancy from the perspective of two related theoretical perspectives. The first of these perspectives derives from Johnson’s (2005a) updated version of Johnson and Morton’s (1991; Morton & Johnson 1991) two-process model. The original two-process theory sought to reconcile apparently conflicting lines of evidence about the development of face processing by postulating the existence of two systems: a tendency for newborns to orient to faces (Conspec), and an acquired specialisation of cortical circuits for other aspects of face processing (Conlern). We hypothesised that Conspec served to bias the input to developing cortical circuitry over the first weeks and months of life thus ensuring that
appropriate cortical specialisation occurred in response to the social and survival-relevant stimulus of faces.

The Conspec notion occupies the middle ground between those who argue that face preferences in newborns are due to low-level psychophysical biases, and others who propose that infants’ representations of faces are, in fact, richer and more complex than we supposed (which we describe in detail in the next section). Johnson and Morton (1991) speculated that Conspec was mediated largely, but not necessarily exclusively, by subcortical visuo-motor pathways. Briefly, this proposal was made for several reasons: (a) that the newborn behavioural preference declined at the same age as other newborn reflexes assumed to be under sub-cortical control, (b) evidence from the maturation of the visual system indicating later development of cortical visual pathways (e.g. Johnson, 1990), and (c) evidence from another species (the domestic chick) (see Horn, 2004).

Subsequently, cognitive neuroscience, electrophysiological and neuropsychological studies with adults provided evidence for a rapid, low spatial frequency, sub-cortical face detection system in adults that involves the superior colliculus, pulvinar and amygdala (Johnson, 2005). For example, evidence from patients with hemi-spatial neglect indicates that they display visual extinction to stimuli in the neglected field unless arranged in the pattern of a face (Vuilleumier, 2000; Vuilleumier & Sagiv, 2001). Vuilleumier uses this and other evidence to propose that the human amygdala is part of a low spatial frequency sub-cortical face processing route that boosts cortical processing of faces when these stimuli are most relevant – a system for “emotional attention.” These and other lines of evidence strongly support the view that there is a ‘quick and dirty’ neural route for face processing in adults, and that this same system operates in young infants, including newborns (Johnson, 2005).

The second theoretical perspective that motivates this review of the social brain in infancy concerns a model of functional organisation in the cerebral cortex called “Interactive Specialisation.” Johnson (2001) presented three distinct, but not necessarily mutually
exclusive, frameworks for understanding postnatal development in humans. According to a
“Maturational” perspective, functional brain development involves the sequential coming “on
line” of a number of modular cortical regions. The maturation of a given region is thought to
allow or enable advances in the perceptual, cognitive, or motor abilities of the child. As
applied to the neurodevelopment of social cognition, this implies that more complex aspects
(such as “theory of mind” computations) will depend on the maturation of associated cortical
regions (possibly within the prefrontal cortex). The second perspective, “Skill learning,”
argues for parallels between the dynamic brain events associated with the acquisition of
complex perceptual and motor skills in adults, and the changes in brain function seen as
infants and children come to successfully perform simpler tasks. From this perspective, it has
been argued that some cortical regions become recruited for processing social information
because typical humans become perceptual experts in this domain (Gauthier & Nelson, 2001).
A third perspective, “Interactive specialisation,” posits that functional brain development, at
least in cerebral cortex, involves a process of specialisation in which regions go from initially
having very broadly tuned functions, to having increasingly finely tuned (more specialised)
functions (Johnson, 2001). A consequence of increased specialisation of cortical regions is the
increasingly focal patterns of cortical activation resulting from a given task demand or
stimulus. By this view, some regions of cortex gradually become increasingly specialised for
processing social stimuli and thus become recruited into the social brain network.

In this chapter, we use the sub-cortical route hypothesis, and the interactive
specialisation model, to review and interpret studies on the development of face perception
and identity, eye gaze perception, and the perception of emotional expressions. We then
discuss how these different facets of social brain function interact, and speculate on useful
directions for future research.
II. Face processing

The face provides a wealth of socially relevant information. To detect and recognise faces is therefore commonly considered to be an important adaptation of social animals. From birth, human infants preferentially attend to some face-like patterns (Morton & Johnson, 1991; Johnson, 2005) suggesting a powerful mechanism to bias the input that is processed by the newborn’s brain. Perhaps the most controversial aspect of the two-process theory was the proposal that “infants possess some information about the structural characteristics of faces from birth” (Morton & Johnson, 1991, p. 164). Although this committed us to the view that the general configuration that composes a face was important, we did not commit to a specific representation. Nevertheless, our empirical observation from the early experiments with newborns, and evidence from other species, indicated that a stimulus with three high-contrast blobs corresponding to the approximate location of the eyes and mouth might be sufficient. A stimulus that appeared minimally sufficient to activate this representation (Figure 1) was termed “Config”. We did not claim that this stimulus was identical to the representation supporting the hypothesised Conspec system, and we remained open to the view that more ideal stimuli for activating Conspec could potentially be created. At the time of publication the idea that infants were born with face-specific information had been rejected by most in the field, largely on the basis of experiments with 1- and 2-month-old infants that failed to show face preferences (see Johnson & Morton 1991 for review). Because infants beyond the newborn stage did not prefer schematic faces over scrambled faces, it was concluded that newborns could not discriminate faces from other stimuli.

---

INSERT FIGURE 1 about here

---

At present, the Conspec notion occupies the middle ground between those who argue that face preferences in newborns are due to low-level psychophysical biases, and others who
propose that infants’ representations of faces are, in fact, richer and more complex than we supposed. We will defend the view that configurational information about faces is important, and argue that this configurational information is sufficient to account for most of the currently available evidence pertaining to schematic and naturalistic face-like images.

Between 1991 and 2007, there were at least 16 studies conducted on face preferences in newborns (see Johnson, 2005 for review). All of these studies found some evidence of discrimination of face-like patterns. However, one study failed to demonstrate a preference for face configuration (Easterbrook et al., 1999). In this study the authors failed to find evidence of preferential tracking of face-like patterns, but did find discrimination between face and non-face-like patterns in a habituation-dishabituation procedure. In considering the reasons for this apparent failure to replicate it is useful to refer to the reasons why Johnson et al. (1991) changed their procedure for testing infants after the newborn period. Specifically, they found that the procedure of moving the test stimulus around the newborn’s head, and measuring the extent of eye and head turning to keep the stimulus in view, was not effective after the immediate newborn period due to maximal (ceiling) tracking to all patterned stimuli. They therefore used a different testing procedure in which the infant was oriented relative to the stimulus. In the experiments conducted by Easterbrook and colleagues, all of the patterned stimuli used elicited greater degrees of orienting than in prior studies. Although the precise reasons for this greater extent of tracking in Easterbrook’s study are not known, differential responding is unlikely when near-ceiling performance is achieved. This methodological point aside, several types of explanation have been advanced for newborn face preferences that have been observed in at least five independent laboratories and the vast majority of studies conducted.

A. The Sensory Hypothesis: A number of authors have advanced the view that newborn visual preferences, including those for face-related stimuli, can be accounted for simply in terms of the relative visibility of the stimuli. For example, Kleiner and Banks (1987, Kleiner
1993) originally argued that the “linear systems model” of infant visual preferences could be used to account for newborn face preferences. Specifically, by this account newborn face preferences could be accounted for in terms of a simple psychophysical measure, the energy in the amplitude spectrum of the stimulus (following Fourier decomposition). Morton et al. (1990) reviewed critical experiments on this topic, and concluded that even when amplitude is held constant, phase information (roughly speaking, configuration) still influences preference toward faces. In addition to these arguments, Johnson and Morton (1991) conducted spatial frequency analyses on the range of schematic face stimuli discussed in their book. This analysis revealed that although spatial frequency is an important predictor of newborn visual preferences, stimuli with the phase appropriate for a face were always more preferred than predicted by spatial frequency alone.

In a later revival of the sensory hypothesis, Acerra and colleagues (2002) constructed a neural network model of visual cortical processing in infants. This model, based on four properties of the infant visual system (the spatial frequencies related to an immature retina; the tuning properties of neurons in area V1; the activity-dependent learning properties of both feedforward and lateral connections in cortical areas; and the absence of interhemispheric connections) successfully simulated the results of some of the experiments on newborns preferences for schematic face stimuli. In one critical comparison, that between the upright and inverted “Config” stimulus (see figure 1), differences in the spacing of high-contrast elements relative to the outline frame yields slight differences in spatial frequency that can be detected and amplified within the model. To produce the correct predictions for the behavioural data simulated, however, depends precisely on the exact details of the stimuli used in one experiment (Valenza et al. 1996). Specifically, Valenza et al. (1996) inverted the internal facial features in their stimuli in a way that resulted in the inverted stimulus having less regular spacing between the white and black areas than the upright case. The model exploited and amplified this small difference in spatial frequency profile, and was not tested
with the equivalent patterns from other newborn experiments (such as Johnson et al. 1991). Furthermore, Bednar and Miikkulainen (2003) doubt that the Acerra et al. model will be able to simulate results from experiments involving real faces, and especially the complex visual scenes that newborns are exposed to in real life. This is because very small irrelevant differences between stimuli (such as where the hairline falls on a face) would have amplified effects on visual preferences.

**B. Non-face structural preferences:** Debates about the mechanisms underlying face preferences in infants have often revolved around a contrast between structural and sensory preferences. For much of this debate, structural preferences referred to information about the configuration of features that compose a face. An alternative proposal is that infants’ preference behaviour to face and non-face stimuli can be explained by “domain-general” structural preferences, such as those based on known adult “Gestalt” principles (see Simion et al., 2003). Specifically, Simion and colleagues argue that “newborns’ preference for faces may simply result from a number of non-specific attentional biases that cause the human face to be a frequent focus of newborns’ visual attention” (p. 15, Simion et al., 2003). These authors still believe that the purpose of these biases is to direct attention to faces in the natural environment of the newborn, so that they are debating the nature of the representation(s) that underlie this bias, not the existence of the mechanism itself.

The behavioural data discussed by Simion and colleagues is consistent with a preference for up-down asymmetrical patterns with more elements or features being in the upper half of a bounded object or area. Such preferences are sometimes said to be due to an upper visual field bias (Simion, Valenza, Macchi Cassia, Turati, & Umiltà, 2002). However, all the experiments to date indicate a crucial interdependency between the borders of the object/area and the elements within it. Indeed, some existing experiments from this group suggest that the shape of boundary needs to correspond with the number of elements enclosed within it (Macchi, Valenza, Simion & Leo, submitted). Experiments that independently manipulate
upper visual field elements and bounded areas, and experiments that measure eye movements sufficiently to control upper/lower visual field presentation have not yet been done. Therefore the attentional bias being discussed by these authors - a bounded object or area with a greater number of elements or features in the upper half – is likely the product of one system or representation rather than being independent non-specific biases.

Skepticism about the “top-heavy” account of newborn face preferences seems warranted for a number of additional reasons. First, an assumption behind the “top-heavy” view is that a series of non-specific biases is a computationally simpler and more parsimonious account of newborn preference than Conspec. However, as discussed previously, the extent literature suggests that the hypothesized underlying mechanism has, at a minimum, to integrate a boundary with the inner elements/features. More likely, object segregation followed by numerical assessment of the number of elements present will be required. In contrast, Conspec has been simulated by computational neural network model in which a basic Config representation arises within the artificial cortex as a result of spontaneous activity (Bednar & Miikkulainen, 2003).

Another reason for doubting that an upper visual field bias alone is sufficient to account for newborn preferences comes from studies of the effects of phase contrast (Farroni et al., 2005). In these experiments newborn preferences for upright compared to inverted Config patterns and photographic face images were assessed with both black elements on white (as in previous studies) and the converse (see figure 2). If the newborns are seeking elements or features then phase contrast should make either no difference, or possibly cause them to prefer white elements on a black background (because lighter elements are typically closer to the viewer in natural scenes). In contrast, if the purpose of the representation is to detect faces then black elements on white should be more effective, because the eyes and mouth region are recessed into the face, and appear in shadow under natural (top-down) lighting conditions. Consistent with the latter view, we (Farroni et al., 2005) found the preference for an upright
face only under conditions of black on white or when smaller dark blobs are placed inside the light elements. Infants preferred face-like configurations only when some contrast relations within the image resembled to the natural difference between the darker iris and the lighter sclera of human eyes. This was also shown to be the case with photographic images. Newborns seem to seek for face-like stimuli that provide them with gaze information as well.

Another area that casts doubt on the “top heavy” hypothesis is evidence taken to support the existence of complex face processing abilities in newborns. This evidence is discussed next.

---

insert Figure 2 about here

---

C. Newborns have complex face representations: Some empirical results have lead to the hypothesis that newborns already have complex processing of faces (for review see Quinn & Slater, 2003). These findings, usually obtained with naturalistic face images, include a preference for attractive faces (Slater et al. 1998, 2000), data indicating that newborns are sensitive to the presence of eyes in a face (Batki et al., 2000), and that they prefer to look at faces with direct gaze that engage them in eye contact (Farroni et al., 2002). Such findings have led some authors to conclude that “…face recognition abilities are well developed at birth” (Quinn & Slater, 2003, p.4).

The usual interpretation of attractiveness preferences in older infants is that these stimuli are seen as more face-like because they best match an average or prototypical face (e.g. Langlois & Roggman, 1990). We note in passing that the newborn effect is not just due to increased symmetry because no preference is found with inverted stimuli. However, if we assume that Conspec is activated most strongly by the most typical spatial arrangement of features (rather than equidistant blobs), then a preference for more average of typical arrangements of facial features would be expected. Furthermore, inspection of realistic face
images through the appropriate spatial frequency filters for newborns reveals that a mechanism sensitive to the arrangement of high-contrast, low-spatial frequency components of a face would be preferentially activated by (a) the typical configural arrangement of eye and mouth regions, (b) the presence (or absence) of open eyes, and (c) direct versus averted gaze (see Farroni et al., 2002). Clearly, the extent to which Conspec can explain newborn responses to a variety of naturalistic faces requires further investigation. However, it is very difficult to see how the “top-heavy” hypothesis can explain this body of data, suggesting that this theory can only account for a narrow range of preference phenomena in newborns. Whatever the exact basis of the newborn preference, most investigators agree that it is sufficient to cause newborns to orient to faces in their early visual environment.

Besides the detection of faces, another important aspect is to recognise familiar faces. Newborns first recognise their mother’s face on the basis of information from the outer contour of the head, hairline, and the internal configuration of eyes, nose, and mouth (Bushnell, Sai, & Mullin, 1989). But not until 6 weeks of life are infants able to do this recognition solely on the face’s internal configuration (de Schonen & Mathivet, 1990). The face preference observed in newborns is thought to be guided largely by subcortical brain structures, but the maturation of visual cortical areas is necessary for the emergence of the more sophisticated competencies underlying identity recognition from faces (for a discussion, see Johnson, 2005).

D. Developing cortical areas for face processing: Investigations of brain areas involved in face processing in infants have been limited by ethical and technical issues (de Haan & Thomas, 2002). One exception has been a study by Tzourio-Mazoyer and colleagues, who took advantage of the opportunity to perform positron emission tomography (PET) on infants in an intensive care unit as part of a clinical follow-up (Tzourio-Mazoyer et al., 2002). In this small-scale study, a group of six 2-month-olds were imaged while they watched a face or an array of coloured diodes used as a control stimulus. A subtraction analysis revealed that faces
activated a network of areas in 2-month-old infants’ brains similar to that described as the core system for face processing in adults (Haxby, Hofman, & Gobbini, 2000). More specifically, activation was reported in an area in infants’ right inferior temporal gyrus, which is thought to be the homologue of the adult fusiform face area (FFA) (Kanwisher, 2000; Gauthier et al., 1999). It is interesting to note that a cortical region that at the age of 2 months is neuroanatomically immature (Huttenlocher & Dabholkar, 1997; Huttenlocher, 2002) and has only a low level of metabolic activity (Chugani & Phelps, 1986; Chugani, Phelps, & Mazziotta, 1987) can exhibit functional activity. Furthermore, as we will see later, activation of an area in response to faces does not mean that the area is specifically tuned for this function.

Face perception also activated bilateral inferior occipital and right inferior parietal areas in infants. The former has been implicated in early perceptual analysis of facial features, whereas the latter is thought to support spatially directed attention in adults (Haxby et al., 2000). Interestingly, and contrary to what is known from adults, face processing in 2-month-olds also recruited areas in the inferior frontal and superior temporal gyrus, which have been identified as a part of the adult language network. One possible interpretation is that the coactivation of face and future language networks has a facilitative effect on social interactions guiding language learning by looking at the speaker’s face (Tzourio-Mazoyer et al., 2002).

While bearing in mind the small sample size and non-optimal control stimulus, the study by Tzourio-Mazoyer and colleagues provided important information on a structural level by identifying the neural substrates of the infant face processing network. However, as mentioned earlier, due to ethical and technical concerns with the use of neuroimaging methods like PET and functional magnetic resonance imaging (fMRI) in infants, evidence coming from these kinds of studies is rare. Therefore, most studies that have examined infant
brain function rely on the more readily applicable recording of EEG measures, which provide excellent temporal resolution but relatively poor spatial resolution.

We now discuss briefly the empirical evidence available on infants’ face processing using event-related brain potentials (ERPs; for a more detailed review, see de Haan, Johnson, & Halit, 2003). In adults, human faces elicit an N170 response, which is most prominent over posterior temporal sites and is larger in amplitude and longer in latency to inverted than to upright faces (Bentin et al., 1996; de Haan, Pascalis, & Johnson, 2002). This component is not modulated by the inversion of monkey faces (de Haan et al., 2002), nor when upright objects are compared to inverted objects (Bentin et al., 1996). This selective inversion effect for human faces has been taken as evidence for a face-related processing mechanism generating the N170.

On the basis of waveform morphology and some of its response properties, it has been suggested that the infant N290 is a precursor to the adult N170. In these studies, infants’ and adults’ ERPs were measured in response to upright and inverted human and monkey faces (de Haan et al., 2002; Halit, de Haan, & Johnson, 2003). The infant N290 is a negative-going deflection observed over posterior electrodes whose peak latency decreases from 350 ms at 3 months to 290 ms at 12 months of age (Halit et al., 2003). This is consistent with the latency of many prominent ERP components reducing with increasing age during childhood. The results of these studies indicate that at 12 months of age the amplitude of the infant N290, like the adult N170, increases to inverted human but not inverted monkey faces when compared to the upright faces. However, the amplitude of the N290 was not affected by stimulus inversion at an earlier age (3 and 6 months).

By applying new source separation and localisation methods (Richards, 2004) to the infant ERP data, Johnson et al. (2005) identified candidate cortical sources of the face inversion effect. This analysis suggested that generators in the left and right lateral occipital area, the right FFA, and the right superior temporal sulcus (STS) discriminated between
upright and inverted faces in 3- and 12-month-olds. All three cortical areas have been
implicated in the core face processing system in adults (Haxby et al., 2000), and they also
generally overlap with the areas identified in the previously discussed PET study with 2-
month-olds (Tzourio-Mazoyer et al., 2002).

The development of the brain processes reflected in the N170/N290 continues well
beyond infancy (for a review, see Taylor, Batty, & Itier, 2004). Although latency of the adult
N170 is delayed by face inversion, no such effect is observed for the latency of the infant
N290 at any age (de Haan et al., 2002; Halit et al., 2003) and, in fact, this latency effect may
not appear until 8 to 11 years (Taylor et al., 2004). Another developmental change is that
although the amplitude of the adult N170 is larger to the monkey faces, infants’ N290 shows
the opposite pattern. A completely adult-like modulation of the amplitude of the N170 has not
been reported until 13 to 14 years (Taylor et al., 2004). The reason for these changes in the
N170 response profile during mid to late childhood remain unclear.

To date, face processing, besides speech perception (Kuhl, 2004; Dehaene-Lambertz,
Dehaene, & Hertz-Pannier, 2002), represents the best-studied aspect of the infant social brain.
The evidence available from PET and EEG/ERP studies suggests that most of the brain areas
and mechanisms implicated in adult face processing can be activated relatively early in
postnatal life. However, there are some additional effects, such as the activation of the inferior
frontal and superior temporal gyrus in 2-month-olds (Tzourio-Mazoyer et al., 2002), and the
STS generator of the inversion effect found in 3- and 12-month-olds (Johnson et al., 2005),
that do not directly map onto the mature face processing system. In addition to the extra
regions involved while infants perceive faces, another important observation in the infant ERP
work is that the infant face processing system possesses much broader response properties
which are not yet as finely tuned to upright human faces. This suggests that despite the
gradual cortical specialisation seen throughout the first year of life, the system continues to
specialise well beyond infancy. These changes in the degree of specialisation of processing,
and the spatial extent of cortical activation, are in accordance with the interactive
specialisation perspective alluded to earlier (Johnson, 2001).

III. Eye gaze processing

An important social signal encoded in faces is eye gaze. The detection and monitoring
of eye gaze direction is essential for effective social learning and communication among
humans (Bloom, 2000; Csibra & Gergely, 2006). Eye gaze provides information about the
target of another person’s attention and expression and also conveys information about
communicative intentions and future behaviour (Baron-Cohen, 1995). Eye contact is
considered to be the most powerful mode of establishing a communicative link between
humans (Kampe, Frith, & Frith, 2003). From birth, human infants are sensitive to another
person’s gaze as reflected in their preference to look at faces that have their eyes open rather
than closed (Batki et al., 2000).

A. Newborns detect direct gaze: We (Farroni, Csibra, Simion, & Johnson, 2002) tested
newborn infants by presenting them with a pair of stimuli, one a face with eye gaze directed
straight at the newborns, and the other with averted gaze (see Figure 3). Fixation times were
significantly longer for the face with the direct gaze. Furthermore, the number of orientations
was greater with the straight gaze than with the averted gaze. These results demonstrate
preferential orienting to direct eye gaze from birth. The preference is probably a result of a
fast and approximate analysis of the visual input, dedicated to finding socially relevant stimuli
for further processing.

---

insert Figure 3 around here

---

To examine the specificity of this newborn preference, we conducted two further
experiments. Our goal was to see whether inverting faces has any effect on gaze perception in
newborns. If the preference for direct gaze is not found under conditions of inversion, then we
may conclude that low-level aspects of the faces, such as symmetry or local spatial frequency, are not the basis for the newborn preference for direct gaze in upright faces previously observed. Furthermore, an absence of the effect with inversion would indicate that the factors responsible for the preference require the eyes to be situated within the configuration of an upright face. Newborns did not show significant looking time or orientation differences between direct and averted gaze conditions when the faces were presented upside down. These results rule out symmetry and local spatial frequency as possible explanations of the newborn effect.

Two underlying mechanisms could account for the observed gaze preferences in newborns. By one account, even newborns have sophisticated face processing abilities sufficient to extract gaze direction when presented in the context of a face. By an alternative account, the preferences of newborns are based on a primitive “Conspec” detector that responds to an optimal configuration of high-contrast elements. Straight-on faces with direct gaze better fit this simple template than do faces with averted gaze (see Farroni et al. 2002, 2003). To test these hypotheses, we conducted a second experiment that involved similar face stimuli, but with averted head angles. We reasoned that a sophisticated face processing system would be able to extract gaze direction even when head angle varied. In contrast, a simple Conspec mechanism may only produce a preference for direct gaze under conditions where the spacing between eyes and mouth is optimal. Changing head angle alters the relative spacing of the two eyes and mouth, and thus may disrupt the preference seen with a straight head. In this experiment newborns looked equally at the direct gaze and at the averted gaze, and they oriented equally to both. The results of these experiments show that the strong preference for faces with direct gaze depends on the eyes being situated within the context of an upright straight-ahead face. This finding simultaneously rules out many low-level explanations of the original result along with the suggestion that newborns may have sophisticated eye gaze perception abilities. Rather, they are consistent with the view that
newborns orient to direct gaze due to a primitive configuration detection system (such as “Conspec”).

**B. Gaze following:** Eye gaze has also been shown to effectively cue human infants’ attention to spatial locations. Hood et al. (1998) showed that the perception of an adult’s deviated gaze induces shifts of attention in the corresponding direction in 10- to 28-week-olds. In their experiments they modified the standard Posner cueing paradigm (Posner, 1980) using as a central cue the direction of gaze of a woman's face, thus creating a computer generated eye gaze shift (see figure 4 and below for details). Infants’ eye movements toward a peripheral target were faster when the location of the target was congruent with the direction of gaze of a centrally presented face. In subsequent studies, we examined the visual properties of the eyes that enable infants to follow the direction of the gaze (Farroni et al., 2000). We tested 4-month-olds using a cueing paradigm adapted from Hood et al. (1998). Each trial began with the stimulus face eyes blinking (to attract attention), then the pupils shifted to either the right or the left (see Fig 4a). A target stimulus was then presented either in the same position where the eyes were looking (congruent position) or in a location incongruent with the direction of gaze. We found that infants looked faster at the location congruent with the direction of gaze of the face.

---

insert Figure 4 about here

---

In a second experiment, we manipulated the stimulus face so that the whole face was shifted to one side (right or left) while the pupils remained fixed (see Fig. 4b). In this case the infants were faster to look in the direction in which the whole face was shifted, and not the direction where the pupils were directed. Therefore, the infants actually followed the biggest object with lateral motion (i.e., the face) and not the eyes.
In a third experiment, we used the same paradigm as in the first experiment, but this time when the eyes were opened the pupils were already oriented to the left or right, and the infants were not able to perceive the movement of the pupils. In this case the cueing effect disappeared. Up to this point, the results suggested that the critical feature for eye gaze cue in infants is the movement of the pupils, and not the final direction of the pupils.

To try to understand this cueing effect better, we did three further variants of the same procedure (Farroni et al., 2002). In the first experiment in this series we examined the effect of inverting the face on cueing. If infants are merely cued by motion, then an inverted face should produce the same cueing as an upright one. To our surprise, cueing had no effect, suggesting that the context of an upright face may be important. In the next experiment infants saw a face with an averted gaze that then shifted to the centre. If infants are responding just to the motion of elements they should be cued in the direction opposite to the initial gaze. Again, no cueing effect was observed. These results did not support the hypothesis that directed motion of elements is the only determining factor for the cueing effects.

In the last experiment, a more complex gaze shift sequence allowed us to analyse the importance of beginning with a period of mutual gaze: the eyes shifted from centre to averted, and then back to centre. Here we observed a significant cueing effect. Taken together, these results suggest that cueing effects are observed only following a period of mutual gaze with an upright face. In other words, mutual gaze with an upright face may engage mechanisms of attention such that the viewer is more likely to be cued by subsequent motion. In summary, the critical features for eye gaze cueing in infants are (1) lateral motion, and (2) at least a brief preceding period of eye contact with an upright face. Finally, eye gaze has also been shown to effectively cue newborn infants’ attention to spatial locations, suggesting a rudimentary form of gaze following (Farroni et al., 2004a).
C. Neural processing of direct gaze in infants: The question that arose next concerned how the behaviourally expressed preference for mutual gaze and the capacity to follow gaze are implemented in the infant brain. To test infants' sensitivity to gaze direction, we recorded ERPs as 4-month-old infants viewed photographs of faces, some with direct gaze and some with the eyes were averted randomly to the right or left (Farroni et al., 2002). ERPs have been shown to be sensitive to small differences in transient brain activation due to processing of visual stimuli in infants (Csibra, Davis, Spratling, & Johnson, 2000) and they provide the most feasible neuroimaging method to study the brain development of healthy babies. We hypothesized that an early preference for eye contact would facilitate the processing of faces with direct gaze.

---

insert Figure 5 about here

---

In our analyses, we focused on the N290 component described earlier (the precursor to the adult N170). As in previous studies with infants, the N290 component peaked around 240-290 msec post-stimulus. Its amplitude was greater in response to direct gaze than to averted gaze (see Figure 5) with the N290 amplitude being more negative over the medial than over the lateral leads. No differences between left and right hemisphere were observed. By 4 months, the infant brain manifests enhanced processing of faces with direct gaze as indexed by an increased amplitude of the infant N170 when compared to averted gaze (Farroni et al., 2002). This finding is obtained even when the head is averted but direct mutual gaze is maintained (Farroni, et al. 2004b). However, enhanced neural processing to faces with direct gaze is only found when eyes are presented in the context of an upright face. Our results indicated that at this age cortical processing of faces is enhanced with direct gaze under similar conditions to those observed in adults. Overall, these results further the view that
relatively simple perceptual biases in the newborn are replaced by adult-like perceptual processing of gaze within a few months of birth.

Furthermore, source separation and localization methods were used to identify the cortical sources of the scalp-recorded ERP effects sensitive to eye gaze (Johnson et al., 2005). Contrary to adults, who show specific activations associated with eye gaze perception in the STS (Allison, Puce, & McCarthy, 2000), cortical generators localised in the fusiform gyrus discriminated gaze direction best in infants. Although the amplitude of the N170 in infants is modulated by eye gaze (Farroni et al., 2002; 2004b) and face orientation (de Haan et al., 2002; Halit et al., 2003), the amplitude of the adult N170 is only affected by face inversion but not by direction of gaze (Grice et al., 2005; Taylor, Itier, Allison, & Edmonds, 2001). Taken together, the difference in the response properties of the infant N170 (N290) versus the adult N170 suggest that face and eye gaze share common patterns of cortical activation early in ontogeny which later partially dissociate and become more specialised, a developmental change that is consistent with the Interactive Specialisation view introduced earlier (Johnson, 2001).

Another technique that can reveal brain activation missed by averaging methods is the analysis of high-frequency oscillations in the gamma band (20-100 Hz). It is thought that brain oscillations in the high frequency range reflect neural mechanisms by which activity of small neuronal networks is synchronized, whereas very large networks are recruited during slow oscillations (Buzsaki & Draguhn, 2004). Synchronous activity of oscillating networks is a prominent feature of neural activity throughout the animal kingdom (Sejnowski & Paulsen, 2006), and it is viewed as the critical middle ground linking single-neuron activity to behavior (Herrmann, Munk, & Engel, 2004; Engel, Fries, & Singer, 2001). Such oscillations are either time-locked to eliciting stimuli (evoked gamma activity) or can be detected as induced gamma activity consisting of oscillatory bursts whose latency jitters from trial to trial and its temporal relationship with the stimulus onset is fairly loose. Hence, induced gamma activity is not
revealed by classical averaging techniques and specific methods based on time-varying spectral analysis of single trials are required to detect it (Tallon-Baudry & Bertrand, 1999).

As a theoretical framework, which attempts to assign functional significance to early evoked and late induced gamma-band responses, Herrmann and colleagues (2004) have put forward a Match-Utilization-Model (MUM). According to this model, the early gamma-band response reflects the matching of stimulus-related information with memory contents in primary sensory cortex. Once a stimulus has been identified through this matching stage, this information can be used in all kinds of more complex cognitive operations involving other brain areas and the late induced gamma response might be a signature of such a utilization process.

Gamma oscillations are also of special interest because they have been found to correlate with the BOLD response used in fMRI as shown in invasive work with animals (Niessing et al., 2005) and non-invasive studies combining EEG and fMRI in humans (Foucher et al., 2003; Fiebach, Gruber, & Supp, 2005). In this work it also has been shown that whereas fMRI BOLD correlated with gamma-band activity such a correlation was not found for ERP measures (e.g. Foucher et al., 2003). These findings are consistent with a biophysical model which suggests that increases in hemodynamic signals as measured by fMRI are associated with a shift in the spectral mass from low to high frequencies as measured with EEG (Kilner, Mattout, Henson, & Friston, 2005).

Grossmann, Johnson, Farroni, and Csibra (in press) examined gamma oscillations and its relation to eye gaze perception in 4-month-old infants. This study analyzed previously published EEG data sets taken from Farroni et al.’s (2002) and (2004b) ERP studies in which infants were presented with upright images of female faces directing their gaze toward them or to the side. We predicted a burst of gamma oscillation over prefrontal sites to direct gaze if gamma oscillations are indeed related to detecting eye contact/communicative intent as suggested by adult fMRI work (Kampe et al., 2003; Schilbach et al., 2006). Because the right
intraparietal sulcus (IPS) and right STS are sensitive to averted gaze in the adult brain (Hofman & Haxby, 2000) we hypothesized that some activity over right temporo-parietal regions would be associated with the perception of averted gaze. In addition, another group of 4-month-old infants were presented with the same face stimuli upside-down, which is thought to disrupt configural face processing (Rodriguez et al., 1999; Turati, Sangrioli, Ruel, & de Schonen, 2004) and infants’ preference for mutual gaze (Farroni et al., 2006). Thus, we predicted that inverted faces would not induce activity in the gamma band that differs as a function of eye gaze.

The data revealed that gamma oscillations varied as a function of gaze direction in the context of an upright face, which extends previous ERP and source localization results (Farroni et al., 2002, 2004b; Johnson et al., 2005). In support of our hypotheses, specific effects with distinct spatial and temporal characteristics were observed depending upon whether gaze was directed at or directed away from the infant. Direct gaze compared to averted gaze evoked early (100 ms) increased gamma activity (20-40 Hz) at occipital channels. Short-latency phase-locked oscillatory evoked gamma responses have been described in the visual modality in response to brief static stimuli in infant and adult EEG (Csibra et al., 2000; Tallon-Baudry & Betrand, 1999). In adults, evoked gamma activity is significantly larger for items that match memory representations (Herrmann, Lenz, Junge, Busch, & Maess, 2003; Herrmann, Munk, & Engel, 2004). For infants a face with direct gaze may represent a more familiar and prototypical face (Farroni, Massaccesi, Menon, & Johnson, 2007), which is closer to what is represented in memory than a face with averted gaze, and therefore elicits an enhanced evoked oscillatory response. This interpretation is supported by, and might be linked to, findings showing an enhanced neural encoding (Farroni et al., 2002) and better recognition of upright faces with direct gaze in infants (Farroni et al., 2007). The modulation of the evoked gamma response in this study was observed much earlier (100 ms) than the effect in the previous ERP study (290 ms, Farroni et al., 2002, 2004). This indicates
that gamma activity is a more sensitive measure of some aspects of very early brain processes related to the discrimination of gaze direction.

As predicted, direct gaze also elicited a late (300 ms) induced gamma burst over right prefrontal channels. Directing eye gaze at someone (i.e., making eye contact) serves as an important ostensive signal in face-to-face interactions that helps establishing a communicative link between two people. Successful communication between two people may well depend crucially on the ability to detect the intention to communicate conveyed by signals directed at the self such as making eye contact (Kampe et al., 2003). On a neural level, the medial prefrontal cortex (MPFC) is consistently activated when gaze is directed at, but not when gaze is averted away from, the self (Kampe et al., 2003; Schilbach et al., 2006). Because gamma oscillations measured with EEG are correlated with the BOLD response used in fMRI (Fiebach et al., 2006; Foucher et al., 2003), eye contact detection in 4-month-old infants may well recruit very similar brain mechanisms as in adults. The gamma burst distributed over right frontal cortex in infants might reflect less localized functional activity than in adults, which is in accordance with current views of functional brain development (Johnson, 2001), suggesting a more diffuse to a more focal pattern cortical activity with age.

Averted gaze also serves an important function during social communication by directing the perceiver’s attention to certain locations or objects, and behavioral evidence indicates that 4-month-olds are sensitive to this aspect of eye gaze (Farroni et al., 2003; Hood et al., 1998). The right IPS and right STS have been identified as sensitive to averted gaze in the adult human brain (Haxby, Hofman, & Gobbini, 2000; Hofman & Haxby, 1999). Activity in the IPS may be specifically recruited when perceived eye gaze direction elicits a shift in spatial attention, whereas STS is more generally associated with eye and mouth movements (Haxby et al., 2000). Our finding of a late (300 ms) induced gamma burst in response to averted gaze over right occipito-temporal-parietal regions might reflect similar but perhaps more diffuse brain activations in infants.
In another study (Grossmann, Johnson, Farroni, & Csibra, unpublished), we further examined how head orientation would influence the infants’ brain responses to eye gaze direction cues observed in the gamma band. This study was based on a time-frequency analysis performed on previously published EEG data (Farroni et al., 2004b). Infants were presented with upright images of female faces orienting their head away from the infant but either directing their gaze towards them or away from them. Corresponding with the findings reported for straight head angle faces, direct gaze compared to averted gaze elicited early (100 ms) increased gamma activity (20-40 Hz) at occipital channels. However, contrary to the previous findings (Grossmann et al., 2007), no induced gamma burst was observed over prefrontal channels to direct gaze when the head was averted.

This suggests that although infants at the age of 4 months can discriminate between direct and averted gaze in the context of averted heads as indicated by the increased evoked occipital gamma activity to direct gaze, in this context, they do not yet recruit brain processes associated with detecting eye contact as a communicative signal. Hence, it follows that by 4 months a frontal face is required to elicit activity in prefrontal brain structures involved in social communication. This finding stands in contrast to fMRI work showing that adults show specific activity in the prefrontal cortex in response to direct gaze even when the head is averted from the perceiver (Kampe et al., 2003) and therefore suggests that development must occur after 4 months which enables the human brain to detect mutual gaze from static face representations regardless of head orientation. In a follow-up study (Grossmann, Johnson, Farroni, & Csibra, 2007), we were able to show that when infants at this age are presented with dynamic eye gaze cues directed at them in the context of an averted head a burst of prefrontal gamma activity was observed. This suggests that biological motion cues (i.e., gaze shifts) help infants to detect communicative signals and recruit frontal brain regions, which play a fundamental role for the development of social perception and cognition (Anderson et al., 1999).
Furthermore, as in the previous study using faces with a frontal orientation (Grossmann et al., in press), we observed a late (300 ms) induced gamma burst over right temporo-parietal regions (Grossman et al., 2007), but whereas when frontal faces were used this burst was evoked by averted gaze, for oriented faces this activity occurred in response to direct gaze. These seemingly contradictory findings need explanation. Our suggestion is that the gamma activity observed over right temporo-parietal channels is associated with neural computations integrating eye gaze direction information in relation to head angle. When the results are revisited from this perspective, it appears that this gamma burst is observed when eye direction is different/incongruent from the head direction (i.e., averted gaze in a frontal face and direct gaze in an oriented face). This view is in accordance with adult fMRI data showing increased activity in the right STS to averted gaze in frontal face (Hoffman & Haxby, 1999) and to direct gaze in averted face (Pelphrey et al., 2004).

The finding that inverted faces did not elicit gamma band responses that differed between direct and averted gaze is in line with, and adds further developmental evidence to, the notion that face inversion disrupts face processing (Rodriguez et al., 1999; Turati et al., 2004). This indicates that relatively early in development cortical structures involved in face processing are already somewhat specialized to extract information about gaze direction from upright faces. It further shows that the gamma band effects observed in response to direct and averted gaze are not simply driven by ‘lower level’ perceptual parameters (e.g., symmetry [direct gaze] and asymmetry [averted gaze]) because then they should have occurred in the inverted condition as well. These gamma band findings in infants show a high degree of correspondence in terms of timing and frequency content with previous findings in adults (Tallon-Baudry & Bertrand, 1999). This suggests continuity throughout development, and further underlines the functional importance of gamma band oscillations also for social perception.
D. Understanding eye-object relations: Another communicative function of eye gaze is to direct attention to certain locations, events, and objects. Understanding the relations between eye gaze and target objects is particularly important for aspects of development such as word learning. Comprehending that another’s gaze direction refers to a specific object allows the child to associate the object with a name or emotional expression (Baldwin & Moses, 1996). Adults’ gaze has been found to facilitate object processing at a neural level in infants as young as 4 months (Reid, Striano, Kaufman, & Johnson, 2004). In this ERP study, objects that were previously cued by eye gaze elicited a diminished positive slow wave observed between 700 and 1000 ms over right fronto-temporal channels. A diminished positive slow wave is thought to indicate deeper memory encoding (Nelson & Collins, 1991). This suggests that eye gaze as a social cue facilitates brain processes involved in memory encoding that might assist infants’ learning.

In another ERP study, 9-month-old infants and adults watched a face whose gaze shifted either towards (object-congruent) or away from (object-incongruent) the location of a previously presented object (Senju, Johnson, & Csibra, 2006). This paradigm was based on that used in an earlier fMRI study (Pelphrey, Singerman, Allison, & McCarthy, 2003) and designed to reveal the neural basis of “referential” gaze perception. When the ERPs elicited by object-incongruent gaze shifts were compared to the object-congruent gaze shifts, an enhanced negativity around 300 ms over occipito-temporal electrodes was observed in both infants and adults. This suggests that infants encode referential information of gaze using similar neural mechanisms to those engaged in adults. However, only infants showed a fronto-central negative component that was larger in amplitude for object-congruent gaze shifts. Thus, in the less specialised infant brain, the referential information of gaze may be encoded in broader cortical circuits than in the more specialised adult brain.

In summary, infants show specific modulations of cortical responses associated with eye gaze perception, which differ from what is seen in adults. Nevertheless the findings reviewed
also suggest that they successfully discriminate between direct and averted gaze, and can make use of gaze cues when encoding objects in memory. This indicates that, although the cortical structures involved in the perception of gaze direction may not be fully differentiated in infancy, they are at least partially functional. This may be critical for social communication and learning from others (see Discussion).

IV. Perception of emotions

Social behaviour is coupled with emotion. During social interactions, emotions are overtly expressed by the infant, whereas displays of emotion are normally more regulated in adults. Nevertheless, most brain structures that are important in processing emotions in adults are also essential for social behaviour (for a review, see Adolphs, 2003). Detecting emotion in others through their facial expressions is an important way to communicate emotions in social interactions (Izard, 1991). Recognising facial expressions permits us to detect another person’s emotional state and provides cues on how to respond in these social interactions.

A. Processing emotion in the face. In an ERP study in which 7-month-olds watched happy versus fearful faces, (Nelson & de Haan, 1996), fearful faces elicited an enhanced negative component (Nc) peaking around 500 ms. The Nc has its maximum at frontal and central sites and has been thought of as an obligatory attentional response sensitive to stimulus familiarity (Courchesne, Ganz, & Norcia, 1981; Snyder, Webb, & Nelson, 2002). Dipole modeling has revealed that the cortical sources of the Nc can be localised in the anterior cingulate and other prefrontal regions (Reynolds & Richards, 2005). This suggests that 7-month-old infants in Nelson and de Haan’s study allocated more attentional processing resources to the unfamiliar fearful than to the familiar happy expression as indicated by an enhanced Nc.

Another important question is whether infants can discriminate between different negative emotional facial expressions. De Haan and Nelson (1997) found no difference in 7-month-olds’ ERP responses to fearful and angry faces and suggested that infants did not
display different brain responses to the two expressions because they perceived the signal value of both expressions as “negative” or because they perceived both expressions as equally unfamiliar. In a subsequent study (Kobiella, Grossmann, Reid, & Striano, in press) using another set of facial stimuli and measuring from more electrode positions than de Haan and Nelson (1997), infants’ N290, P400 and Nc differed between emotions (for a discussion of how the neural processing of angry in contrast to fearful facial expressions develops during infancy, see Grossmann, Striano, & Friederici, 2007). This suggests that infants discriminate between emotions of negative valence, which confirms findings that infants at this age discriminate between the angry and fearful expressions behaviorally (Serrano, Iglesias, & Loeches, 1995). It further suggests that the infant N290—which is viewed as the developmental precursor to the adult N170—is sensitive to emotional information, whereas the adult N170 does not show an emotion-dependent modulation (see also, Leppänen et al., 2007). This suggests that early in life the N170 is sensitive to other aspects of face processing such as eye gaze (Farroni et al., 2002) and emotional expression (Kobiella et al., in press; Lapanen et al., 2007), a notion consistent with the Interactive Specilization view of social brain development (Johnson, 2001, 2005).

Processing of facial expressions has been studied as a function of particular experiences such as early institutional rearing (Parker, Nelson, & the Bucharest Intervention Core Group, 2005) and maternal personality (de Haan et al., 2004). In general, these studies suggest that experiential factors influence the ways that infants process facial expressions. It is important to note that the reverse may also be true—that is, neural development occurring at the end of the first year (Diamond, 1991, 2000; Johnson, 2001, 2005) may impact infant behaviour and subsequently infants’ experiences with others. Although the control of infants’ orienting is partly in the hands of the caregivers’ presentation of relevant information, the infant is clearly involved in soliciting attention and information from adults (Baldwin & Moses, 1996; Stern, 1985). With the development of the anterior attention system during the
first year of postnatal life, more direct control of attention passes from the caregiver to the infant (Bush, Luu, & Posner, 2000; Posner & Rothbart, 2000), which might influence and regulate infants’ emotional exchanges and experiences.

B. Processing emotion in the voice. Emotion produces not only characteristic changes in facial patterning but also in respiration, phonation, and articulation, which in turn determine the acoustic signal (Scherer, 1989). Thus speech melody, also referred to as emotional prosody, serves as a signal for various emotions. Adult listeners across cultures can reliably and readily recognise different emotions on the basis of vocal cues (Banse & Scherer, 1996; Scherer, Banse, & Wallbott, 2001) and do so using some of the same brain structures used to recognise facial expressions (Adolphs, Tranel, & Damasio, 2002). How the processing of emotional prosody develops over the course of ontogeny is only poorly understood. Unlike the extensive behavioural work on processing facial expressions in infancy, studies on infants’ perception of vocal expressions of emotion are relatively scarce (Walker-Andrews, 1997).

In one of the few studies of infants’ processing of emotional prosody (Grossmann, Striano, & Friederici, 2005), infants heard semantically neutral words spoken with either a happy, angry, or neutral voice. Angry prosody elicited a negative component peaking around 450 ms over fronto-central sites in infants’ ERPs, suggesting greater allocation of attention to the angry prosody. This is consistent with the notion of an evolutionarily driven propensity to react more strongly to negative than to positive or neutral stimuli (Cacioppo & Berntson, 1999). A positive slow wave (700-1000 ms) was elicited by angry and happy prosody over temporal electrode sites, whereas words with neutral prosody returned to baseline. This indicates an enhanced processing in auditory (temporal) brain structures of the emotionally loaded stimuli (happy and angry), which is in line with recent fMRI work on adults’ processing of emotional speech (Grandjean et al., 2005; Mitchell et al., 2003). These findings suggest that very early in development, the human brain detects emotionally loaded words and shows differential attentional responses depending on their emotional valence. The three
emotional speech categories used in the Grossmann et al. (2005) study significantly differed in duration and fundamental frequency but did not differ with respect to their mean intensity. Effects due to these perceptual differences, however, are likely to be limited to early (exogenous) ERP components (earlier than 200 ms). Therefore, acoustic differences alone cannot account for the late ERP effects observed in response to the various emotions.

C. Integration of emotional information from face and voice. The way that emotions are perceived when communicated either by the face or the voice has been studied in each modality separately. However, in most social interactions, emotional information is communicated simultaneously by different modalities such as the face and voice. Thus the question arises: how is emotional information from the face and voice integrated? To address this question, processing of emotionally congruent and incongruent face-voice pairs has been studied using ERP measures (Grossmann, Striano, & Friederici, 2006). In this study, infants watched facial expressions (happy or angry) and, after a delay of 400 ms, heard a word spoken with a prosody that was either emotionally congruent or incongruent with the facial expression being presented. Emotionally congruent face-voice pairs elicited similar ERP effects as recognised items in previous memory studies with infants, children, and adults in which a picture (visually presented)-name (acoustically presented) matching paradigm was used (Friedman, 1991; Friedman et al., 1991, 1992; Nelson, Thomas, de Haan, & Wewerka, 1998; Rugg & Coles, 1995). Thus, the ERP effects in Grossmann et al.’s (2006) study provides electrophysiological evidence that 7-month-olds integrate emotional information across modalities and recognise common affect in the face and voice, a finding that is in line with previous behavioural studies (Soken & Pick, 1992; Walker, 1982; Walker-Andrews, 1986).

Because the face-voice pairs presented to the infants were novel to them, these findings not only indicate that infants recognised common affect but, moreover, that they applied their knowledge about emotions in face and voice to draw inferences about what might be
appropriate emotional face-voice pairings. Extending previous behavioural findings, the ERP
data revealed insights into the time course and characteristics of the processes underlying the
integration of emotional information from face and voice in the infant brain. Emotionally
incongruent prosody elicited a larger fronto-central negative component (400-600 ms) in
infants’ ERPs than did an emotionally congruent prosody, indicating increased allocation of
attention. Conversely, the amplitude of infants’ positive component (600-1000 ms), with a
maximum over parietal channels, was larger to emotionally congruent than to incongruent
prosody, indexing recognition of the congruity between face and voice (Grossmann et al.,
2006).

Interestingly, the body of evidence presented here shows that effects of emotion, in
contrast to the ‘early’ effects elicited by face and gaze manipulations alone, are reflected in
specific modulations of components that occur relatively ‘late’ in the infants’ ERP (> 300 ms;
in contrast, some emotional expressions can have effects at very short-latency ERP
components in adults, Eimer & Holmes, 2002). This might be related to the fact that learning
to identify an emotion requires more complex cognitive and brain operations, which integrate
various perceptual elements provided by face and (or) voice, and retrieve specific information
about the emotion stored in memory. The studies that have been conducted so far have only
looked at a limited set of emotions (happiness, fear, and anger). This work should be extended
to other basic (disgust, sadness, surprise) and complex (social) emotions in order to
understand the emotion-specificity of the effects summarized here.

V. Interactions between face identity, eye gaze, and emotion

Following from the two theoretical frameworks that have motivated our review so far,
several specific hypotheses emerge with regard to the interaction, or lack of it, between the
different aspects of social brain function reviewed above. In this section we begin by
reviewing evidence in adults of partially independent processing streams for processing
different kinds of information from the face (identity, eye gaze, emotion). Interaction
between these streams appears to occur in a restricted and task-dependent way. We then go on to review evidence from infants in support of the prediction from the interactive specialisation view that as the social brain network develops processing streams will fractionate and become increasingly specialised and dissociable. A specific prediction is that infants may have a common processing stream for all aspects of face processing.

A. Dissociable neural pathways for processing faces. In adults, emotion, eye gaze, and identity are processed by partially independent neural routes, and particular structures become specialised for these different computations. One driver for these differential patterns of specialization may be that different, and even conflicting, invariances need to be extracted from the same basic visual input. For example, recognizing the identity of a face requires the invariant aspects of facial structure configuration to be extracted from the dynamic changes associated with eye gaze shifts and emotional expressions. An example of the neural dissociation of these pathways comes from neuropsychological studies that provide evidence that specific brain damage can impair the perception of facial expressions without impairment in detecting face identity (Kurucz and Feldmar, 1979) or the face inversion perception (Bruyer et al., 1983). Furthermore, neuroimaging studies provide evidence of a spatial and temporal dissociation between face identity and emotion perception (Munte et al., 1998; Sergent, Ohta, MacDonald & Zuck, 1994). Finally, different facial expressions may be processed by different neural networks that involve both cortical regions (prefrontal, frontal and orbital frontal cortex, occipital temporal junction, cingulated cortex and sensorial motor cortex) and subcortical structures (amygdala, insula and basal ganglia) (Damasio et al., 2000; Gorno-Tempini et al., 2001; Kesler-West et al. 2001; Lane et al., 1997; Streit et al., 1999).

B. Interactions between face processing pathways in adults. Although there are dissociable routes for different aspects of face processing in adults, there are also specific interactions between these streams of processing. These interactions can be specifically restricted to particular stages of processing, and can be specific to particular emotions,
directions of eye gaze, or identities.

We begin by considering the effects of gaze direction on other aspects of face processing. George et al. (2001) investigated how gaze direction (direct or averted), influences face processing in adults using a gender recognition task. They presented a face with direct or averted gaze, and the face was either a frontal view or tilted at 45 degrees. Specific regions of the fusiform gyrus yielded stronger responses to faces when these looked directly at the participant (regardless the orientation of the head). Because the fusiform response can be sensitive to recognition of individuals (Gauthier et al., 2000, 1999; George et al., 1999; Sergent et al., 1994, 1992), and is stronger for attended faces (Wojciulik et al., 1998), the authors concluded that the stronger activity found for faces with direct gaze could be interpreted in terms of enhanced attention and deeper encoding for faces with direct gaze (George et al. 2001).

Furthermore, it has been previously argued that direct gaze is likely involved in the communication of increased emotional intensity, regardless of the particular emotion being expressed (Kimble & Olszewski, 1980; Kleinke, 1986). Despite the intuitive appeal of the long-standing notion that direct gaze may facilitate emotional expression processing, brain and behavioural research with adults has shown that the way in which gaze direction influences emotion perception depends on the emotion in question (Adams & Kleck, 2003, 2005; Adams et al., 2003). There are some reasons for supporting an interaction between eye gaze and emotional expression processing. First, analysis of facial expression and gaze direction may be partly processed by the same anatomical structures. Areas in the superior temporal sulcus are critical in the processing of changeable features of faces, such as expression and gaze direction (Allison et al., 2000; Hoffman & Haxby, 2000; Perrett et al., 1992). Neuroimaging studies have also shown that the amygdala plays a central role in processing emotional facial expressions, particularly fear expression and fearful eyes, (Adolphs, Tranel, Damasio, & Damasio, 1995; Breiter et al., 1996; Whalen et al., 1998, 2001)
and monitoring gaze direction (George et al., 2001; Wicker et al., 2003). Adams et al. (2003) also highlighted amygdala’s involvement in the combined processing of facial expression and gaze direction. Namely, they demonstrated that the amygdala is differentially responsive to anger and fear expressions as a function of gaze direction.

Both eye gaze and emotional expression are reported to elicit rapid and automatic spatial orienting (Driver et al., 1999; Öhman et al., 2001). As mentioned previously, eye gaze can trigger orienting in automatic fashion, in that individuals shift their attention to the location of a looked-at target in a fast and efficient manner, even when eye gaze direction is uninformative and even counterpredictive of the target location (Driver et al., 1999; Friesen and Kingstone, 1998, 2003). As regards emotion, many behavioural observations indicate that people more readily pay attention to emotional than neutral stimuli. These effects also arise in a reflexive or involuntary manner. In visual search tasks where a unique target must be found among distractors, detection times are faster when the target has some emotional values, such as an angry or happy face among neutral faces (Eastwood et al., 2001; Fox, 2002), or a snake or spider among flowers (Ohman et al., 2001). Furthermore, fearful faces presented peripherally trigger orienting of attention toward their location, thus facilitating target discrimination at this location, more efficiently so than neutral or happy faces do (Pourtois et al., 2004).

Finally, another reason for supporting that emotional faces may promote attention to gaze direction is that one could expect that an expressive face (e.g., happy or fearful) gazing in a certain direction would be a more powerful attention orienting stimulus than a comparably gazing neutral face. Such a stimulus would provide information about another individual’s direction of attention and, in addition, about the emotional significance of the object or person he or she is attending to. A smiling face with an averted gaze would inform the observer that the other person is looking at something nice or funny, whereas a comparably gazing fearful face would suggest the presence of something threatening. So,
logic suggests that the gaze of an emotional face should be a more effective cue to attention. This should be particularly true for fearful faces. Because another’s fearful gaze often signals a source of danger, learning and/or natural selection may have particularly favoured attention to the same location. By comparison, averted gaze in happy faces may indicate sources of reward important for interpersonal communication, but it has less immediate relevance to personal safety or survival.

We have seen that in adults there are complex and restricted interactions between the different streams of face processing. Interactions can either can early in processing, such as modulation by a fearful expression, or after the independent parallel processing of identity, gaze and expression. These interactions are selective with, for example, specific emotions interacting with specific directions of eye gaze.

C. Interacting face pathways in infants. A specific prediction of the Interacting Systems model is that dissociable pathways involving specialized cortical structures in adults may start life as un-differentiated highways of common processing in infants (Johnson 2001, 2005; Johnson & Munakata, 2005). If this is the case we should expect neural and perceptual interactions between identity, eye gaze and emotion. However, in contrast to adults, this interaction may be general and non-specific. In Farroni et al. (2006), we examined how facial identity and gaze direction may interact early in infancy. The question was whether direct gaze evokes a deeper analysis of other aspects of the person’s face such as its identity. Specifically, we predicted that infants who had previously seen a face accompanied by direct gaze would subsequently demonstrate recognition of that face by displaying a novelty preference for an unfamiliar face. In contrast, we predicted no preference when infants initially see a face with averted gaze. These predictions were confirmed. In addition, post-hoc analyses revealed that the order in which the two conditions were presented made a difference. If the experiment began with the direct gaze condition then the averted gaze condition also elicited evidence of subsequent individual recognition. In contrast, if the
experiment began with the averted gaze condition, then only the later occurring direct gaze condition resulted in recognition. These effects were not due to differences in looking time between the conditions during the habituation phase. These results confirm the hypothesis that direct gaze is an important modulator of face and social information processing early in life. We concluded that from at least 4 months of age infants show facilitation of cueing of spatial attention, improved individual recognition, and enhancement of neural correlates of face processing, when accompanied, or preceded by, direct gaze.

The question of how eye gaze interacts with emotional expression has also been examined in young infants using ERP measures (Striano, Kopp, Grossmann, & Reid, 2006). In this study, 4-month-old infants were presented with neutral, happy, and angry facial expressions when accompanied with direct and averted gaze. Neural processing of angry faces was enhanced in the context of direct gaze as indexed by an increased amplitude of a late positive fronto-central ERP component, whereas there was no effect of gaze direction on the processing of the other expressions.

This neural sensitivity to the direction of angry gaze can be explained in two different ways. First, infants at this age who have very rarely seen angry faces (Campos et al., 2000) might be primed to detect novel/unfamiliar emotions that are directed toward them. Second, infants may detect the potential threat conveyed by an angry face with gaze directed at them, which is reflected in an enhanced neural response. To find out which explanation (‘novelty’ versus ‘threat’) can account for these ERP findings, it would be necessary to examine the processing of a facial expression that is equally unfamiliar to the infant but whose perception is enhanced in the context of averted gaze; one such emotion is fear (Adams et al., 2003, Adams & Kleck, 2005). Nevertheless it is important to note that by using ERP measures it was possible to show that the neural processing of facial expression is influenced by the direction of eye gaze.
In summary, converging lines of evidence suggest more generalized interactions between different aspects of face perception (identity, eye gaze, emotion) in infants than in adults. These generalized interactions may be indicative of less differentiated processing streams for these functions in infants than in adults. Greater common processing of the different computations applied to faces in early development is a prediction of the Interactive Specialisation approach.

VI. Conclusions

In this chapter we have illustrated how adopting a developmental cognitive neuroscience approach sheds light on how the social brain network emerges during infancy. We believe that the combination of neuroscience, cognitive and behavioral studies has significantly advanced our understanding in this foundational area of neuroscience. Our work, and that of other labs, has been presented within two frameworks; the two-process model of the development of face-processing originally presented by Johnson & Morton (1991; Morton & Johnson 1991) and the interactive specialization model of functional brain development (Johnson 2001, 2005b).

With regard to the two-process model we reviewed a number of studies of newborn face-related preferences, most of which supported the view that newborns have a bias to orient toward faces in their natural visual environment. Although the exact mechanisms that underlie this bias remain the topic of some debate, the proposal that best accounts for the majority of the data currently available is that there is a “quick and dirty” sub-cortical route for face detection that is activated by a face (or eye)-like phase contrast pattern within a bounded surface or object (Johnson, 2005a). Interestingly, some evidence suggests a broader role for Conspec than originally envisaged because strong evidence from adults reveals that the “quick and dirty” sub-cortical route modulates processing in cortical regions within the social brain network. This indicates that the mechanisms that underlie the orienting toward,
and foveating of, faces with direct gaze in young infants also facilitate the activation of relevant cortical regions, providing an important foundation for the emerging social brain.

Turning to the development of social perception over the early months and years of life, we have reviewed evidence broadly consistent with predictions from the Interactive Specialisation approach. According to this view, most parts of the social brain network can be activated in infants, though activation may also extend to other regions not activated under these circumstances in adults. Furthermore, the social brain regions activated may have broader functions (be less finely tuned) than in adults. Some of the evidence consistent with this “Interactive Specialisation” view is that, compared to adults, infants activate regions in addition to, and surrounding, the core face processing network. Furthermore, face and eye gaze perception have been shown to share common patterns of cortical activation early in ontogeny, which later partially dissociate and become more specialised (Farroni et al., 2002; 2004; Grice et al., 2005; Johnson et al., 2005; Taylor et al., 2001). These findings support the view that structures in the social brain network initially have more homogeneous response properties, with common processing of many aspects of faces, bodies, and actions. With experience, these structures may become more differentiated and specialised in their response properties, finally resulting in the specialised patterns of activation typically observed in adults.

This view has implications for atypical development in that some developmental disorders that involve disruption to the social brain network, such as autism, may be characterised in terms of failures or delays of the specialisation of structures on the cortical social brain network (see Johnson et al., 2005, for further discussion). For example, Grice et al. (2005) found ERP evidence consistent with common processing of eye gaze and other aspects of face perception in young children with autism, at an age at which there is evidence for different streams of processing having emerged in typically developing children.
We have described several neuroimaging methods that have been used to study different aspects of social information processing in infants. The available evidence on the neural processes related to face, gaze, emotion, biological motion, action, and joint attention discussed revealed how the infant brain processes information about the social world. However, many areas of infant social cognition, such as imitation, social (complex) emotions, and ‘theory of mind’ remain unexplored (for recent behavioural studies on infant theory of mind, see Onishi & Baillargeon, 2005; Southgate, Senju, & Csibra, in press; Surian, Caldi, & Sperber, in press). Future research is needed to examine the neural correlates of these more complex aspects of social cognitive development, and it is very likely that this kind of work will reveal greater differences between adult and infant/child brain function.

Another important issue is that although the cross-talk between developmental psychologists and social cognitive neuroscientists has begun on a theoretical level (Decety & Sommerville, 2004; Meltzoff & Decety, 2003), there is very little infant brain research that is more directly informed and motivated by already existing theories of infant social cognitive development (Csibra & Gergely, 2006; Meltzoff, 2002, 2005; Tomasello et al., 2005). The available theoretical frameworks explaining the developmental trajectories of social cognitive capacities provide a rich source of hypotheses that are testable using the neuroimaging tools. In this context it will be of particular importance to identify the neural processes that underlie known social behavioural and social cognitive transitions. Thus, further progress in the developmental cognitive neuroscience of the social brain network crucially depends upon a closer integration of human functional brain development with theories of social cognitive development.
References


Figure captions

Figure 1. Examples of schematic face stimuli used in the newborn experiments.

Figure 2. Newborns’ looking time measured in response to schematic face stimuli which differed in their contrast polarity (Figure adapted from Farroni et al., 2005).

Figure 3. Results of the Farroni et al. (2002) preferential looking study with newborns. (a) Mean looking times spent at two stimulus types. (b) Mean number of orientations toward each type of stimulus. (c) Filled triangles represent reference score for the direct gaze over the averted for each individual newborn and open triangles represent average preference scores.

Figure 4. Example of the edited video image illustrating the stimulus for Experiment 1 in Farroni et al. (2000). In this trial the stimulus target (the duck) appears on the side which is incongruent with the direction of gaze.

Figure 5. Four-month-old infants’ ERPs recorded to faces with direct and averted gaze in Farroni et al. (2002). An enhanced N170 component (peaking around 240 ms after stimulus onset) was observed over occipital cortex in response to faces with direct gaze when compared to faces with averted gaze.
Figure 1
Figure 2
Figure 3
Figure 5