SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Psychology

PhD program in *Psicologia Clinica, dello Sviluppo e Neuroscienze cognitive* Cycle: *XXIX°*

Curriculum in *Mente, Cervello e Comportamento*

Processing of human gestures and emotions early in life.

Surname: **Addabbo** Name: **Margaret**

Registration number: **035053**

Tutor: **Chiara Turati**

Coordinator: **Maria Teresa Guasti**

ACADEMIC YEAR 2015/2016

Index

Abstract

From very early in life, infants start developing some understandings about the meaning of others' gestures and emotions. The overall aim of this thesis is to provide new insights in the field, striving to increase our understanding on processing of gestures and emotions early in life. In Section 1, we have highlighted the fundamental role of sensory-motor experience in the development of early sensitivity to human gestures. We have shown that biomechanically possible grasping gestures, but not impossible gestures, can trigger 6-month-olds shifts of attention as reflected by modulation of the amplitude of the early N290 ERP component (Chapter 1). Also pre- and postnatal sensorimotor experiences seem to play a crucial role in gesture understanding. In fact, 2-day-old newborns are able to discriminate between touching and non-touching gestures, manifesting a visual preference to the former than to the latter. Moreover, at 3 months of age spontaneous preferential responses change possibly as a result of further experiences accumulated during development (Chapter 2). Further, newborns can successfully match familiar and experienced facial gestures (i.e., yawns and hiccups) to the corresponding sound both in the presence and absence of temporal synchrony (Chapter 3). Findings in Section 2 shed new light on emotional processing in infancy. Specifically, our results have shown that already at birth newborns are able to differentiate between dynamic facial expressions of happiness and disgust (Chapter 4) and, at 6 months of age, the emotional context can modulate at the neural level the processing of an action (Chapter 5).

General introduction

Humans are social creatures and, from very early in life, they are surrounded by a complex social world full of stimuli coming from different senses and sources. Since birth, we devote a great deal of time attending to faces, bodies, and actions. Interpretation of others' facial expressions, gestures and actions results, thus, foundational to learn about the social world. Social understanding includes different fundamental capacities through which we can make sense of the social events that surround us. Among these, there are two abilities that play a distinctive role in social understanding and that will be the main focus of the present thesis: i) the ability to understand others' gestures and to use this information to predict where people are focusing their attention in the environment; ii) the ability to detect emotions and to use this information to predict others' behavior.

Studying how these abilities develop is fundamental to shed light on the mechanism behind the capacity to process social information. The main questions that arise are: how the ability to process the social world develops, and which are the developmental precursors of adult capabilities in social processing?

1. Processing the social world.

From the earliest stages of life, we are immersed in a world filled with people and one of the major functions of our cognitive system is to allow us to interact with them. Foundational to this ability is to recognize socially relevant information coming, for example, from faces, voices and gestures and to make use of it to understand others' behavior and intentions. Attending to social information allows us to learn from other humans and promotes attachments with others. Accordingly, being able to relate with others is essential for a healthy development and is fundamental to achieve an optimal functioning in life.

Several findings support the hypothesis that there is a network of regions in the adult cortex preferentially involved in social perception, the so-called 'social brain'. This 'social' network includes the superior temporal sulcus (STS), the fusiform face area (FFA) and the orbitofrontal cortex (for a review see Adolphs, 2003). Three distinct frameworks tried to explain the origins of the 'social brain'. *Maturational viewpoints* relate brain development to the coming on-line at different post-natal ages of particular regions of the cortex. Differently, according to the *skill learning perspective,* some brain regions are recruited to process social information because humans become experts in that domain. This viewpoint argues that the brain regions active at the onset of new simple abilities in infancy are identical to the ones involved at the onset of new, but more complex skills, in adulthood (Gauthier & Nelson, 2001). Lastly, a third framework called *interactive specialization* posits that during development some brain regions become more specialized to process certain type of information, like social stimuli. According to this perspective, the acquisition of a new ability in infancy reflects changes in the activity over networks involving different brain regions (Johnson, Grossman, & Cohen Kadosh; Johnson, 2001).

Thanks to the work that has been done in the last decade to explore social information processing in infancy, to date appears clear that from very early in life infants are already attuned to the social world. The purpose of the present thesis is to explore the earliest precursors of social understanding at birth and how the brain deals with social information in infancy within two main domains of social cognition: perception of gesture and emotion processing.

1.1 Perception of human gestures.

Fundamental to social understanding is the ability to predict others' behavior. To do so, we need first of all to identify and select a social stimulus and relevant agents in the world that surrounds us, understand where is the focus of their attention and, accordingly, infer their intentions.

1.1.1 Newborns' early attunement to social information.

Social information coming from the body, head, hands and the face exerts a key role in understanding others' behaviors because from observing them, we readily detect where others' are directing their attention in the environment. From birth, we already possess some early sensitivities to our conspecifics (Johnson & Morton, 1991; Simion, Regolin & Bulf, 2008; Craighero et al., 2011; Longhi et al., 2015). Thus, from very early in life we can detect relevant social information essential to develop the ability to predict others' actions. One of the best-studied aspect of social understanding in infancy is face processing. Converging evidence has shown that at birth newborns attend preferentially to human faces (Morton & Johnson, 1991). Further, newborns' attention is also biased towards familiar faces, like the mothers' face (Sai et al., 2005). According to Johnson (2005) these early preferences towards faces are guided by subcortical brain routes that direct attention to social stimuli (Johnson et al., 2015). Detecting biological motion of conspecifics is also a crucial ability when we have to deal with the social world. Behavioral findings with newborn babies have shown that shortly after birth newborns can discriminate biological motion from other forms of motion (Simion et al., 2008). More recently, evidence has shown that newborns seem to be sensitive also to goal-directed actions (Craighero et al., 2010) and hand movements (Longhi et al., 2015). Specifically, neonates seem to be able to discriminate goal-direct from non-goal-direct movements, showing to prefer the former to the latter (Craighero, Leo, Umiltà, & Simion, 2011). Moreover, at birth newborns can discriminate between biomechanically possible and impossible whole hand grasping movements (Longhi et al., 2015). These findings led to the intriguing suggestion that newborns' early discrimination abilities of goal/non goal-directed actions and possible/impossible hand movements could be based on the sensorimotor knowledge they have accumulated during pre- and early postnatal life (Craighero et al., 2011; Longhi et al., 2015). Recent challenging findings have shown that already within the confines of the womb fetuses perform coordinated and organized hand movements (Myowa-Yamakoshi & Takeshita, 2006; Zoia, Blason, D'Ottavio, Bulgheroni, Pezzetta, Scabar, & Castiello, 2007). For instance, more than 50% of fetal arm movements are directed towards the mouth (Myowa-Yamakoshi & Takeshita, 2006) and fetuses open their mouth before their hands reach it, thus anticipating hands' arrival (Myowa-et al., 2006; Reissland, Francis, Aydin, Mason & Schaal, 2013). Furthermore, by 22 weeks of gestation fetuses seem to coordinate their arm movements towards the facial region by modifying the kinematic according to the goal of the action (i.e., the mouth or the eye) (Zoia et al., 2007). These early sensorimotor experiences might have provided the basis for the come into place of primitive forms of sensorimotor association at birth, thus allowing newborns to develop early sensitivities to visual cues that hint both the presence of goal-directed movements and the violation of the biomechanical constrains of human hands. However, further evidence is needed to support the existence of a link between prenatal sensorimotor experience and early postnatal perceptual abilities. First of all, it might be of great interest to explore whether pre- and postnatal experience can affect early visual preferences across the first moths of life. This issue will be assessed in Chapter 2 of section 1, by testing infants' visual preferences to touching gestures at birth and at 3 months of life. Secondly, it would be crucial to explore if early perinatal experiences could facilitate newborns' early perceptual abilities at birth. This issue will be tackled in Chapter 3 of section 1, by testing newborns multisensory abilities using stimuli that are part of their own sensorimotor repertoire, like hiccups and yawns.

1.1.2 Following the direction of human gestures.

Following the direction of others' gaze, body orientation and gestures represents another distinct and essential process to action prediction. A rudimentary form of gaze following can be found at birth (Farroni, Massaccesi, Pividori & Johnson, 2002). Later in development, infants' shifts of attention can be also triggered by social cues like hand grasps and points (i.e. 4-months-olds). Further, recent evidence has shown that action priming emerges at the age at which infants themselves start performing the gesture. Thus, gestures comprehension is deeply shaped by sensorimotor experience (Lagdon & Smith, 2005; Daum, & Gredeback, 2011; Gredeback, Melinder, & Daum, 2010; Bakker, Daum, Handl, & Gredeback, 2014). Covert shifts of attention primed by social and non-social cues have been previously investigated in adults using the Posner paradigm. Studies using electroencephalography (EEG), have found that priming effect in adults occurs very rapidly, within 300 ms from cue onset over posterior sites (Langdon & Smith, 2010). To date, the neural underpinnings of covert shifts of attention triggered by hand gestures in infancy is still unknown. Infants brain processing have been predominantly explored with ERPs techniques. An advantage of using electroencephalography (EEG) is that it provides data with excellent temporal resolution and, thus, is a fruitful approach to study very fast neural signals, like the ones elicited by covert attention. In Chapter 1 of Section 1, EEG technique will be used to deepen our comprehension of action priming effects in the infant brain and the role of sensorimotor experience in hand gestures processing.

1.1.3 Action prediction: links between action and observation.

Both detection of relevant social information and action priming, are processes essential to predict others' actions (Gredeback & Daum, 2015). Infants start to predict simple and familiar actions performed by others by 6 months of age (Woodward, 1998; Hunnius & Bakkering, 2010, Kanagogi & Itakura, 2011; Kochukhova & Gredeback, 2010) and, only at 12 months, but not at 6 months, they start predicting more complex actions like moving an object in a container (Cannon, Woodward, Gredeback, von Hosten, & Turek, 2012; Falck-Ytter, Gredeback, & von Hosten, 2006). These findings showed that infants can't predict an action that they can't perform themselves and has been taken as evidence of the mirror neuron view of action understanding.

Since their discovery in the early 1990s, mirror neurons have been considered as a mechanism by which humans can understand others' actions (Rizzolatti et al., 2001). These neurons were first identified in the premotor cortex F5 of the Macaque monkeys and were demonstrated to be functional to action understanding because they were active during both execution and observation of goal directed actions (e.g. Rizzolatti et al., 2001). In the following years, evidence using neuroimaging technique has shown that neurons with similar mirror properties were also contained in several areas of the human brain (Buccino et al., 2001; Keysers & Gazzola, 2010), including the inferior frontal gyrus, the ventral premotor cortex and IPL. Mirror neurons involvement in action understanding has also been observed via EEG, indexed by the desynchronization of the mu rhythm in both adults and infants (Hari, 2006; Nyström, 2008). After the discovery of mirror neurons, it has been suggested that brain activation elicited by action observation might serve to facilitate action understanding (Rizzolatti et al., 2001). Different theories were then built to explain the role of mirror neurons in action understanding. The *direct-matching hypothesis* posits that when we observe somebody else performing an action, we automatically and directly map the observed action in our own motor system, without the need of any cognitive or inferential processes (Rizzolatti et al., 2001). This means that the motor representation activated during the observation of an action matches the motor representation that would be active if the observer performed the same action. However, the direct matching hypothesis was called into question for several reasons. First of all, the direct-matching mechanism cannot explain why mirror neurons are active also during observation of actions performed by other species (Buccino et al., 2004) or by a mechanical claw (Gazzola et al.,2007). Secondly, evidence has found dissociations between the ability to produce and to recognize actions in people with apraxia (Negri et al., 2007). Recent theories suggested that there is a second mechanism called 'mentalizing network' that represent the underlying intention of an action (Van Overwalle et al., 2009) while others have suggested that within the mirror system there are neurons that respond specifically to the goal of an action (Gazzola et al., 2007). According to all these theories, the action-perception link is an innate evolutionary endowment. An alternative account suggests that action-perception couplings are acquired through *associative learning* thanks to the accumulation of sensorimotor experience (Heyes, 2010). Differently, according to the *Teleological stance* viewpoint (Csibra et al., 2007) the goal of an action is inferred outside the motor system and only subsequently duplicated in the mirror system through emulative processes (Csibra et al., 2007). Recently, a neuroconstructivist framework (Quadrelli & Turati, 2016) suggests that the direct-matching and the teleological views might not be considered contrasting theories, but, instead, as complementary mechanisms both involved in action understanding. For example, when the action is familiar direct mirroring might be a more efficient way to understand the goal of the action. In the case of unfamiliar actions, the goal of the action could be reconstructed by higher level inferential paths.

1.2 Processing of emotional facial expressions.

A cornerstone of social understanding is emotion processing. Express and detect emotions in others is crucial in social interactions because it allows us to communicate our and understand others' emotional states. Detect, discriminate and categorize facial expressions of emotion are crucial processes during infancy that can provide infants with essential information to understand, later in development, how to act in the external world and predict what others are going to do.

1.2.1 Early biases towards emotions.

Infants' attention seems to be biased towards happy faces from the very first stages of postnatal life (Farroni, Menon, Rigato, & Johnson, 2007; Rigato, Menon, Johnson, Faraguna, & Farroni, 2011). These findings show that newborns, despite the relatively little visual experience possessed at birth and their immature visual system, seem to be already well attuned to relevant emotional information present in their surroundings. This early attunement might provide a fertile ground for the development of later and more complex abilities acquired later on through experience. For example, significant postnatal experiences are required to enrich infants' understanding of emotions. In fact, only at 7 months of age infants start to categorize facial expressions (de Haan, & Nelson, 1998; Leppänen, Richmond, Vogel-Farley, Moulson, & Nelson, 2009) and start manifesting an adult-like negativity-bias, showing to be more attracted by negative compared to positive emotions (Hoel et al.,2008; Leppanen et al., 2007; Peltola et al., 2009). These perceptual and attentional biases are thought to rely on a distributed neural network involving areas that overlap with the social brain (Adolphs, 2003). However, most of the current knowledge about infants' processing of emotions comes from studies that used static stimuli. To better understand how emotional stimuli are processed early in infancy, it is crucial to use stimuli that better approximate real life. This could be done, for instance, by using dynamic facial expressions of emotions as stimuli. This would allow researchers to shed light on how infants process the world as they actually see it. In Chapter 4 of section 2, a study that explores visual discriminative abilities of dynamic facial expressions at birth will be presented.

Available evidence suggests that there are some aspects of emotion recognition that emerge very early in development, like early attentive bias towards emotional stimuli, and other aspects, like emotion regulation, that needs more time to become fully mature. Which is the mechanism at the basis of such developmental pattern?

1.2.2 Mechanisms of emotion recognition.

Different mechanisms can be considered at the basis of emotion recognition. The first mechanism considers recognition as part of perception: we can discriminate, categorize and identify emotions relying solely on the perceptual/structural properties of the emotional stimulus (see Adolphs, 2003). This mechanism might be functional from very early in life, representing a prerequisite for the ability to recognize emotions that emerge later in development. A second possible mechanism is recognition of emotions via simulation: observing a facial expressions might activate sensorimotor representations involved in the execution of that expression, which, in turn, might facilitate emotion recognition (Adolphs, 2006). This proposal is supported by recent findings that have shown activation of a shared neural network during observation, imitation and execution of facial expressions (Carr et al., 2002). Recently, evidence of the mirror system as a potential mechanism of emotion recognition has been found also in 30-month-olds infants, by showing desynchronization of mu rhythm during observation of facial expressions (Rayson, Bonaiuto, Ferrari & Murray, 2016). Further, at 5 months of age infants show rapid facial mimicry responses to dynamic multimodal emotional stimuli (Isomura & Nakano, 2016).

How does the ability to recognize emotions come into place? The evidence of the presence already at birth of anatomical connections between emotion- and attention-related areas (Macado & Bachevalier, 2003) and of early attentive biases towards salient emotional expressions has been taken as evidence of experience-expectant processes in emotion recognition. Then, through exposure to species-typical emotional signals the immature neural circuitries become more specialized and refined (Leppanen & Nelson, 2009). Infants' face-to-face interactions with the caregiver from 9 months onwards might be crucial for this refinement process leading to a strengthening of neural connectivity between areas involved in different aspects of emotional processing: experience (amigdala, anterior insula), production (premotor cortex), perception (temporal cortex) and regulation (prefrontal cortex) of emotion (Leppanen, 2011).

To sum up, sensitivities to others' gestures and emotions are both key processes in social understanding and are supported by brain regions which were identified as being part of the 'social brain' (Adolphs, 2003). Much of the current knowledge has focused on adults processing of the social world while little is known about the development of these abilities. However, available evidence has shown that precursors of the social brain network can be found already in infancy, demonstrating early adaptation of the human brain to the social context. Interestingly much of the current knowledge has explored each of the several domains that are part of the social brain separately. Advances in the studies of social cognition might be done by exploring how these domains interact and affect each other. This issue will be asses in Chapter 5 of Section 2, by exploring potential modulations in infancy of the emotional context on action processing trough EEG measurements.

Aims of the study.

The first year represents an exceptional time of growth during which important social skills emerge, like gestures and emotion understanding. To date, some aspects of infants' and newborns' early sensitivity to human gestures and emotions need to be further explored. The current investigation includes two main sections in which we will shed more light on newborns' and infants' gestures and emotion understanding.

Section 1 focuses on newborns' and infants' understanding of human gestures coming from the face, the hand, and from touch. This section includes 3 Chapters, each one tackling an intriguing issue regarding gestures processing in infancy with special attention on the role of sensorimotor experiences in the development of this ability. In particular, in *Chapter 1* will be described a study that employs Electroencephalography (EEG) with 6 month-old infants. The main aim of this study is to identify the neural underpinnings of covert shifts of attention triggered by possible and impossible grasping gestures. The manipulation of the biomechanical properties of the human gesture will allow us to deepen our comprehension of the role of early sensorimotor experience in action understanding. Chapter 2 and 3 describe two studies that sought to investigate whether pre- and postnatal sensorimotor experience with touching gestures and facial gestures can boost early perceptual visual abilities at birth. In particular, in *Chapter 2,* 2-day-old newborns and 3-month-old infants will be tested using a visual preference procedure to examine their ability to discriminate between touching and no-touching gestures. Lastly, the study presented in *Chapter 3* aims to explore whether newborns could benefit by the presence of audio-visual information that is part of their sensorimotor experience to match a facial gesture to the corresponding sound produced by it even when synchrony cues are not available.

Section 2 focuses on emotion processing in infancy. In an effort to conduct experimentally rigorous and well-controlled experiments, many of the studies that have explored emotion processing in infancy have used as stimuli static and disembodied faces. Thus, the studies presented in the two chapters of this section will attempt to improve the ecologic validity of the stimuli used to better understand how infants process the world as they actually see it in everyday life. At birth, newborns are mostly exposed to dynamic facial expressions and not to static photographs. *Chapter 4* describes a study that aims at exploring whether 2-day-old newborns can discriminate between dynamic facial expressions of happiness and disgust using both visual preference and visual habituation procedures. This study will shed light on the early ability to detect morphological changes that take place during the unfolding of emotional expressions. To date, most of the current studies have addressed facial expressions in isolation. However, actions are most of the time embedded in an emotional context and it is fundamental in life to be able to link effective- to action-related information. To understand whether infants can bind affective-related to action-related information in Chapter 5 will be present an EEG study with 6-month-olds infants that aims at exploring whether the emotional context (happy, negative and neutral facial expression) can modulate the processing of an action.

Section 1. Understanding human gestures in the first year of life.

Introduction

Human infants possess some knowledge of their social world and, from very early in life, they start developing expectations and understandings about the meaning of others' actions and gestures. In particular, human gestures are pervasive behaviors which are fundamental for human communication. Humans are exposed to gestures since they are born and they start to communicate through facial and body gestures before they learn to express their intentions through speech.

Some human gestures are fundamental to understand where an interactor is orienting his attention in the environment. To understand where others are directing their attention we rely on different clues like eye gaze, head turning, body orientation and referential manual gestures as well (i.e., pointing and grasping) (Lagdon & Smith, 2005; Daum, & Gredeback, 2011). For example, when we observe somebody who looks on the right, we immediately shift our attention in the same direction even if we are not overtly looking there. To capture these covert shifts of attention Posner (1980) introduced a priming paradigm in which a centrally displayed cue was followed by a target that could appear in the same direction of the cue (Congruent condition) or in the opposite direction (Incongruent condition). Typically we are faster in orienting our attention towards the target in the congruent condition compared to the incongruent one. This priming effect is the result of a covert shift of attention in the cued location and it has been observed both with inanimate cues (arrows) or human gestures like, for example, eye gaze (Lagdon & Smith, 2005) and grasping gestures (Daum, & Gredeback, 2011). Interestingly, converging behavioral evidence have shown that priming effects can be found very early in life. For example, by 6-months of age infants are faster to make saccades towards a peripheral target cued by the direction of eye gaze (Farroni, Johnson, Brockbank & Simion, 2000), hand grasp (Daum, & Gredeback, 2011; Wronsky & Daum, 2014), human walking point-light-displays (PLDs) (Bardi, Di Giorgio, Lunghi, Troje, & Simion, 2015) as well as pointing gestures (Rohlfing, Longo, & Bertental, 2012). These findings show that since the first year of life infants are sensitive to referential human gestures like pointing or grasping. Yet, the neurophysiological underpinnings of action priming and the properties of gestures that might be crucial for it remain unknown. We will address these issues in Chapter 1, where will be described a study in which we recorded electroencephalographic activity (EEG) from 6-month-old infants.

Touching gestures also play a crucial role in our life. Besides being extremely important as a means to explore and learn about our surrounding environment, the sense of touch is also fundamental in our social interactions. When we observe a touching gesture, like a caress or a slap, we attribute to it different meanings. Thus, touching gestures represent an important non-verbal channel through which we can communicate our intention and affect to others. Recent evidence in human adults indicates that touch is crucial not only for own corporeal sensations, but also for understanding others tactile sensations likely through mirroring mechanisms (i.e. Keysers, Gazzola, Anton, Fogassi, Gallese, 2004; Blakemore, Bristow, Bird, Frith, Ward, 2005; Ebisch, Perrucci, Ferretti, Del Gratta, Romani & Gallese, 2008). Humans start to use the sense of touch as a means to explore the surrounding word already in the confines of the womb. During the gestational period fetuses perform different hand movements directed towards parts of their own body or features in the environment, such as the umbilical cord and the uterine wall (Sparling, 1999, Sparling e Wilhelm 1993). Further, between the 21st and the 25th week of gestational age fetuses respond to external tactile stimulations (Marx & Nagy, 2015) and at birth the cortex is developed enough to produce somatosensory responses even if this cortical processing is immature (Pihko, Nevalainen, Stephen, Okada & Lauronen, 2009). Moreover, Filippetti and colleagues (2013) provided the first evidence of newborn detection of visual-tactile synchrony of a stimulation on their own body and an observed tactile stimulation. However, how touching gestures are visually processed in infancy is a topic that remains still unexplored. A handful of studies have highlighted the role of pre- and postnatal sensory-motor experiences in the development of some visual perceptual abilities at birth (Craighero, Leo, Umiltà, & Simion, 2010; Longhi et al., 2015; Guellaï, Streri & Yeung, 2014). Therefore, it is reasonable to assume that the extensive somatosensory-motor experience accumulated during fetal and postnatal life through the sense of touch could provide a base at birth for the development of the ability to visually recognize touching gestures. This issue will be tackled in Chapter 2, in which will be presented a couple of studies to assess 2-day-olds and 3-month-olds ability to visually process gestures involving a tactile component.

Lastly, among all the visual stimuli that surround us, faces are one of the most salient. Since birth, newborns are particularly attuned to human faces: a few hours after birth their attention is preferentially attracted by faces compared to other visual stimuli (Fantz, 1961; Johnson & Morton, 1991; Macchi Cassia, Turati, & Simion, 2004) and, further, they prefer the mother's face compared to a stranger's face (Bushneil, Sai & Mullin, 1989; Field, Cohen, Garcia, & Greenberg, 1984; Pascalis et al., 1995). Facial gestures are also extremely attractive: from the first time a newborn is held, he/she sees faces that are talking, expressing emotions or performing non-emotional gestures like yawning or hiccupping. Typically, these facial gestures are accompanied by sounds and, thus, newborns' first experiences with faces are multimodal in nature and involve both vision and audition. Evidence of an early sensitivity for facial gestures comes from studies on newborn imitation (Meltzoff & Moore, 1983). Within the first 71 hours after birth newborns seem to be able to imitate adults' facial movements like mouth opening and tongue protrusion (Meltzoff & Moore, 1983). This has been interpreted as indicating that shortly after birth newborns are able to identify a facial gesture and then to repeat it by matching their own facial movements to the ones of the live model. So far, neonatal imitation isn't a phenomenon unanimously accepted given that further attempts to replicate this seminal finding yielded mixed results (see for an overview Anisfeld, 1991; 1996; Ray & Heyes, 2012). However, despite the intense debate on newborns' imitative abilities, a recent finding have shown that newborns imitated significantly more quickly adults' facial gestures when they are presented along with the sound that they produce (Coulon, Hemimou, & Streri, 2013). Further evidence that supports newborns' early attuning to facial gestures is provided also by studies that explored whether they are able to match a facial gesture with the corresponding sound. Past and more recent findings demonstrate that newborns seem to be sensitive to some audio-visual correspondences coming from face and voice already in the first days after birth (Aldridge, Braga, Walton, & Bower, 1999; Lewkowicz, Leo, & Simion, 2010; Guellaï et al., 2016). Then, through infancy, the ability to integrate facial gestures to a voice becomes more sophisticated thanks to significant experience acquired postnatally (Lewkowicz, 2000). However, all the studies that have explored audio-visual matching abilities at birth have used stimuli that weren't part of newborns' own sensorimotor repertoire (Slater, 1999; Aldridge, Braga, Walton, & Bower, 1999; Lewkowicz, Leo, & Simion, 2010; Guellaï et al., 2016). In Chapter 3 will be presented a study in which it will explore whether the presence of audio-visual information that newborns have experienced can affect their early multisensory abilities.

To sum up, a great amount studies show that infants are particularly attuned to their social world. By 6 months of age, they are able to follow the direction of referential manual gestures (i.e. Daum, & Gredeback, 2011). Since the very first days of life, they can detect temporal synchrony between an observed touching gesture and a felt body-related touch (Filippetti et al. 2013), and also audiovisual correspondences from facial gestures and sounds (Guellaï et al. 2016). However, some aspects of infants' and newborns' early sensitivity to human gestures need further exploration and the following three studies of this first section will provide new interesting insights about this issue.

Chapter 1. Spatial orienting following possible and impossible grasping gestures: an EEG study with 6-months-old infants.[1]

1.1 Introduction

Following the direction of others' manual gestures is a fundamental skill that allows us to orient our attention in relevant locations of the surrounding environment. In the last decade a growing amount of studies started to explore how this important skill develops during infancy and have found that human manual gestures can trigger shifts in attention very early in life. Here we will add an important contribution to the current body of knowledge by investigating the neural underpinnings of this ability through EEG measurements and, further, we will provide further evidence on how sensorimotor experience affects action priming. Several studies on action processing in infancy have emphasized the role of action experience in the ability to predict the goal of an observed action (Stapel, Hunnius, Meyer, & Bekkering, 2016; Sommerville, Woodward, & Needham, 2005; Falck-Ytter, Gredebäck, & von Hofsten, 2006; Ambrosini, Reddy, de Looper, Costantini, Lopez, & Sinigaglia, 2013). For instance, infants' goal inferences improve when the action that they observe is a movement that they can perform (Cannon & Woodward, 2012; Falck-Ytter et al., 2006; Kanagogi & Itakura, 2011; Stapel et al., 2016). Also action priming is modulated by infants' own experience with the gesture. Daum and colleagues (2011) have

---

[1] Results presented in this chapter have been previously published in "Natale, E., Addabbo, M., Marchis, I. C., Bolognini, N., Macchi Cassia, V., & Turati, C. (2016). Action priming with biomechanically possible and impossible grasps: ERP evidence from 6-month-old infants. Social neuroscience" doi:10.1080/17470919.2016.1197853."

shown that a grasping hand triggers shifts in attention of 5-month-old infants, that is at the age when they start to learn how to perform grasping actions. Modulating effects of motor skills on priming effect were also found in studies that have investigated the neural underpinnings of action priming (Senju, Johnson, & Csibra, 2006; Gredeback, Melinder, & Daum, 2010; Bakker, Daum, Handl, & Gredeback, 2014). These studies have used a procedure in which, unlike the Posner paradigm, the peripheral target preceded the gesture and ERP responses were then measured at the onset of the referential action/eye gaze. Two key event-related EEG components were modulated by the congruency of the action with the object: the posterior 290 for eye gaze (Senju et al., 2006) and the P400 for referential actions (Gredeback et al., 2010; Bakker, Daum, Handl, & Gredebäck, 2014). In particular, enhanced posterior N290 responses were found in response to object-incongruent gaze shifts (Senju et al., 2006). The posterior P400 ERP component differentiated congruent from incongruent condition for power grasp only at 6 months, age at which infants were proficient at grasping. Conversely, no congruency effect was found for precision grasp, gestures that infants weren’t able to perform. With respect to pointing, the congruency effect was found on the P400 at 13 months, but not at 6 months (Melinder, Konijnenberg, Hermansen, Daum, & Gredebäck, 2015).

Action priming effects in infancy seem to be also influenced by two other main factors: the motion of the cue and the animacy of the cue. Accordingly, priming effect is enhanced when infants observe dynamic gestures compared to static ones (Farroni, Mansfield, Lai, & Johnson, 2003; Rohlfing, Longo, & Bertenthal, 2012; Wronsky & Daum, 2014). For example, only dynamic pointing gestures, but not static ones, could trigger 4.5 and 6.5-month-olds infants’ shifts of attention towards the target (Rohlfing et al., 2012).

With respect to the animacy of the cue, action priming is less likely to occur for pseudo-social cues (i.e., foils; Bertenthal, Boyer, & Harding, 2014), or non-human cues (i.e., mechanical claws; Daum, & Gredeback, 2011).

So far we don't know whether and to what extent infants' ERP responses may reflect facilitation in orienting the attention towards the cued target, as observed in eye-tracker studies (Berthental et al., 2014; Daum & Gredeback, 2011; Daum et al., 2013; Rohlfing et al., 2012 Wronsky & Daum, 2014). Given that the paradigm to study action priming adopted in previous EEG studies (Gredeback et al., 2010; Bakker et al., 2014) is remarkably different from the one used in behavioral studies, the link between priming effect and P400 appears merely speculative. More direct effects of covert orienting of attention towards a location cued by an arrow on early components of the ERP have been found in several studies that have used the spatial cueing paradigm in adults (Eimer, 1996, 1998; Hillyard, Luck, & Mangun, 1994; Mangun & Hillyard, 1991). Congruency effect occurred on the occipital P1 and N1 which amplitudes were larger on congruent trials than on incongruent ones. Neural basis of priming effect was investigated for the first time also in infancy by Richards (2000) who measured ERP responses to targets that appear in attended and unattended locations. As in adults, he identified two early components, namely the P1 and N1 over contralateral occipital sites. Congruency effect was found only on the P1 component for 4.5-month-olds and 6-month-olds, but not for 3-month-olds. Thus, in both adults and infants priming effect arises from a sensory facilitation in processing visual information at a cued location, indexed by an increase of efficiency at early stages of visual processing.

Further, it is still unclear from literature which properties of human gestures are critical in triggering action priming effects. In this regard, available evidence in very young infants has shown that the motion of the cue seems to be necessary for action priming (Rohlfing et al., 2012; Wronsky & Daum, 2014). With respect to static presentations, dynamic gestures might provide additional cues on which infants can rely to detect the directionality of the gesture and then rapidly shift their attention accordingly. Static images of a gesture contain information about the grasping direction while dynamic stimuli provide a second additional cue, that is the moving direction. Young infants, unlike older ones, might not be able to use each of the two cues in isolation and both of them might be needed to prime efficiently their attention (Wronsky & Daum, 2014). Nevertheless, besides grasp and movement direction, infants could rely on information coming from the biomechanical properties of the gesture to process action direction. Available evidence have shown that by 6-months of age infants can detect violation of human grasp biomechanics (Geangu, Senna, Croci, & Turati, 2015) and, further, they can attribute goals to biomechanically impossible actions (Southgate, Johnson, & Csibra, 2008). Still, it is not known whether biomechanically properties may act as a cue priming orienting responses towards targets congruent with the direction of the action.

In the current study we aim at exploring whether ERP responses may reflect facilitation in orienting attention towards the cued target, as observed in eye tracking studies (Berthental et al., 2014; Daum & Gredebäck, 2011; Daum, Ulber, & Gredebäck, 2013; Rohlfing et al., 2012; Wronsky & Daum, 2014). As in previous behavior studies on action priming, we will use a modified version of the Posner Paradigm (Posner, 1980) in which infants will be presented with a centrally displayed grasping that is followed by a target that could appear in the same direction indicated by the hand gesture (Congruent condition) or in the opposite direction (Incongruent condition). Sensory facilitation at the cued location should be indexed by the modulation of early posterior ERP components contralateral to target presentation (i.e., increased P1 and/or N290 amplitude on congruent trials) (Richards, 2000). Importantly, here we manipulated the biomechanical properties of the grasping, thus, in a condition the gesture was executable by a human hand while in the second condition the gesture was impossible to be executed by a human hand. It is worth noting that biomechanically possible and impossible grasps are both characterized by a movement direction, although impossible grasps violate anatomical constraints. If the information about movement direction is sufficient to trigger a shift of attention, we would expect to find similar ERP modulation as a function of the congruency of the trial, regardless of the biomechanical plausibility of the hand movement. On the contrary, if infants heavily rely on information about the biomechanical properties of human motion, we would expect to observe ERP modulations as a function of the congruency only for biomechanically possible grasps, but not for impossible grasps.

1.2 Experiment

1.2.1 Methods

Participants

Twenty-four healthy, full-term 6-month-olds (11 females, mean age = 6 months 4 days, range = 150–198 days) were randomly assigned to the two experimental conditions: 12 (6 females, mean age = 6 months and 5 days, range = 150–198 days) to the biomechanically possible hand action condition, and 12 (5 females, mean age = 6 months and 2 days, range =

167–196 days) to the biomechanically impossible hand action condition. An additional 28 infants were also tested, but were not included in the final sample due to fussiness and no completion of an adequate number of trials to be considered for data analysis (i.e., 10 trials per condition; N = 19 infants), or eye and body movements that resulted in excessive recording artifacts (N = 9 infants). The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethical Committee of the University of Milano-Bicocca. Parents gave their written informed consent.

Stimuli, apparatus, and procedure

Testing took place in an electrically shielded and dimly illuminated room. Infants were seated on a parent's lap approximately 60 cm from a 24-inch screen used for stimulus presentation. A video camera installed above the screen recorded a video of the infant, which was synchronized with stimulus presentation for off-line coding of eye and body movements occurring during each trial. As shown in Figure 1, the latter began with an animated fixation point, randomly selected among sixteen different animations, which was displayed at the center of the screen. As soon as the infant looked at it, this was replaced by the visual-spatial cue, namely a dynamic gesture of a human hand randomly grasping either toward the left or the right side (subtending maximum 5 degrees of visual angle along the horizontal axis and 15 degrees along the vertical axis). The grasping action lasted for 1200 ms and, after a variable delay (range: 300-500 ms), a target stimulus was displayed for 200 ms. The target consisted of the static image of a colorful ball, which was randomly selected among sixteen images of a colored ball (all 1.75 cm in radius), in order to have a different target for each trial within the trial block. The target could be randomly presented at a peripheral location (~10 degrees of visual angle from the center of the screen) either congruent (i.e., valid) or incongruent (i.e. invalid) with the grasping direction. Two types of grasping gesture were presented: a biomechanically possible hand action and a biomechanically impossible action. The biomechanically impossible action was obtained by making the virtual hand's fingers bending in the opposite direction with respect to that in the possible grasp, that is toward the back of the hand. In the impossible grasping, all angular displacements were clearly beyond the natural limits of the metacarpophalangeal joints, thus clearly violating the biomechanical constraints of the phalanges. The two types of gestures were otherwise matched for low-level visual features as well as cinematic aspects, i.e. velocity and duration of the hand movement. In particular, the angle of fingers' and phalangeal joints' displacements was matched frame by frame between the two types of grasping, see Figure 1. For each type of grasping, stimuli were presented in blocks of 16 trials, 8 valid (4 with left- and 4 with right-sided targets) and 8 invalid (4 with left- and 4 with right-sided targets). There was no restriction in number of blocks or trials shown, i.e., they could be played indefinitely. The experimental session terminated when infants looked away from the screen during five consecutive trials. On average, 48 trials (range = 33-72) were presented for the biomechanically possible grasping and 44 trials (range = 34-57) for the biomechanically impossible grasping. The total number of trials was not significantly different for the two types of action ($p > .43$). Also, there was no between- or within-group difference in the number of valid and invalid trials that were presented (possible action: 25

valid and 23 invalid trials; impossible action: 22 valid 12 and 22 invalid trials; all ps > .11). The sequence and timing of the stimuli were controlled by Eprime 2.0 software.

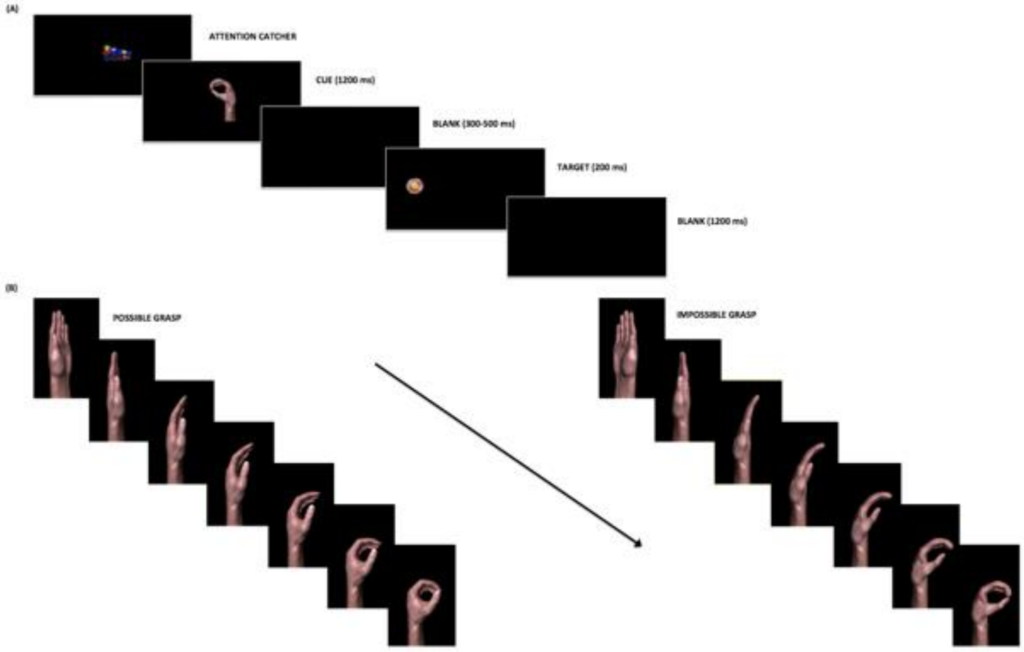


Figure 1. A) A schematic representation of the sequence of events presented on each trial. This included a dynamic grasp displayed at the center of the monitor, which cued a peripheral target in either a valid way, as shown in example, or an invalid way.

ERP recording and analysis

Continuous scalp EEG was recorded from a 128-channel HydroCel Geodesic Sensor Net (Electrical Geodesic, Eugene, OR) that was connected to a NetAmps 300 amplifier (Electrical Geodesic, Eugene, OR) and referenced on-line to a single vertex electrode (Cz). Channel impedance was kept at or below 100 KΩ and signals were sampled at 500 Hz. EEG data were pre-processed off-line using NetStation 4.5 (Electrical Geodesic, Eugene, OR). The EEG signal was segmented to 1200 ms, post-stimulus onset, with a baseline period beginning 100 ms prior to target onset. Data segments were filtered using a 0.3–30 Hz band-pass filter and baseline corrected using mean voltage during the 100 ms pre-stimulus period. Automated artifact detection was applied to the segmented data to detect individual sensors that showed > 200 uV voltage changes within the segment period. The entire trial was excluded if more than 18 sensors (15%) overall had been rejected. Data were then inspected manually to mark as bad segments containing drift and eye blinks. We also manually marked as bad segments belonging to trials in which the infant did not look at the cue or did not keep central fixation at least until the target off-set, as assessed by off-line coding of the infant’s video. Bad segments identified by either procedure, i.e. automatic and manual, were 36% of the total number of presented trials for the biomechanically possible grasp and 34% for the biomechanically impossible grasp, and they were excluded from further analysis. Of the remaining trials, individual channels containing artifacts were replaced using spherical spline interpolation. For each participant, average waveforms were generated within each experimental condition (valid and invalid separately for left- and right-sided targets) only if at least 10 artifact-free trials were overall available per condition collapsed across the target-side. Averaged data were then re-referenced to the average reference. Overall, for the biomechanically possible grasp, each subject contributed with a mean number of 15 trials (range 13-24) to the valid condition average and a mean number of 15 trials (range 10-25) to the invalid condition average ($p > .5$), whereas for the biomechanically impossible grasp, each subject contributed with a mean number of 15 trials (range 12-20) to the valid condition average and a mean number of 13 trials (range 10-21) to the invalid condition average ($p = .01$). The mean number of trials per condition did not differ between the two groups (all ps > .45). Inspection of the grand-averaged waveforms revealed that the two ERP components of interest, i.e. the P100 and N290, were elicited at the target onset over occipito-temporal scalp sites contralateral to the hemifield of stimulus presentation, but not over homologous ipsilateral sites (Figure 2). Thereby, ERP responses to the target were analyzed from posterior scalp sites of the contralateral hemisphere only. Based on visual inspection of both the grand averaged and individual waveforms, twelve electrodes were identified for the ERP analysis: electrodes 71-70(O1)-66-65-64-58(T5) over the left hemisphere to analyze ERP in response to valid and invalid right-sided targets, and electrodes 76-83(O2)-84-90-95-96(T6) over the right hemisphere to analyze ERP in response to valid and invalid left-sided targets (Figure 2). For each participant, signal from homologous electrodes of the two hemispheres (71/76, 70(O1)/83(O2), 66/84, 65/90, 64/95, 58(T5)/96(T6)) was averaged to analyse ERP responses to valid and invalid targets across the whole visual field. Based on grand-averaged data and individual data, peak amplitude and latency of P100 and N290 were extracted within a time window of 100 to 230 ms and of 200 to 330 ms, respectively. For each ERP component, a repeated-measures Analysis of Variance (ANOVA) was separately carried out on latency and/or amplitude (as obtained by collapsing together averaged amplitude and latency from each pair of homologous electrodes) with target validity (valid vs invalid) as within-subjects factor and grasp type (biomechanically possible and impossible) as between-subjects factor.

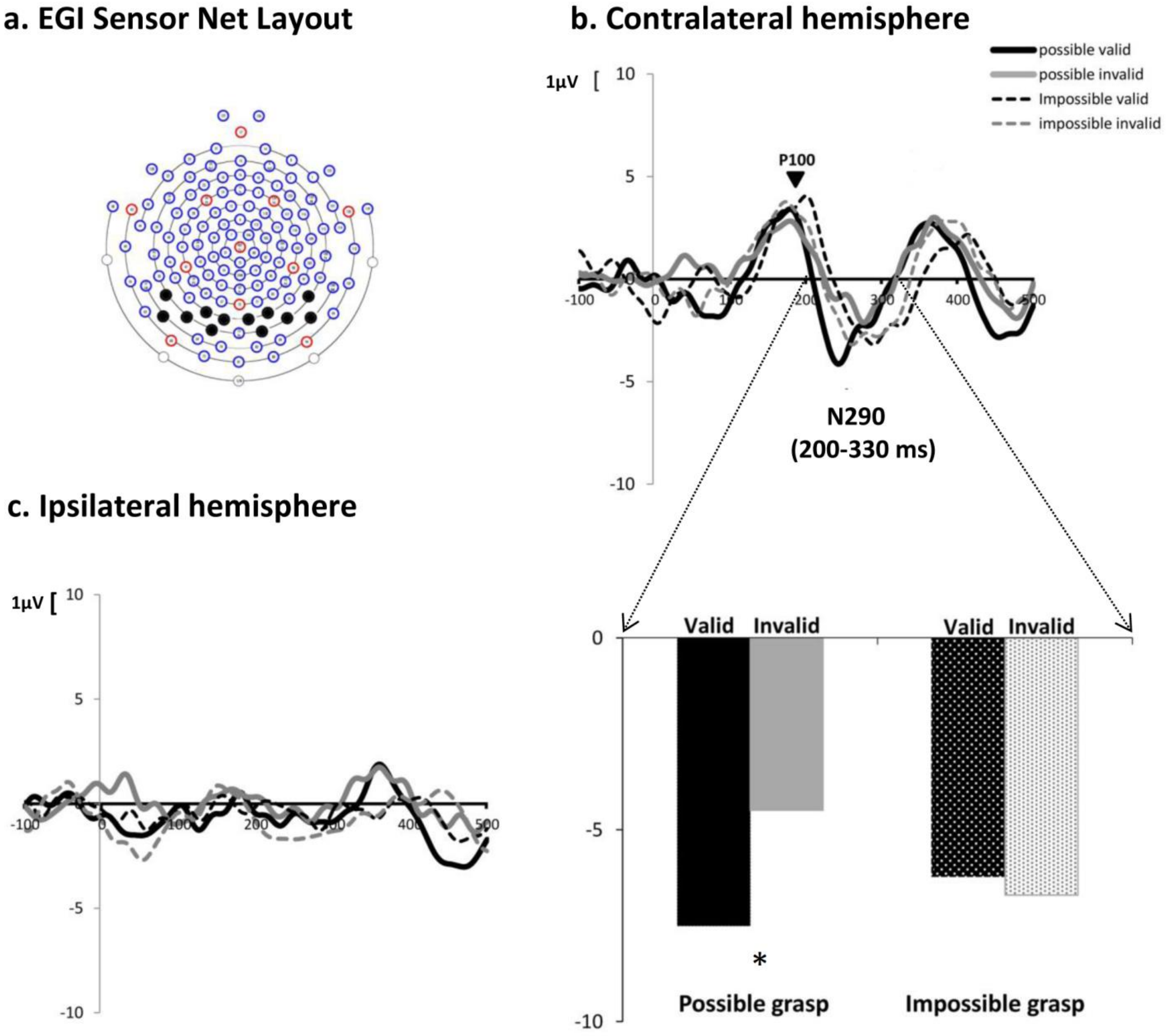


Figure 2. A) ERPs waveforms elicited by the target cued by possible grasps (solid lines) and impossible grasps (dashed lines) on valid trials (black lines) and invalid trials (grey lines) in the hemisphere contralateral (left panel) and ipsilateral (right panel) to the hemifield of target presentation. ERPs were averaged across the hemifield of target presentation and the electrodes selected for the analysis. The number of trials going into the grand averages were 183 for the possible valid, 175 for the possible invalid, 182 for the impossible valid, and 161 for the impossible invalid condition. B) Mean amplitude of the N290 component in response to valid and invalid targets cued by possible and impossible grasps. Error Bars = SEM; *= p < .01. C) The recording layout of the EGI sensor net. The two electrodes on the outside of the eyes (125, 128) and the two below the eyes (126, 127) were not used and are marked in grey in the layout. The posterior electrodes, marked in black, were selected for the ERP analysis.

1.2.2 Results

P100 (100-230 ms)

The ANOVAs did not reveal any significant main effect of grasp type and target validity as well as any no significant grasp type by target validity interaction for both the peak amplitude and latency of P100 (all ps > .13).

N290 (200-330 ms)

As for the peak amplitude of N290, the ANOVA revealed no main effect of grasp type (p > .65), a marginally significant main effect of target validity, F(1,22)=4.3, p =.051, η2=.16, but a significant Grasp Type by Target Validity interaction, F(1,22)=9.1, p < .01, η2=.29. Bonferroni post-hoc test indicated that the amplitude of the N290 was larger in response to valid (-7.5 µV) than invalid (-4.5 µV) targets cued by a biomechanically possible grasp, p < .01, whereas there was no difference between the amplitude of the N290 in response to valid (-6.2 µV) and invalid (-6.7 µV) targets cued by a biomechanically impossible grasp, p = .9, see Figure 2. Bonferroni post-hoc test also indicated that the amplitude of the N290 in response to valid and invalid targets did not differ between infants cued with biomechanically possible and impossible grasps, ps>.8. As for the peak latency of N290, no main effects or interactions attained statistical significance (all ps > .09).

1.3 Discussion

In the present study, we have observed two ERP components in response to the target onset over posterior scalp sites contralateral to the hemifield of stimulus presentation: the P100, peaking at about 165 ms and the N290, peaking at about 270 ms. Although this latter component likely corresponds to the N1 reported by Richards (2000), we preferred to label it N290 in accord with the infant ERP literature (e.g. Nelson, Thomas, & de Haan, 2006). The ERP congruency effect modulated the amplitude of N290 only when infants were presented with the possible gesture, but not with the impossible gesture. Specifically, for biomechanical possible grasps the amplitude of N290 was larger in response to targets at the cued than at the uncued location. This is in line with previous ERP studies employing a spatial cuing paradigm with exogenous, non-social cues (Hillyard & Anllo-Vento, 1998; Richards, 2000). Therefore, modulations of the ERP component N290 can be considered as the result of the shift of visual-spatial attention triggered by the valid cue, which enhanced sensorial analysis of stimuli at the attended location.

Here we have shown for the first time that, early in development, grasping gesture can affect neural processing of stimuli that appear in a cued location in the space. This hypothesis is supported by ERP evidence indicating enhanced processing of objects previously displayed at the location looked by an adult face (Reid et al., 2004) as well as by behavioral evidence showing that human actions trigger faster orientations to targets displayed at the congruent spatial position (Bardi et al., 2015; Bertenthal et al., 2014; Daum & Gredebäck, 2011; Daum et al., 2013; Farroni et al., 2000; 2003; Hood et al., 1998; Rohlfing et al., 2012; Wronsky & Daum, 2014). Unlike previous studies on infants' covert spatial attention here no ERP validity effect was found at the level of the early P100 component (Richards, 2000). This could be due to several methodological differences between the two studies: first of all, the nature of the cue (i.e., social in the present study, non-social in the previous study). Indeed, the N290 component is specifically sensitive to social information in infancy (Csibra, Kushnerenko, & Grossman, 2008; de Haan, Johnson, & Halit, 2003; Luyster, Powell, Tager-Flusberg, Nelson, 2014; Nelson et al., 2006). Nonetheless, it is important to consider that in developmental studies the N290 has been identified as a component with topographical and functional analogies to the adult face-sensitive N170 (Csibra et al., 2008; de Haan et al., 2003; Luyster et al., 2014; Nelson et al., 2006). Second, in previous studies that have used static actions (Bakker et al., 2015; Gredebäck et al., 2010; Melinder et al., 2015) validity effects have been found on a later component, namely the P400. This could be explained by the dynamic nature of the stimuli used in the present study which may trigger earlier ERP effects compared to those elicited by static actions. Similarly, eye-tracking studies have shown faster saccadic responses when targets were preceded by dynamic as compared to static action (Rohlfing et al., 2012).

Crucially, in the current findings no congruency effect was found in the group of infants who saw biomechanically impossible gestures. This is also in line with previous evidence showing no ERP modulation when the hand gesture was not part of the infant's motor repertoire (Bakker et al., 2015; Melinder et al., 2015). This finding is also in line with recent findings demonstrating that 6-month-olds make more anticipatory gaze shifts toward the goal of a possible, as compared to an impossible, action, showing, thus, to be able to discriminate between biomechanically possible and impossible hand movements (Geangu et al., 2015). The current results are the first to demonstrate that the biomechanical plausibility of a hand gesture can influence the action priming effect. One possible interpretation is that infants quickly discriminate between possible and impossible actions and the outcome of this first evaluation allows them to extract information about directionality from the possible, but not the impossible, action. A second alternative explanation is that hand actions which violate biomechanical properties may represent a novel and more salient visual stimulus compared to the possible hand grasping thus yielding to a greater difficulty in disengaging attention from it.

Overall, the present finding indicates that grasping gestures can trigger 6-month-olds shifts of attention as reflected by modulation of the amplitude of the early N290 ERP component. Importantly, action priming effects were observed only in infants who were presented with biomechanical possible hand gestures, finding that highlights the role of body biomechanical information in action understanding.

2

Chapter 2. Visual discrimination of touching and no-touching gestures early in life.[2]

2.1 Introduction

Results presented in Chapter 1 demonstrated that manual gestures, like grasping hand movements, represent powerful means which can drive infants' attention in relevant locations of the environment from very early in life. Crucially, the study described in Chapter 1 supports recent findings that have highlighted the fundamental role of sensory-motor experience in the development of early sensitivity to human referential manual gestures (Bakker et al., 2015; Melinder et al., 2015). Interestingly, infants start accumulating a great amount of sensory-motor experience already inside the womb and in the first days after birth. For example, recent evidence has demonstrated that organized and coordinated movements can be found even in utero (Myowa-Yamakoshi & Takeshita, 2006; Zoia et al., 2007; Sparling & Wilhelm, 1993). More than 50% of fetal arm movements are directed toward the mouth and it has been shown that fetuses anticipate the hand arrival to their mouth by opening it (Myowa-Yamakoshi & Takeshita, 2006). By 22 weeks of gestation fetuses seem to plan hands movements toward their facial region and adapt the kinematic of their movements according to the part of the face that is going to be touched (i.e., eyes or mouth). Possibly due to these early sensory-motor experience, newborns possess some sophisticated visual perceptual abilities already at birth (Craighero et al., 2010; Longhi et al.,

---

[2] Results presented in this chapter have been previously published in "Addabbo, M., Longhi, E., Bolognini, N., Senna, I., Tagliabue, P., Cassia, V. M., & Turati, C. (2015). Seeing touches early in life. PloS one, 10(9), e0134549."

2015). For instance, neonates seem to have some knowledge of at least some basic characteristics belonging to goal-directed actions (i.e., presence of an object, hand shape, direction of the hand movement related to the body) (Craighero, Leo, Umiltà, & Simion, 2011). Moreover, at birth newborns can discriminate between biomechanically possible and impossible whole hand grasping movements (Longhi et al., 2015).

A primary means of learning, exploring and acquiring sensory-motor experience from foetal life throughout infancy is represented by the haptic modality, which is, among other sensory modalities, the first to develop in utero (Field, 2003). In the womb foetuses perform different hand movements directed towards the environment, like the uterine wall, the umbilical cord or their own body, in particular the face (Sparling, Van Tol, & Chescheir, 1999; Sparling & Wilhelm, 1993). These movements appear to be well organized and coordinated (Zoia et al., 2007, Myowa-Yamakoshi & Takeshita, 2006; Reissland, Francis, Aydin, Mason & Schaal, 2013). At 24 weeks of gestation, half of the foetal arm movements result in hands touching the mouth (Myowa-Yamakoshi & Takeshita, 2006) with the frequency of touch for lower and perioral regions of the face increasing significantly with gestational age. Hand-to-face gestures represent one of the most frequent activities during foetal life, and they remain a prominent behaviour during the earlier stages of postnatal life (Kurjak, Stanojević, Andonotopo, Scazzocchio-Duenas, Azumendi, & Carrera, 2005). In fact, neonates often touch their face, and especially their mouth, during their waking hours (Butterworth & Hopkins, 1988; Lew & Butterworth, 1995). Although it is commonly recognized that touch is very important early in life (Field, 2003; Hertenstein, 2002), we know very little about how infants visually perceive others being touched.

Thanks to prenatal extensive somatosensory-motor experiences neonates seem to possess some cross-modal abilities immediately after birth (Filippetti et al., 2013;2015; Streri & Gentaz, 2003; 2004; Rochat & Hespos, 1997). For instance, in a seminal study conducted by Streri and Gentaz (2003, 2004) it has been shown that newborns are capable to transfer the information of the shape of an object from the tactile to the visual modality. In this study, 2-day-old newborns were first habituated to the shape of an object through the tactile modality and then tested with a visual paired-preference task, in which they were visually presented with a novel and a familiar object. In the visual test phase, newborns prefer to look at the novel object and this was taken as the evidence of crossmodal recognition of the object from the hand to the eye. At birth, newborns are also able to discriminate between external touch and self-stimulation, displaying rooting responses when their cheek is touched by the finger of somebody else, but not by their own finger (Rochat & Hespos, 1997). Likewise, newborns imitate observed facial and hand movements (e.g., Meltzoff and Moore, 1997; Nagy, Pal, & Orvos, 2014). With respect to the visual processing of tactile bodily signals, Zmyj and colleagues (Zmyj, Jank, Shütz-Bosbach, & Daum, 2011) have shown that 7- and 10-month-old infants are sensitive to intersensory visual-tactile contingency. When presented with two videos displaying a lifelike-baby doll whose legs are touched by a hand, infants look longer at the video in which the touch is contingent with a tactile stimulation on their own leg than at a non-contingent touch. By contrast, when presented with a video showing oblong wooden blocks rather than doll legs, infant do not show any visual preference, even in the case of contingent tactile stimulation. A recent study by Filippetti et al. (2013) further showed that 1-to-4-day old newborns look longer towards a video displaying a paintbrush stroking an infant’s cheek in synchrony with a tactile stimulation on their own cheek, than towards a video in which the stroke was asynchronous. Newborns show such a preference only when the observed face is depicted in its canonical upright orientation, but not when it is inverted (Filippetti et al., 2013). The ability to detect intersensory synchrony is seen as crucial in the development of an early sense of one's own body. In fact, the early ability to match the rhythm of a corporeal sensory event to a non-corporeal one provides infants with critical information fundamental to perceive their own body as a differentiated object among other objects in the world (Bahrick, 2013).

Unlike cross-modal visual-tactile capabilities, so far infants' ability to recognize others' gestures, like touching gestures, has not been investigated. We speculate that the extensive somatosensory-motor experience that newborns have accumulated pre- and postnatally can help them a few days after birth to understand and discriminate others' touching gestures on the basis of visual information alone.

We have addressed this issue in two experiments using a visual preference task in which two dynamic images depicting a touching gesture and a no-touching gesture involving a face, a hand or an object were presented; looking times and orienting responses were measured. In Experiment 1 of Chapter 2, 2-day-old (i.e., newborn) and 3-month-old infants were shown gestures involving a moving hand approaching a static face. In Experiment 2, 2-day-old infants were assigned to two conditions, different with respect to the agent performing the gesture: in one condition newborn infants were presented with gestures involving a hand approaching another hand, while in the other condition the agent that performed the gesture was an object.

2.2 Experiment 1

Here we explored the ability of newborns and 3-month-old infants to discriminate between dynamic hand-to-face touching and no-touching gestures. Participants were tested with an infant-control visual preference paradigm, in which they were simultaneously presented with two dynamic images depicting a hand moving towards a static face. In one video, the hand movement ended up with a hand-to-face contact (i.e., touching gesture), in the other video the hand movement terminated before the hand-to-face contact occurred (i.e., no-touching gesture). If prenatal experience is sufficient for developing the ability to discriminate between gestures with or without a tactile component, newborns would discriminate between the two hand gestures. Discrimination would be implied by a significant preference for either the touching gesture, which might be associated with sensorimotor experiences and rewards (i.e. affective touch), or the no-touching gesture, which might be perceived as an unfamiliar, unexpected event.

2.2.1 Method

Participants

The final sample included 18 healthy full-term newborns (10 girls; mean age: 48 h, range: 26-85, mean birth weight: 3177g, Apgar score: at least 8 after 5 minutes) and 18 3-month-old infants (10 girls; mean age: 94 days, range: 82-103 days). Only one 3-month-old and none of the newborns had a twin.

Newborns were recruited at the maternity unit of the San Gerardo Hospital in Monza while 3-month-old infants were recruited via a database of parents who had agreed to participate in the study. Five additional newborns and six 3-month-olds were tested but excluded from the final sample due to fussiness (*n*= 3 newborns; *n*= 4 3-month-olds) or position bias (i.e., looking more that 85% of the time in one direction, *n*= 2 newborns, *n*= 2 3-month-olds). All newborns and infants were tested when they were awake and in an alert state. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committees of the San Gerardo Hospital and the University of Milan-Bicocca. Written informed consent was obtained from parents before the beginning of the study.

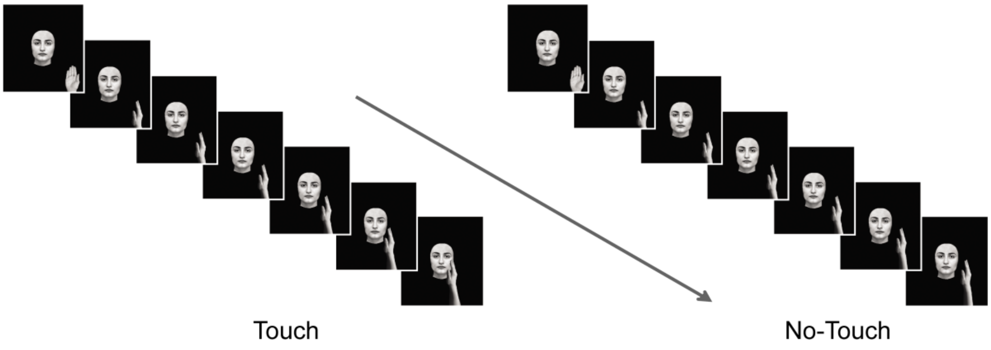
Stimuli

Stimuli were two greyscale videos showing a hand moving towards a static young woman face on a black background. One video showed the hand reaching for the face and touching it on the cheek (touching gesture), while the other video showed the hand stopping at a distance of about 2.5 cm from the face (no-touching gesture). Each video comprised of seven frames. All frames depicted the face in the same frontal pose and central position within the image, whereas the hand changed position across the frames. The first two frames were the same for the touching and no-touching stimuli: Frame 1 depicted the hand in the lower corner of the image with the palm facing the observer; as the hand made a 90° rotation on the vertical axis, Frame 2 presented a sideway hand with the thumb in front and the other fingers aligned vertically. The following 5 frames showed the hand moving diagonally towards the static face, with an angle of 44° in the touching gesture and of 65° in the no-touching gesture (Fig. 1). The amount of movement performed by the hand was the same on the vertical axis in the touching and no-touching stimuli, and differed between the stimuli on the horizontal axis (i.e., 0.5 cm). Both stimuli were presented simultaneously on the screen, and had the same duration (Video A in S1 File): each of the seven frames lasted 571 ms, for a total duration of 4 s. Luminance, contrast, and hue, as well as saturation, were kept constant between frames and stimuli. The videos of touching and no-touching gestures were presented bilaterally and played continuously, in a loop. The dimension of the hand at a distance of 40 cm from the screen ranged between 6.4° and 7.9° of visual angle in width, and between 14.2° and 24.7° of visual angle in height. The face was 14° wide and 21° high, and the distance between the faces depicted in the bilaterally presented videos was 39.2°. The, external contour of the eyes subtended a visual angle of 3.7° X 2°, and the eyes iris was 1 cm in diameter. The face stimulus was taken from our own database (Macchi, Turati, & Simion, 2004).

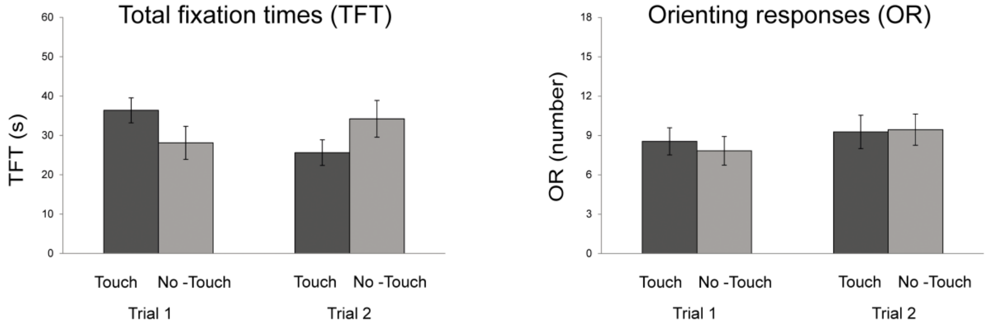
Procedure

Newborns were tested in a dedicated room at the hospital, whereas 3-month-olds were tested in a testing room at the University of Milan-Bicocca. An undergraduate student, blind to the aim of the study, sat with the newborn on the lap in front of the stimulus presentation monitor (27” screen size, 1920 X 1080 pixel resolution, 60 Hz) at a distance of about 30-40 cm. A video camera recording the newborns’ gaze direction was positioned above the monitor; the video camera sent a visual input to a laptop controlled by a second experimenter, who coded online newborns' gaze and ran the experiment, which was designed with E-Prime 2 (Psychology Software Tools). Three-month-olds sat in an infant seat in front of the stimulus presentation monitor (24" screen size, 1600 X 1200 pixel resolution, 60 Hz), at a distance of about 40-50 cm. Both groups were tested using a preferential looking paradigm with an infant-control procedure. Stimulus presentation began as soon as the infant looked at a red flickering circle appearing in the center of the screen; after its disappearance, the two experimental trials were presented. In each trial, touch and no-touch stimuli were shown simultaneously and bilaterally on the screen. The left/right position of the stimuli in the first trial was counterbalanced across participants, and reversed between the first and the second trial. Between the two trials, the central circle re-appeared to attract the infants' gaze. Each trial ended when the infant watched each stimulus at least once, and shifted their gaze away for more than 10 s. On average, newborns watched 16 (SD = 5.29) video sequences in the first trial and 15 (SD = 4.04) in the second trial, while 3-month-olds watched 13 (SD = 4.23) video sequences in the first trial and 16 (SD = 7.49) in the second one. Gaze direction and fixation times were coded online by the experimenter, who was blind to the left/right position of the stimuli on the screen; the number of orienting responses and total fixation times (i.e., sum of all fixations) on the stimuli were recorded as dependent variables (Cohen, 1973). The experimenter used right and left buttons of the mouse to code the corresponding newborns' gaze direction (right, left or none); the duration of infants' fixations corresponded to the duration of button pressing by the experimenter. Video-recordings of eye movements were coded offline for 50% of the infants by an observer, blind to the hypotheses of the study and the stimuli shown. Inter-coder agreement (Pearson correlation) was .91 for total fixation time and .82 for number of orientations for the newborns, and .97 and. 99, respectively, for the 3-month-olds. The Intra-Class Correlation (ICC) coefficient was .94 for total fixation time and .87 for number of orientations for the newborns, and .98 and. 99, respectively, for the 3-month-olds.

a. Stimuli



b. Results - Newborns



c. Results - 3-month-olds

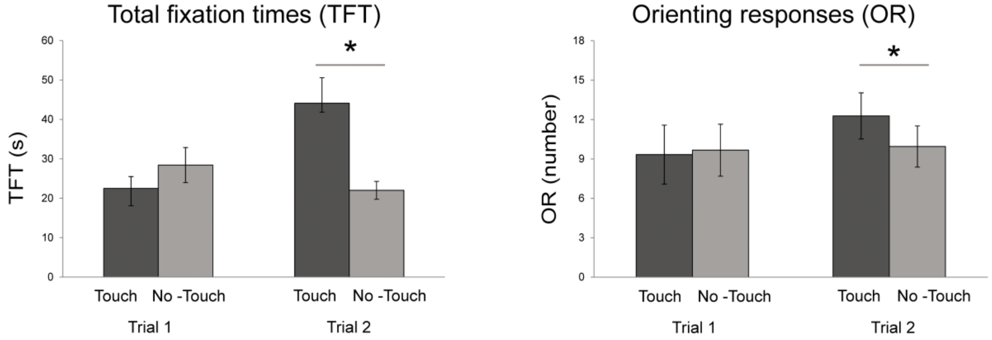


Fig 1. Stimuli and results of Experiment 1. (a) Frames composing the hand-to-face touching gesture (left) and the no-touching gesture (right) videos. Frames are presented in succession in the direction indicated by the arrow. The two stimuli were presented simultaneously on the screen (b) Newborns' total looking times (left) and orienting responses (right) towards the touching and no-touching gesture during trial 1 and 2. (c) 3-month-old infants' total looking times (left) and orienting responses (right) towards the touching and no-touching gesture during trial 1 and 2.
Error bars refer to the standard errors of the mean. * = $p < .05$

2.2.2 Results

To determine whether newborns and 3-month-old infants were able to discriminate between touching and no-touching gestures, two separate repeated-measures Analyses of Variance ($_{rm}$ANOVAs) were performed on total fixation times and number of orienting responses, with *trial* (first vs. second) and *gesture* (touching vs. no-touching) as within-subjects factors, and *age* (newborns vs. 3-month-olds) as the between-subjects factor. The $_{rm}$ANOVA on total fixation times showed a significant *Trial x Gesture x Age* interaction, $F_{1,34} = 9.12$, $p = .005$, $\eta^2 = .093$, as well as a *Trial x Age* interaction, $F_{1,34} = 5.20$, $p = .03$, $\eta^2 = .182$. The 3-way interaction was explored through separate 2-way ANOVAs, one for each age group, with *trial* and *gesture* as within-subjects factors. For newborns, the analysis showed no significant main effects or interaction (all *p*s > .1) (Fig. 1). For 3-month-olds, the analysis revealed a main effect of *trial*, $F_{1,17} = 5.37$, $p = .033$, $\eta^2 = .046$, *gesture*, $F_{1,17} = 5.03$, $p = .038$, $\eta^2 = .053$, and a significant *Trial x Gesture* interaction, $F_{1,17} = 6.62$, $p = .020$, $\eta^2 = .160$. The main effect of *gesture* showed that, overall, infants looked longer at the touching ($M = 66.61$, $SD = 29.78$) than at the no-touching gesture ($M = 50.44$, $SD = 19.48$). The preference for the touching gesture was apparent during the second trial, as revealed by the significant *Trial x Gesture* interaction. In fact, multiple post-hoc comparisons (by means of the Newman-Keuls test) showed that, during the second trial, infants looked longer at the touching gesture (*M* = 44.10, *SD* = 27.45) than at the no-touching gesture (*M* = 22.02, *SD* = 9.61), (p = .047, Cohen's *d* = .78), whereas in the first trial they looked equally long at the touching (*M* = 22.51, *SD* = 12.65) and the no-touching gesture (*M* = 28.42, *SD* = 18.77) (p= .45, Cohen's *d* = .22) (Fig. 1). Looking times to the touching gesture were also significantly longer in the second trial (*M* =44.10, *SD* = 27.45) than in the first trial (*M* = 22.51, *SD* = 12.65), (p = .031, Cohen's *d* = .70). All the others comparisons failed to reach significance (all *p*s > .05).

The analyses on orienting responses confirmed and extended the results obtained for looking times. The 3-way $_{rm}$ANOVA revealed a significant *Trial x Gesture x Age* interaction, $F_{1,34}$ = 4.15, *p* = .049, η^2 =.005. Separate ANOVAs, with *trial and gesture* as within-subjects factors, were then performed for each age group. For newborns, the analysis did not reveal significant main effects or interactions (all *p*s > .2), whereas for 3-month-olds there was a main effect of *gesture,* $F_{1,17}$ = 6.44, *p* = .021, η^2 =.031, as well as a significant *Trial x Gesture* interaction, $F_{1,17}$ = 5.94, *p* = .026, η^2 = .056. The main effect of *gesture* showed that infants oriented their gaze more frequently towards the touching gesture (*M* = 21.61, *SD* = 16.23) as compared to the no-touching gesture (*M* = 19.61, *SD* = 14.43). However, multiple post-hoc comparisons (Newman-Keuls test) for the *Trial x Gesture* interaction showed that it was during the second trial that infants oriented their gaze more often to the touching gesture (*M* = 12.28, *SD* = 7.46), compared to the no-touching gesture (*M* = 9.94, *SD* = 6.65), (p = .008, Cohen's *d* = .67). By contrast, in the first trial they oriented equally towards the touching (*M* = 9.33, *SD* = 9.52) and no-touching gesture (*M* = 9.66, *SD* = 8.39) (p = .67, Cohen's *d* = .16) (Fig. 1). Orienting responses in the second trial were also more frequently directed towards the touching gesture (*M* = 12.27, *SD* = 7.46) as compared to both the touching (*M* = 9.33, *SD*

= 9.52) (p = .007, Cohen’s *d* = .55) and the no-touching gesture (*M* = 9.66, *SD* = 8.39) (p = .01, Cohen’s *d* = .61) of the first trial. All the others comparisons failed to reach significance (all *p*s > .6).

Three-month-olds, but not newborns, were able to differentiate between the two stimuli, as shown by longer looking times and more frequent orientations towards the touching than the no-touching hand-to-face gesture. A possible interpretation of the lack of discrimination in newborns may refer to the saliency of the face stimulus, which may trigger newborns’ attention, preventing them to differentiate the touching vs. no-touching gesture. Indeed, since the first hours of postnatal life, faces are highly salient to infants, capturing their attention under a variety of conditions (Johnson & Morton, 1991; Macchi Cassia, Simion, Umiltà, 2001; Valenza, Simion, Cassia, & Umiltà, 1996). An alternative interpretation of the null result could be that the differences between the two gestures are not marked enough in our stimuli to be detected by newborns' immature visual system. The differences between the two gestures became more evident during the last frames of the videos and, thus, newborns may not have had enough time to detect them. With the aim of disentangling between these different interpretations, in Experiment 2 we investigated newborns’ ability to discriminate between touching/no-touching gestures when the tactile events are directed towards a body part other than the face.

2.3 Experiment 2

Experiment 2 aims at exploring whether newborns are able to discriminate touching from no-touching gestures when presented with hand gestures directed towards a non-face body part, such as the hand. A further aim of Experiment 2 is to examine whether the social nature of the agent performing the gesture is critical in triggering newborns' visual sensitivity to touch. In order to address this issue, Experiment 2 introduced a new condition: a second group of newborns was presented with two dynamic images depicting an inanimate object approaching or touching a static hand.

2.3.1 Method

Participants

The final sample included 34 healthy full-term newborns (13 girls; mean age: 46 h, range: 24-101, mean birth weight: 3410g, Apgar score: at least 8 after 5 minutes), recruited from the maternity unit of the San Gerardo Hospital in Monza. None of the newborns was reported to have a twin.

Newborns were assigned to two different experimental conditions: 17 newborns belonged to the social condition, 17 to the non-social condition. Eight additional newborns were tested but excluded from the final sample due to fussiness (*n* = 4) or position bias (i.e., looking more that 85% of the time in one direction) (*n* = 4). All newborns were tested when they were awake and in an alert state. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committees of the San Gerardo Hospital and the University of Milan-Bicocca. Written informed consent was obtained from parents before the beginning of the study.

Stimuli

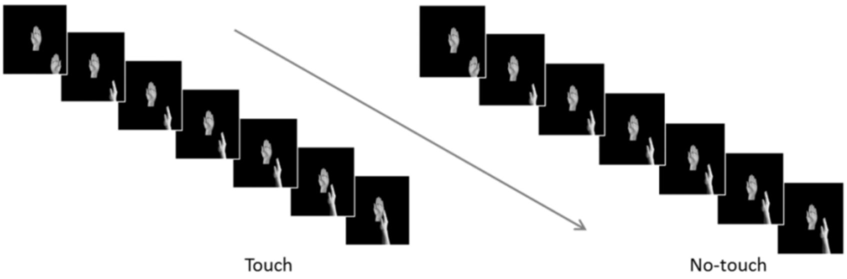
The technical features of the stimuli, as well as the experimental approach, were the same as in Experiment 1. Half of the newborns were presented with touching/no-touching gestures in which a hand moved towards a static hand (social condition), while the other half was presented with a spoon approaching a static hand (non-social condition) (Fig.2). Touching and no-touching stimuli were presented simultaneously on the screen (Videos B and C in S1 File). The dimension of the moving hand at a distance of 40 cm from the screen was the same as in Experiment 1, while the dimension of the spoon ranged between 4.3° and 9.3° of visual angle in width and between 14° and 24.7° of visual angle in height. The static hand was positioned with the palm facing the observer and was 7.9° wide and 17.8° high. The distance between the static hands depicted in the bilaterally presented videos was 39.2°. The moving spoon and the moving hand were positioned in each frame at the same distance from the static hand.

Procedure

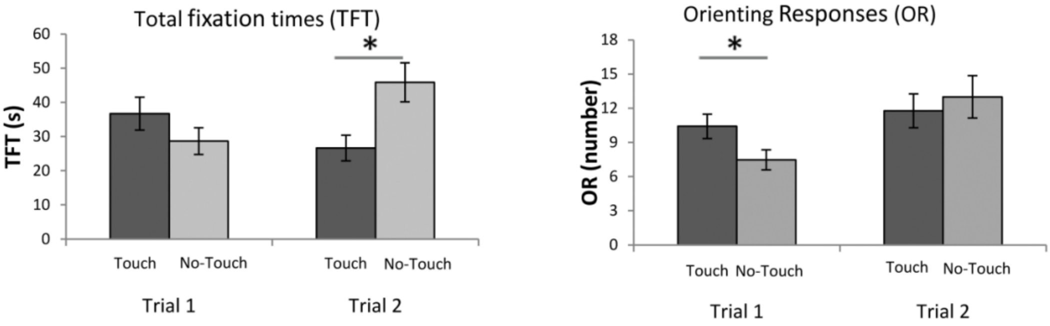
The procedure was the same as in Experiment 1. Newborns watched an average of 14 (SD = 5.72) video sequences in the first trial and 14 (SD = 7.73) sequences in the second trial. The Intercoder agreement (Pearson correlation) on 50% of the participants was .82 for total fixation time and .86 for number of orientations. The ICC coefficient was .99 for both total fixation time and number of orientations.

a. Social Touch

a1. Stimuli

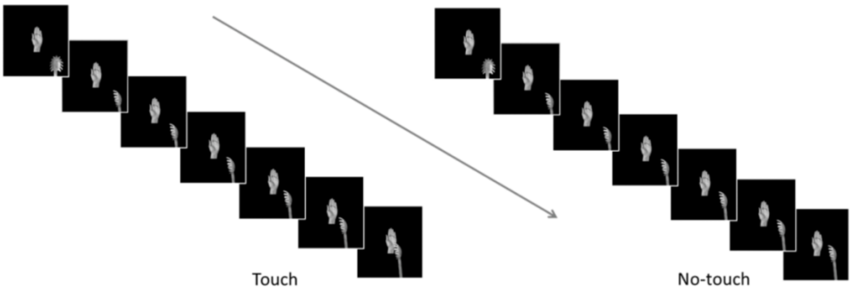


a2. Results



b. Non-Social Touch

b1. Stimuli



b2. Results

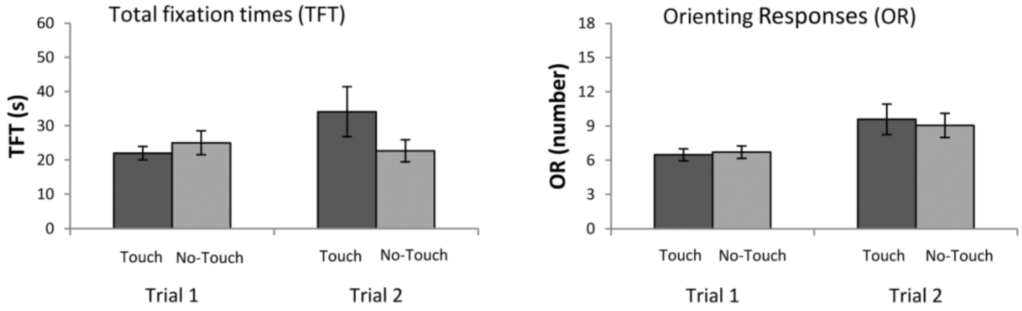


Fig 2. Stimuli and results of Experiment 2. (a) Stimuli and results of the social condition. (a1) Frames composing the hand-to-hand touching gesture (left) and the no-touching gesture (right) videos. Frames are presented in succession in the direction indicated by the arrow. The two stimuli were presented simultaneously on the screen (a2) Newborns' total looking times (left) and orienting responses (right) towards the touching and no-touching gesture during trial 1 and 2 in the social condition. (b) Stimuli and results of the Non-social

condition. (b1) Frames composing the object-to-hand touching gesture (left) and the no-touching gesture (right) videos. (b2) Newborns' total looking times (left) and orienting responses (right) towards the touching and no-touching gesture during trial 1 and 2 in the Non-social condition.
Error bars refer to the standard errors of the mean. * = $p < .05$

2.3.2 Results

The $_{rm}$ANOVA on total fixation times with *trial* (first vs. second) and *gesture* (touching vs. no-touching) as within-subjects factors, and *condition* (social vs. non-social) as the between subject-factor showed a significant main effect of *condition*, $F_{1,32} = 5.78$, $p = .022$, $\eta^2 = .055$, and a significant *Trial x Gesture x Condition* interaction, $F_{1,32} = 7.42$, $p = .010$, $\eta^2 =.08$. The 3-way interaction was explored through separate 2-way ANOVA, with *trial* and *gesture* as within-subjects factors. For the non-social condition, the analysis showed no significant main effects or interaction (all *p*s > .1) (Fig. 2). For the social condition, the analysis revealed a significant *Trial x Gesture* interaction, $F_{1,16} = 6.71$, $p = .020$, $\eta^2 = .120$. Based on the results obtained in Experiment 1, we explored this interaction with a set of planned comparisons, by means of t-test with the Bonferroni correction. During the first trial newborns looked equally longer at the touching ($M = 36.68$, $SD = 19.83$) and the no-touching gesture ($M = 28.68$, $SD = 16.08$), $t_{16} = 1.299$, $p = .424$, Cohen's $d = .315$ (two-tailed), whereas in the second trial they looked significantly longer to the no-touching ($M = 45.87$, $SD = 23.43$) compared to the touching gesture ($M = 26.64$, $SD = 15.41$), $t_{16}= 2.977$, $p = .018$, Cohen's $d = .722$ (two-tailed) (Fig. 2).

The 3-way $_{rm}$ANOVA performed on orienting responses revealed a significant main effect of *condition*, $F_{1,32} = 4.32$, $p = .046$, $\eta^2 = .068$, of *trial*, $F_{1,32} = 17.58$, $p = .0002$, $\eta^2 =.088$, and a significant *Trial* x *Gesture* x *Condition* interaction, $F_{1,32} = 6.54$, $p = .015$, $\eta^2 =.014$. Separate ANOVAs for each *condition* (Social, Non-social) with *trial* and *gesture* as within-subjects factors were then performed. For the non-social condition, the analysis revealed only a significant main effect of *trial*, $F_{1,16} = 12.83$, $p = .002$, $\eta^2 =.215$. Newborns oriented more frequently in the second trial ($M = 18.65$, $SD = 8.35$) than in the first one ($M = 13.18$, $SD = 3.36$), independently of the type of gesture. Differently, for the social condition there was a significant main effect of *trial*, $F_{1,16} = 7.46$, $p = .015$, $\eta^2 = .204$, as well as a significant *Trial* x *Gesture* interaction, $F_{1,16} = 12.95$, $p = .002$, $\eta^2 = .075$. Planned paired t-test (with Bonferroni correction) showed that in the first trial newborns oriented their gaze more frequently towards the touching ($M = 10.41$, $SD = 4.40$) than the no-touching gesture ($M = 7.47$, $SD = 3.61$), $t_{16} = 3.178$, $p = .012$, Cohen's $d = .77$ (two-tailed), whereas in the second trial they oriented their gaze almost equally towards the touching gesture ($M = 11.76$, $SD = 6.13$) and the no-touching gesture ($M = 13$, $SD = 7.63$), $t_{16} = 1.182$, $p = .508$, Cohen's $d = .28$ (two-tailed) (Fig. 2).

Overall, these findings indicate that newborns are actually able to visually distinguish between touching and no-touching gestures involving two hands. Crucially, newborns do not manifest any preference when presented with a non-social touch, such as that provided by an inanimate object like a spoon.

2.4 Discussion

The study reported in Chapter 2 explored the developmental origins of the ability to visually recognize touching gestures involving human body parts (face, hand) and/or an object (spoon). In Experiment 1, only 3-month-olds, but not newborns, manifested a visual preference for a human hand-to-face touching gesture over a no-touching gesture. This indicates that 3-month-olds differentiated between a hand-to-face gesture that led to touch and a comparable gesture in which touch did not occur. Three-month-olds' discrimination between the hand-to-face touching and no-touching gesture was accompanied by a visual preference towards the touching gesture, which was evident for both fixation times and orienting responses.

In Experiment 2 we removed the potential interference effect generated by newborns’ sensitivity to faces by comparing newborns' gaze and looking behavior while watching touching and no-touching gestures directed towards a different human body part, namely a hand. Under this condition, 2-days-old newborns were able to differentiate between touching and no-touching gestures. Specifically, during the first trial newborns' attention was attracted by the touching gesture, as shown by newborns’ orienting responses. Then, in the second trial their attention was held for longer time by the no-touching hand-to-hand gesture, as testified by their total looking times. Crucially, newborns’ preference vanished when the agent of the gesture was an inanimate object, namely the spoon.

A possible interpretation of 3-month-olds’ spontaneous preference for the touching hand-to-face gesture is that gestures that comprise a tactile event are those that infants commonly experience during their daily interactions with others, and that provide them with both somatosensory and affective/communicative information. When observing a hand approaching a face, infants might expect the moving hand to fulfill its communicative and affective goal through touch, consequently making the hand-to-face touching gesture particularly salient and attractive for infants.

Three-month-olds' preference for the touching gesture was apparent in the second trial. This might be due to the saliency of face, which may have captured infants' attention, to the point of masking the difference between the two hand-to-face gestures. Accordingly, infants appreciated the differences between the two hand gestures only during the second trial, when they shifted their attention from the face to the gesture. Future studies might confirm this interpretation by using an eye-tracker procedure to record infants' scanning pattern on the stimuli.

Interestingly, newborns were able to distinguish touching from no-touching gestures when two hands were involved. Newborns' ability to discriminate between these two gestures likely relies on the somatosensory-motor experience accumulated in the womb and in the first hours after birth. Such an experience might drive newborns' expectation that, when a hand is moving towards another hand, the approaching gesture will lead to contact, i.e. a touching event. In this vein, the switch in the direction of newborns' preference between the first and the second trial would imply a switch from a familiarity preference (i.e., for the familiar touching event) to a novelty preference (i.e., for the unexpected no-touching event). Irrespective of the factor driving the change in the direction of newborns' preference across trials, the crucial finding here is that newborns can discriminate touching versus no-touching hand gestures.

It is noteworthy that, in infant research, looking time is typically considered as a more sensitive measure of infants' visual processing, than number of gaze orientations (Cohen, 1972). The direction of visual preference expressed by looking times in the current study was different for newborns and 3-month-olds infants, as the preference was towards the no-touching gesture for newborns in Experiment 2 and towards the touching gesture for 3-

month-olds in Experiment 1. Unlike 3-month-olds, newborns have limited experience with interpersonal bodily interactions involving tactile contact and, thus, they might have also limited ability to decode the affective and communicative implications of touch. During prenatal life, fetuses’ tactile contacts are mainly related to the exploration of their own body and of the surrounding womb environment; within 3-months of postnatal life tactile experiences acquire an affective valence. The affective/communicative relevance that touch acquires through early interpersonal bodily experiences might explain the different visual behavior of newborns and 3-month-olds. On the other hand, there are indications that already during prenatal life fetuses have a natural predisposition to social interactions: when fetuses have to share the uterine environment with their co-twin, they touch and explore their twins’ body and these social contacts increase during the second semester of gestation (Castiello et al., 2010). This raises the question of the effect of such prenatal interpersonal contacts on newborns’ visual processing of touching/no touching gestures.

If newborns successfully discriminate between touching/no-touching body-related, potentially social, gestures thanks to their early somatosensory-motor experiences, their failure to show a preference in the presence of touches when the agent of the touch is an inanimate object may be attributed to their limited experience with gestures involving objects. However, it is important to note that the absence of a preference for either touching or no-touching object-to-hand gestures does not necessarily imply that newborns cannot discriminate them. They just might not have any expectation about the possible outcomes of a gesture that involves an object. The difference in newborns' visual behavior between the hand-to-hand and the object-to-hand condition of Experiment 2 supports the view that, shortly after birth, infants are tuned to human social gestures involving body-parts contact.

The sight of a touching, body-related, gesture provides infants with important information about the social world from very early in life, since touch implies a nonverbal communication of intentions and affect. Given fetuses’ organized and coordinated movements (Reissland et al., 2011) and newborns’ early ability to detect synchrony between an observed and a felt body-related touch (Filippetti et al., 2013) we expected that from the earliest stages of postnatal life infants could visually discriminate between touching and no-touching gestures. Our results support this prediction, highlighting the importance of pre- and post-natal experience in the visual processing of touching gestures involving others' body.

Chapter 3. Matching experienced facial gestures and sounds at birth.[3]

3.1 Introduction

The results reported in Chapter 2 highlighted the importance of pre- and post-natal somatosensory experience in the visual processing of touching gestures involving others' body. However, during the perinatal period newborns acquire also a great amount of orofacial sensory-motor experience. We know from studies employing 4D-sonography that newborns are capable to express positive and negative internal states through facial gestures already during fetal life (Reissland et al., 2011). Moreover, inside the womb they also perform some facial gestures necessary to produce yawns and hiccups and these behaviors become very frequent once they are born (Piontelli, 2014). To our knowledge, to date no study has investigated newborns' processing of audio-visual information coming from facial gestures and sounds that are part of their own sensory-motor experience.

Converging evidence has shown that newborns possess some abilities to match facial gestures to voices already at birth (Aldridge, Braga, Walton, & Bower, 1999; Lewkowicz, Leo, & Simion, 2010; Guellaï et al., 2016). Using an operant sucking procedure Aldridge and colleagues (1999) demonstrated that when newborns were presented with matching or mismatching audio-visual presentations of a face articulating vowels, they preferred the former to the latter. Furthermore, a few days after birth newborns can also perform imitative oral behaviors when hearing vocal and consonant sounds (Chen, Striano, & Rakoczy, 2004) or when seeing face movements (Meltzoff & Moore, 1983). Imitative

---

[3] Results presented in this chapter have been previously presented in a poster in the " XI CEU Conference on Cognitive Development, BCCCD, Budapest, Hungary, January 7-9, 2016".

responses are also quicker when facial gestures are presented along with the sound that they produce (Coulon et al., 2013). Early multisensory abilities have been commonly investigated in infancy using audio-visual matching procedures in which infants are presented with two side-by-side faces each articulating a vowel. The display is accompanied by a sound that matches only one of the two facial movements. If infants are sensitive to the audio-visual correspondence then they are expected to look significantly longer to the facial movement that matches the vowel (Kuhl & Meltzoff, 1982, 1984, 1988; Lewkowicz et al., 2010). Using this procedure, Lewkowicz and colleagues (2010) have found that a few hours old newborns are able to match a monkey's face to the correspondent synchronous vocalization. However, newborns successfully detected the matching audio-visual presentations also when an artificial tone was used instead of the monkey’s natural vocalization. This finding was taken as evidence that very early in life newborns possess some rudimentary multisensory abilities that mostly rely on amodal information redundant across senses like temporal synchrony, specified by the onsets and offsets of the audio and visual inputs (Lewkowicz et al., 2010). A recent study has found that at birth newborns are also sensitive to other amodal information across senses, like prosodic cues (rhythm and intonation) (Guellai, Streri, Chopin, Rider, Kitamura, & 2016). In this study 2-day-old newborns were presented with 2 point-line displays representing a face uttering sentences along with an utterance that matched only one of the two stimuli. To explore if the ability to match directed speech with displays of talking faces was dependent on experience or not, the authors familiarized only one of two groups of infants with the display/sound pairings prior to the testing phase. Both groups of newborns successfully matched the audio-visual presentations regardless of experience.

The importance of redundant stimulations across senses for intersensory development was highlighted by Lickliter and Bahrick (2001) and Bahrick and Lickliter (2002) who developed the “intersensory redundancy hypothesis”, which considers the detection of redundant amodal information (i.e., temporal asynchrony, rhythm, tempo and intensity) as a “cornerstone of perceptual development, allowing optimal deployment of attention and the discovery of higher order perceptual structures” (Bahrick & Lickliter, 2002). Temporal synchrony appears a relatively simple way through which newborns can coherently unify the information coming from different senses.

Despite the fundamental role that audio-visual synchrony cues have in the very first stages of development, infants seem to start to discover higher level audio-visual relations relatively early in life. By 2 months of age, they can match phonetic information from face and voice even in the absence of synchrony cues (Patterson & Walker, 2003) and by 3-5 months of age they can match asynchronous presentations of facial and vocal emotional expressions (Kahana-Kalman & Walker-Andrews, 2001; Vaillant-Molina, Bahrick, & Flom, 2013). Further, 8-month-olds can match gender information across senses (Patterson & Werker, 2002) and, only at the end of the first year of life, infants start to perceive multisensory coherence of fluent audiovisual speech (Lewkowicz, Minar, Tift, & Brandon, 2015) and amodal language identity (Lewcowicz & Pons, 2013).

Together with brain maturation, early multisensory experiences might contribute to the development of the sensitivity to multisensory information (Stein, Perrault, Stanford, & Rowland, 2009; Lewkowicz, 2000). Studies in animals provides strong evidence that multisensory integration is acquired only after accumulating a considerable amount of sensory experience (Stein et al., 2009), but, until now, we know very little about how experience contribute to the development of multisensory functions in humans. Given that some multimodal skills are already present at birth even if they are not necessary fully mature (Lewkowicz et al., 2010), it is plausible to think that prenatal experiences could have a role in the development of these early capabilities. Already during the prenatal period, fetuses are immerse in an environment rich of internal and external stimulations (i.e., tactile, vestibular, chemical, and auditory) (Schaal & Lacanuet, 1999) and they have also some opportunities to experience redundant stimulations across senses. For example, when the fetus hears the mother's speech or non-speech sounds they might also feel a tactile sensation generated by muscles movements involved in producing that sound. Both sound and tactile stimulation might then share the same temporal pattern. Further, when the fetus moves in the womb, this generates proprioceptive feedback as well as a tactile sensation which is the consequence of that motion (Lickliter, 2011). These prenatal multisensory experiences might provide the precursors for the detection of audio-visual contiguous stimulations that newborns will experience only once they are born (Lickliter, 2011). The important role of prenatal experiences on early postnatal multimodal development is well documented in animal-based research. These studies have shown, for example, that modifying normal patterns of prenatal sensory experiences in bobwhite quail chicks alters postnatal multimodal perception (Lickliter, 2000; Markham, Shimizu, & Lickliter, 2008). In humans, important insight on how early postnatal experiences can affect later perceptual abilities, like face recognition, can be found in the study of Sai et al. (2005). The authors have shown that newborn infants manifested a visual preference for their mothers' face compared to a stranger's face only when they had prior experience (from birth to the test) with both her face and voice. Conversely, this early visual preference disappeared when newborns were prevented to hear their mothers' voice. Conversely, Guellaï and colleagues (2016) have demonstrated that newborns' ability to match directed utterances to point-line displays of talking faces does not depend on prior experience but, rather, on prosodic cues redundant across senses.

To sum up, available evidence shows that infants start to learn about the complex multisensory world that surrounds them very quickly and early in life. However, we know very little about multisensory processing at birth. Some multisensory skills are present in some degrees already at birth (Aldridge et al., 1999; Slater et al., 1999; Sai et al., 2005; Lewkowicz et al., 2010; Filippetti et al., 2013,2015; Guellaï et al., 2016) and, a few hours after birth, audio-visual multisensory processing seems to rely on redundant information across senses (Lewkowicz et al., 2000; Bahrick & Lichliert, 2002; Guellaï et al., 2016). However, given the key role of early perinatal experiences on multisensory development, it is reasonable to think that newborns could benefit from the presentation of familiar and experienced audio-visual information to detect multimodal correspondences. To date, no study has investigated newborns' abilities to match facial gestures and sounds that are part of their own experience, like non-speech sounds (yawn, hiccups) and their corresponding facial movement. Crucially, no study has explored whether these familiar audio-visual stimuli could bootstrap newborns' ability to match audio-visual information also in the absence of temporal synchrony.

Here we conducted two experiments to test whether a few hours after birth newborns can match a human face performing a yawn and a hiccup to the corresponding sound in the presence (Experiment 1 of Chapter 2) and absence (Experiment 2 of Chapter 2) of A-V synchrony. Yawns and hiccup were chosen as stimuli because very early in life newborns have already the opportunities to accumulate a great amount of sensory-motor experience with these two behaviors. During the first half of pregnancy, fetuses occasionally yawn and hiccup, while, once they are born, these behaviors become very frequent (Piontelli, 2014; De Vries, & Fong, 2006). Thus, both prenatally and postnatally newborns produce yawns and hiccups. Here we will use an intersensory matching procedure (Lewkowicz et al., 2010) and measure looking times while 2-days-old newborns concurrently viewed two videos of a woman's face performing a yawn and a hiccup in two different experimental conditions: in the absence of auditory stimulation (silent condition) and in the presence of a sound (hiccup or yawn) that matched only one of the two facial gestures (in-sound condition). In the first Experiment sounds will be presented synchronously with the facial gesture while in Experiment 2 the A-V presentation will be asynchronous. In Experiment 1 we expect newborns to be able to associate the sound with the facial gesture, as has been already shown in a previous study with other-species stimuli (Lewkowicz et al., 2010). Crucially, we expect to find the same pattern of responses also in Experiment 2, when sounds and facial gestures are presented asynchronously. We speculate that, with respect to previous literature, the use of familiar stimuli with which newborns have had significant prenatal and postnatal sensorimotor experience can help them to associate a sound to a facial movement even in the absence of temporal synchrony.

3.2 Experiment 1

The aim of this Experiment is to explore whether newborns are able to match a facial gesture producing a yawn and a hiccup with the corresponding sound presented synchronously. We expect them to successfully detect the sound-matching A-V presentations, and, thus, we expect to replicate previous findings that have employed as stimuli unfamiliar and unexperienced face-sound couplings (i.e., monkey faces and monkey vocalizations) (Lewkowicz et al., 2010).

3.2.1 Method

Participants

The final sample included 16 healthy full-term newborns (6 girls; mean age: 43 h, range: 20-79, mean birth weight: 3445 g, Apgar score: at least 8 after 5 minutes). Newborns were recruited at the maternity unit of the San Gerardo Hospital in Monza. Fourteen additional newborns were tested but excluded from the final sample due to fussiness (*n*= 11 newborns) or position bias (i.e., looking more that 85% of the time in one direction, *n*= 3 newborns). All newborns were tested when they were awake and in an alert state. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committees of the San Gerardo Hospital. Written informed consent was obtained from parents before the beginning of the study.

Stimuli

Stimuli were two videos showing a woman face on a black background. One video showed the face performing a yawn, while the other video showed the same actress performing a hiccup. The durations of the yawn and the hiccup were respectively 3000 ms and 1000 ms. Movements produced during yawning and hiccupping comprehend both rigid movements of the whole head and non-rigid motion of the inner features of the face. The videos were presented bilaterally and played continuously, in a loop. The dimension of the face at a distance of 30 cm from the screen was 27.2° of visual angle in width and 38.5° of visual angle in height and the distance between the faces depicted in the bilaterally presented videos was 24.5° (Figure 1).

Procedure

Newborns were tested in a dedicated room at the hospital and an undergraduate student sat with the newborn on the lap in front of the stimulus presentation monitor (27" screen size, 1920 X 1080 pixel resolution, 60 Hz) at a distance of about 30 cm. A video camera recording the newborns' gaze direction was positioned above the monitor; the video camera sent a visual input to a laptop controlled by a second experimenter, which was designed with E-Prime 2 (Psychology Software Tools). A third experimenter then coded offline the total looking times towards the two stimuli. Stimulus presentation began as soon as the infant looked at a white flickering circle appearing in the center of the screen; after its disappearance, four experimental trials were presented. The first two trials were silent (silent condition), thus, only the two faces were shown, without any sound. In the following two trials facial gestures were presented along with a sound that matched only one of them (In-sound condition). During the in-sound condition half of the newborns heard the yawn sound and the other half heard the hiccup sound. Sounds were presented through two speakers positioned on the two sides of the monitor at an equal distance from its center. The onset of the sound was synchronous with both facial movements, but, given the different lengths of the two videos, the offsets were only synchronized to the offset of the corresponding facial gesture. In each trial, stimuli were shown simultaneously and bilaterally on the screen. The left/right position of the stimuli was counterbalanced across participants, and across trials. In each trial, videos were looped continuously for 60 sec. Between each trial, the central circle re-appeared to attract the infants' gaze to the center of the screen. Video-recordings of eye movements were coded offline for 50% of the infants by a fourth experimenter, blind to the hypotheses of the study and the stimuli shown. Inter-coder agreement (Pearson correlation) was .98 for total fixation. The Intra-Class Correlation (ICC) coefficient was .97.

Figure 1. Experimental setting and stimuli presented in Experiment 1 and 2 of Chapter 3.

3.2.2 Results

We used as dependent measure the proportions of looking times towards the sound-matching facial expression calculated separately in each condition (silent and in-sound) by dividing the total looking time towards the matching stimulus by the total looking times to both matching and mismatching stimuli. If newborns can match the facial gesture with the sound, we expect them to look longer towards the matching facial gestures in the in-sound compared to the silent condition (Lewkowicz et al., 2010).

The $_{rm}$ANOVA on proportion of looking times with *condition* (silent vs. in-sound) as within-subjects factor and *sound* (yawn vs. hiccup) as the between subject-factor showed a significant main effect of *condition*, $F_{1,15}$ =8,81, p = .009, η_p^2 = .40. Newborns looked in proportion longer at the matching facial gesture in the in-sound condition (M =60%, SD =

13%) compared to the silent condition (*M* = 46%, *SD* = 17%) (Figure 2). These results confirm that newborns were actually able to match a facial gesture to the correspondent sound and strongly replicate and support previous findings on early intersensory matching abilities at birth (Lewkowicz et al., 2010).

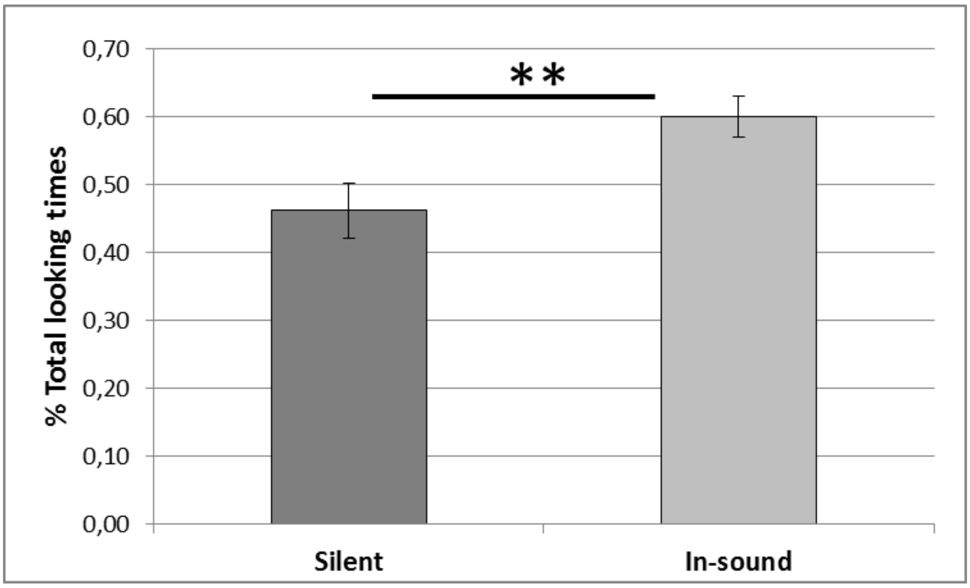


Figure 2. Newborns' proportion of total looking times towards the sound-congruent facial expression in the silent and in-sound condition.
Error bars refer to the standard errors of the mean. ** = p < .01

3.3 Experiment 2

The successful matching obtained in Experiment 1 could be due to the fact that newborns primarily relied on the onsets and offsets of A-V stimulation to make A-V associations. This is basically what Lewkowicz and colleagues (2010) have shown employing as stimuli monkey faces and monkey vocalizations, which are unfamiliar to newborn infants. Our hypothesis is that infants' familiarity and own sensory-motor experience with the A-V information could boost newborns' ability to match the audio-visual inputs even when temporal synchrony is disrupted. We tested this possibility by presenting newborns with the same stimuli of Experiment 1 but the audio track was shifted 700 ms prior to the facial movement, thus making the A-V presentation asynchronous.

3.3.1 Method

Participants

The final sample included 14 healthy full-term newborns (4 girls; mean age: 43 h, range: 21-79, mean birth weight: 3240 g, Apgar score: at least 8 after 5 minutes). Newborns were recruited at the maternity unit of the San Gerardo Hospital in Monza. Sixteen additional newborns were tested but excluded from the final sample due to fussiness (*n*= 11 newborns) or position bias (i.e., looking more that 85% of the time in one direction, *n*= 5 newborns). All newborns were tested when they were awake and in an alert state. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committee of the San Gerardo Hospital. Written informed consent was obtained from parents before the beginning of the study.

Stimuli and procedure

Stimuli and procedure were the same as in Experiment 1 with one exception. The sound track was shifted 700 ms prior to the onset of the facial gestures. This means that differently from Experiment 1, the sound and the facial movement were no longer synchronous. The degree of temporal asynchrony was chosen on the basis of prior studies showing that infants detect A-V asynchrony if the speech and facial movements are separated in time by 666 ms (Lewkowicz, 2000).

Inter-coder agreement (Pearson correlation) was .98 for total fixation. The Intra-Class Correlation (ICC) coefficient was .96.

3.3.2 Results

The $_{rm}$ANOVA on proportion of looking times with *condition* (silent vs. in-sound) as within-subjects factor and *sound* (yawn vs. hiccup) as the between subject-factor showed a significant main effect of *condition*, $F_{1,} = 8.77$, $p = .01$, $\eta_p^2 = .42$. Newborns looked in proportion longer at the matching facial gesture in the in-sound condition (*M* =67%, *SD* = 16%) compared to the silent condition (*M* = 44%, *SD* = 17%) (Figure 3). Thus, even in the absence of temporal synchrony newborns could associate the sound with the matching facial gesture.

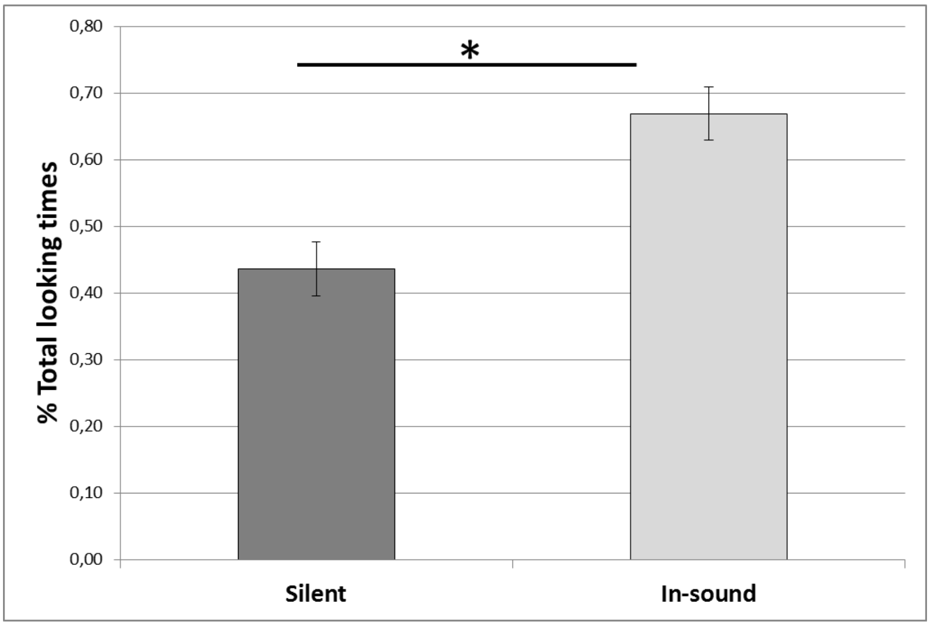


Figure 3. Newborns' proportion of total looking times towards the sound-congruent facial expression in the silent and in-sound condition.
Error bars refer to the standard errors of the mean. * = p < .05

3.4 Discussion

Both prenatal and postnatal environment are rich of information coming from different senses and, since the baby is born, it is extremely important for him/her to respond properly to the surrounding multimodal events. Here we explored the possibility that, at birth, newborns could benefit by the presence of audio-visual information that is part of their sensorimotor experience to detect audio-visual relations even when synchrony cues are not available. In Experiment 1, we replicated previous findings with newborn infants (Lewkowicz et al., 2010) by showing that 2-day-old newborns can match a facial gesture of a yawn and a hiccup to the corresponding sound when they are presented as they naturally occur in everyday life, that is in synchrony. Consistent with our hypothesis, in Experiment 2 we have found that when newborns were presented with familiar multimodal information, temporal audiovisual synchrony seemed no longer necessary to make audio-visual associations. In this second experiment the sound was presented 700 ms prior to the facial gesture and, as in Experiment 1, newborns in proportion looked longer to the sound-matched facial gesture in the in-sound condition compared to the silent condition.

The importance of experience is evident in a couple of studies that have investigated infants' detection of amodal affect from face and voice (Walker-Andrews, 1986; Kahana-Kalman & Walker-Andrews, 2001). These studies have shown that 7-month-olds, but not 5-month-olds, can match a stranger's affectively congruent facial and vocal emotional expression (Walker-Andrews, 1986). Interestingly, when infants were presented with the face of their own mother they can detect amodal affect already at 3,5 months of age (Kahana-Kalman & Walker-Andrews, 2001). According to recent theories of multimodal development (Lewkowicz, 2000; Bahrick & Lichliter, 2002), early in life newborns rely to a great extent on low-level information (i.e., temporal synchrony) to make sense of the myriad of multimodal information that daily reach their senses. Then, gradually and through experience, they discover high-level relations between inputs coming from different sensory channels and the importance of low-level amodal information starts to decline (Lewkowicz, 2000). Here we have demonstrated that multisensory experiences accumulated inside the womb and in the first days of life might be fundamental in the development of early multimodal skills. When newborns observe facial gesture and sounds that they themselves have previously experienced, like yawns and hiccups, they are able to bind audio-visual information even without relying on synchronous cues. Temporal synchrony might be essential to match information across senses early in life only when infants have nor or little experience with that particular audio-visual association. Accordingly, amodal cues might start to become less essential for audio-visual integration after repeated daily exposure with the co-occurent audio-visual inputs. In the present study newborns were presented with A-V information that was part of their sensorimotor repertoire, thus, they might have been a step ahead in the developmental trajectory of multisensory capabilities. Differently, when newborns are presented with unexperienced audio-visual information, they need to rely on temporal synchrony to detect A-V associations.

On which information newborns relied to match audiovisual inputs? Our results don't rule out the possibility that newborns relied on low-level relations -other than temporal synchrony- to match facial gestures and sound of yawns and hiccups. Duration, for example, is another low-level information that newborns could alternatively use to make audio visual associations regardless of amodal temporal cues. However, it is unlikely that they matched the facial gesture to the corresponding sound on the basis of duration given that previous studied with older infants (i.e. 3-month-olds) have shown that they don't use this low-level information to match A-V stimuli (Lewcowicz, 1986).

A possible alternative interpretation is that they relied on high-level relations to make the A-V matching. This means that they recognized the nature of the information coming from the two different channels as belonging to the same event possibly due to early sensorimotor experiences accumulated during fetal life. Fetuses' and newborns' experience with yawn and hiccups derives both from internal and external stimulations. In fact, newborns start to yawn and hiccup already in the womb and these behaviors will become more and more prominent at birth (Piontelli, 2014). Besides these internal self-generated experiences, fetuses also experience their mother producing yawns and hiccup: thus, they will hear the sound of her yawning and hiccupping along with a tactile stimulation due to a change in pressure of their prenatal environment (Lichliter, 2012). However, it is unlikely that newborns also have had the opportunity to acquire A-V experiences by seeing/hearing others yawning or performing hiccups, given their little visual experience in the first few days of life. Thus, most of the multimodal experiences accumulated during fetal life and in the very first few days after birth with yawns and hiccups might derive primarily from sensorimotor stimulations.

All in all, the present study is the first to demonstrate that at birth newborns are able to match facial gestures and non-speech sounds that are part of their experience in the absence of amodal cues. Pre- and postnatal sensorimotor experience is likely involved in the development of these early skills which will refine during development thanks to significant postnatal experience and brain maturation.

Conclusion

Since the first moments after birth humans are introduced in a complex social environment rich of external stimulations. Among all the stimuli that they encounter in their surrounding world, human gestures, like grasping and pointing, are extremely important due the fundamental role that they play during social interactions. The ability to understand the directions of others' gestures undergoes significant development during the first year of life and experience plays a crucial role in this process. In the study presented in Chapter 1 we met two important goals: firstly, we have shown that at 6 months of age, action priming leads to neural sensory facilitation in processing targets that appear in the location cued by a grasping gesture. Secondly, our results are the first demonstrating that the biomechanical plausibility of a hand gesture can influence action priming effect. An impossible grasping hand movement doesn't seem to represent a powerful social signal that can trigger efficiently infants' shifts of attention. Through the sensorimotor experience accumulated in the first 6 months of their life, infants might have acquired some knowledge about the biomechanical constrains of a human hand. This knowledge might have allowed them to understand that only gestures that are possible to perform convey relevant social information. Accordingly, only plausible gestures may have the power to efficiently orient their attention in space. Experience helps infants to make sense of all the social signals that they encounter in their everyday life allowing them to attribute a social meaning only to those gestures that are familiar and part of their sensorimotor repertoire. But also the environment that surrounds newborns in the very first days of life is rich of sensory stimulations. Interestingly, newborns seem to possess already at birth some perceptual abilities that help them to deal with the myriad of sensory information that arrive to their senses. Further, early pre- and postnatal sensorimotor experience might be crucial for the development of these early perceptual skills. Indeed, in Chapter 2 and 3 we supported the idea that sensorimotor experience could be involved in the development of the sensitivity to human gestures already at birth. In particular, in the studies presented in Chapter 2, we have shown that newborns are able to visually discriminate a hand-to-hand touching gesture from a hand-to-hand non-touching gesture, but they are not able to do so when presented with object-to-hand touches. Moreover, infants' spontaneous preferential responses change between the first days of life and 3 months of age possibly as a result of further somatosensory-motor, visual and affective/communicative experiences accumulated during development. The present findings speak in favor of an early ability to visually recognize touching gestures involving the interaction between human body parts. Further, our results highlight the role of somatosensory experience in the development of this early capability. Thanks to the extensive experience acquired in the confines of the womb newborns might possess the ability to transfer the information coming from sensorimotor experience to the visual modality allowing them to visually recognize touching gestures.

Lastly, in the studies described in Chapter 3, we focused on newborns' processing of facial gestures. Specifically, we aimed at exploring whether newborns could benefit by the presence of audio-visual information that is part of their sensorimotor experience to detect audio-visual relations even when synchrony cues are not available. Results confirmed our hypothesis showing that newborns could associate the facial gesture to the sound even without temporal synchrony. Thus, in the presence of experienced A-V information newborns show to be a step ahead in the developmental trajectory of their multisensory capabilities demonstrating to be able as soon as birth to rely on high-level information to make A-V associations rather than on amodal cues.

Overall, the findings described in Section 1 lead to the suggestive hypothesis that from very early in life infants are particularly attuned to social signals like grasping gestures, touching gestures and facial gestures as well. All these gestures represent an extremely relevant source of information in the environment that infants and newborns can use to make sense of the social complex world that surrounds them. Early sensorimotor experiences play a crucial role in the processing of the social world by supporting newborns' ability to recognize others' social gestures at birth and leading, later in development, to a progressive refinement and enrichment of these skills.

Section 2. The processing of dynamic facial expressions at birth and the interplay between action and emotion in infancy.

Introduction

The ability to express and understand others' emotions is distinctly human in nature and plays an extremely important role in our lives since the very first stages of development. The first year of life is an amazing time of growth where babies learn how to show their inner states but also to understand others' emotions. Some emotions, like distress and disgust, occur from birth (Reissland, Francis, Mason, & Lincoln, 2011; Sato, Kanenishi, Hanaoka, Noguchi, Marumo, & Hata, 2014) and from the first days of life newborns also show some visual discriminative abilities of facial expressions (Farroni, Menon, Rigato, & Johnson, 2007; Rigato, Menon, Johnson, Faraguna, & Farroni, 2011). A couple of studies have found that newborns seem to be sensitive to emotional facial expressions already at birth (Farroni et al., 2007; Rigato et al., 2011). At 3 months infants can discriminate between happy and frowning faces (Barrera & Maurer, 1981) and between happy and surprised facial expressions (Young, Browne et al., 1977) but not between happy and sad faces (Young, Browne et al., 1977) while 4- to 6-month-olds can differentiate between angry and fearful expressions (Serrano, Iglesias, & Loceches, 1992). Sensitivity to facial emotional expressions improves during development. By 7 months infants are able to categorize facial expressions (de Haan, & Nelson, 1998; Leppanen, & Nelson, 2006, 2009) and between 5 to 7 months they can recognize facial expressions across variation of identity and intensity (Bornstein, &

Anterberry, 2003; Ludemann, & Nelson, 1988; Nelson, Morse, & Leavitt, 1979). Further, a more recent line of research has explored infants' sensitivity to emotions portrayed by bodies, and has shown, within an intermodal preferential matching procedure, that 6.5 month-olds but not 3.5 months-olds can match emotional body postures to the correspondent emotional vocalization (Zieber, Kangas, Hock, & Bhatt, 2014). Interestingly, evidence demonstrates that early in life infants are more attracted by smiling faces compared to angry and neutral ones (Grossman, Striano, & Friederici, 2007; La Barbera, Izard, Vietze, & Parisi, 1976; Wilcox, & Clayton, 1968). As infants start to crawl and walk they get more exposed to negative expressions from their caregivers and their attention starts to be biased towards negative faces (Vaish, Grossmann, & Woodward, 2008). For example, at 7 months of age infants prefer to look at fearful faces rather that at happy ones (de Haan, belsky, Reis, Volein, & Johnson, 2004; Kotsoni, de Haan, & Johnson, 2001; Peltola, Leppanen, Maki, & Hietanen, 2009). But it is not until they reach the first year of life that they start to use the information coming from others' facial expression to guide their behavior in their environment and develop a fundamental socio-cognitive skill called social-referencing (Campos & Stenberg, 1981). All in all, discriminative abilities of facial and body expressions emerge very early in life. Further, infants' visual behavior towards emotional expressions changes as a result of experience, which seems to play a crucial role since the very first steps of postnatal life (Farroni et al., 2007; Rigato et al., 2011).

However, despite the huge amount of literature that, in the last two decades, has explored emotion processing in infancy, there are still some open questions that need to be further deepened. In particular, a critical point that characterizes current studies on emotion processing in infancy is related to the type of stimuli used. Most of these studies employed stimuli with poor ecologic validity, mostly static stimuli and faces completely isolated from the surrounding context. However, since birth, infants are mainly exposed to emotional expressions that unfold over time. Thus, it is fundamental to explore how infants process emotional facial expressions as they naturally take place in everyday life, that is in motion. It is reasonable to think that a rudimental mechanism that allows them to discriminate dynamic facial expressions of emotion might be already functional at birth and will provide the basis for the development of later and more complex social skills. We know that already at birth newborns are sensitive to motion information. For example, newborns can perceive illusory contours (Valenza & Bulf, 2007) and discriminate between possible and impossible hand movements (Longhi et al., 2015) only when dynamic information is available. In addition, Bulf and Turati (2010) have shown that newborns successfully use information coming from rigid head movements to recognize a face identity posed in a novel viewing perspective. Understanding how newborns process emotions as they actually see them (in motion), which is the aim of Chapter 4, will provide new insights on early perceptual abilities at birth.

Most of the current studies have addressed facial expressions in isolation. However, in everyday interactions facial expressions always refer to someone or something in the environment and provide a context that help to predict others' behaviors. Actions are most of the times embedded in an emotional context and it is fundamental in life to be able to link effective- to action-related information. Recently, Hepach & Westermann (2013) have shown that at the end of the first year of life infants seem to be able to bind a type of action (i.e. patting or thumping a toy) to a specific facial expression. The ability to link actions to emotions emerges as soon as infants start to show social referencing skills. Thus, infants may need more sophisticated social-cognitive skills, acquired only around the end of the first year of life, to develop this ability. However, the precursors of the ability to bind affective to action information might be traced earlier in life. Previous evidence has shown that emotions are extremely salient stimuli that can facilitate infants' processing of relevant information present in the environment. For example, by 3 months of age infants are facilitated in the recognition of face identity when faces display a happy emotional expression (Brenna, Proietti, Montirosso, & Turati, 2013; Turati, Montirosso, Brenna, Ferrara, & Borgatti, 2011). Further, by 3 months of age the processing of novel objects is modulated by emotional expression in combination with eye gaze (Hoehl & Pauen, 2011; Hoehl & Striano, 2010; Hoehl, Wiese, et al., 2008). Thus, we hypothesize that also action processing could be modulated by the emotional context at 6-months of age. This hypothesis will be tested in Chapter 5, by means of Electroencephalography.

Chapter 4. Dynamic facial expressions of emotions are discriminated at birth.[4]

4.1 Introduction

We are born with a natural propensity to communicate our internal states through facial expressions: we wrinkle our nose and elevate the upper lip when we experience disgust while raising the corners of the mouth is a visible evidence of joy. Complex facial movements begin to develop already within the confines of the womb (Reissland et al., 2011; Sato, Kanenishi, Hanaoka, Noguchi, Marumo, & Hata, 2014). By performing these facial movements the fetus provides himself with crucial motor experience for the subsequent emergence of a number of vital functions essential after birth, like breast feeding and vocalizing (Finan, & Barlow, 1998; Green, & Wilson, 2006). Recent studies have also demonstrated that fetuses display facial muscle configurations that can be associated to the expression of distress and positive states and may thus be considered important components for the early interactions with the social world once the baby is born (Reissland et al., 2011). Indeed, fresh from birth newborns are capable to express internal states like distress and disgust (Rosenstein & Oster, 1988), thus being able to communicate and send signals to their parents from the very first stages of their life. Closely tied to the production of facial expressions is the ability to visually discriminate between them. To date, little has been done to investigate this ability at birth.

---

[4] Results presented in this chapter have been previously presented in a poster in the “X° Conference on Cognitive Development, BCCCD, Budapest, Hungary, January 8-10, 2015” and in the "17th European Conference on Developmental Psychology, ECDP , Braga, Portugal, September 8-12, 2015”.

Only a few studies have investigated newborns’ sensitivity to emotional facial expressions (Field, Woodson, Greenberg, & Cohen, 1982, 1983; Farroni et al., 2007; Rigato et al., 2011). Field and colleagues (1982, 1983) have shown that at birth newborns seem to be able to imitate emotional facial expressions like surprise, happiness, or sadness posed by a live model. So far, this result still remains controversial as a further attempt to replicate this finding yielded to contrasting results (Kaitz, Meschulach-Sarfaty, Auerbach, & Eidelman, 1988). More recently, a couple of studies (Farroni et al., 2006; Rigato et al., 2007) have explored newborns’ early ability to distinguish between different static emotional facial expressions through the visual preference and the habituation paradigms. Results from the study of Farroni and colleagues (2006) have shown that when newborns were presented simultaneously with photographs of a happy and a fearful facial expression, they preferred to look at the happy face. Conversely, newborns didn’t show any preference and didn’t even discriminate when the fearful face was compared to a neutral one. With the aim to extend these results, a subsequent study (Rigato et al., 2011) has shown that when facial expressions (neutral, happy or fearful) and gaze direction (averted or directed) were combined, newborns manifested a visual preference only when happy and neutral faces with directed gaze were compared, looking longer toward the happy face (Rigato et al., 2011). The authors suggested that during the very first days of postnatal life newborns are mostly exposed to smiling faces with directed gaze and this facial expression is crucial to promote social interactions (Rigato et al., 2011). This is in line with evidence demonstrating that early in life infants are more attracted by smiling faces (for a review see Hoehl, 2013), and are facilitated in the recognition of face identity when faces display a happy emotional expression (Brenna et al., 2013; Turati et al., 2011). However, both of these studies used as stimuli static photographs in which the perceptual differences between the two facial expressions were very pronounced. Therefore, it could be possible to speculate that newborns were responding to low-level perceptual features present only in the happy face (i.e. the broad toothy smile). Indeed, positive results were found only within visual preference tasks, which are particularly affected by the presence of salient perceptual features early in life (see Slater, 1995 for a review). Also Kestenbaum and Nelson (1990) showed that when a salient feature like a toothy smile was available, discriminative abilities of emotional facial expressions were driven by this perceptual attribute even at 7 months of age.

Most importantly, much of the current knowledge about visual processing of facial expressions in infancy as well as in adulthood comes from studies that have used static stimuli. But, during our daily interactions, especially in our first days of life, we mainly encounter facial expressions that unfold over time. In the last decades many researchers started to comprehend the importance of studying facial expressions in the way they naturally take place in everyday life (i.e. dynamic) (for an overview see Krumhuber, Kappas, & Manstead, 2013) and this shift of attention from static to dynamic stimuli led to a number of relevant findings. First of all, there is evidence that static and moving faces are processed differently: both adults (Võ, Smith, Mital & Enderson, 2012) and infants (Hunnius & Geuze, 2004; Wilcox, Stubbs, Wheeler, & Alexander, 2013; Xiao, Quinn, Liu, Ge, Pascalis, & Lee, 2015) scan differently dynamic and static faces. Further, dynamic displays of facial expressions activate in adults different brain areas compared to still pictures (Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Arsalidou, Morris, & Taylor, 2011). Interestingly, dynamic displays improve adults accuracy in emotion recognition tasks especially when visual information is limited and degraded (Kätsyri & Sams, 2008; Cunningham & Wallraven, 2009). This finding suggests that also newborns, whose visual system in the first few days of life is very immature and characterized by a poor spatial resolution and contrast sensitivity (Atkinson, Braddick, & French, 1979), could benefit from the information coming from facial dynamics when perceiving emotional expressions. In fact, facilitative effects of motion related information can be found already at birth (Valenza & Bulf, 2007; Bulf & Turati, 2010; Longhi et al., 2015).

To sum up, so far there's little evidence concerning newborns' ability to process others' emotional expressions. Moreover, studies that have directly explored newborns' ability to visually discriminate between different emotional expressions have used only static face pictures (Farroni et al., 2006; Rigato et al., 2011). At birth, newborns possess relatively little visual experience and what they actually see in their first days of life are dynamic facial expression and not static faces. To date, it still remains unknown whether at birth newborns are sensitive to others' dynamic facial movements that express emotions. When we express an emotion, a combination of internal features of our face (i.e. eyes, nose, cheeks, mouth) dynamically change configuration over time. Each emotion is the result of a particular combination of these internal features. Are newborn infants able to detect the morphological changes that take place during the unfolding of a facial expression? Are they able to discriminate between different emotional dynamic expressions?

Here we explore whether 2-day-old newborns are sensitive to facial movements expressing happiness and disgust using a visual preference (Experiment 1) and a habituation task (Experiment 2). The happy and disgusted expressions are conveyed by information coming from a combination of facial regions (i.e. eyes, mouth) which changes are comparable in saliency (i.e. narrowing of the eyes, opening of the mouth). Thus, the choice of these facial expressions of emotions reduces the possibility that newborns could manifest a preference or discriminate on the basis of anything other than highly salient features (i.e. eyes wide open, visible teeth).

4.2 EXPERIMENT 1

In this first study we used a visual preference task to explore whether newborns can manifest a visual preference when presented simultaneously with two different dynamic displays depicting a smiling face and a disgusted one. We expect them to show a visual preference for the smiling face, as shown by previous studies with newborn and older infants (Farroni et al., 2007; Rigato et al., 2011; Grossmann, Striano, & Friederici, 2007).

4.2.2 Method

Participants

Eighteen healthy full-term Caucasian newborns (9 girls; mean age: 44 h, range: 21-83, mean birth weight: 3328 g, Apgar score: at least 8 after 5 minutes from birth) recruited at the maternal unit of the San Gerardo Hospital of Monza were tested when they were in an awake and alert state. We have tested other 4 newborns but they were then excluded from the final sample due to fussiness or being not cooperative (n= 3) or to a position biased (i.e. looking towards the right or the left position for more than the 85% of the total looking time) (n= 1). Parental informed consent was obtained before testing began. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committee of the X Hospital.

Stimuli

Newborns were presented simultaneously with two color videos of a woman's face performing a happy and a disgusted expression on a black background. The faces of two different Caucasian women were used (face A and face B). Two face identities were employed as stimuli in order to avoid the possibility that newborns' visual preference could be ascribed to the salient features of one single identity expressing an emotional expression. Half of the newborns saw face A and the other half saw face B, randomly assigned. In both identities women had a direct gaze and their hair, ears and neck as well, were not visible. Each video lasted 4568 ms and was made of 8 frames, each one of the duration of 571 ms. The first two frames depicted the face with a neutral expression and then, in the following 6 frames, the happy/disgusted expression unfolded and reached the maximum intensity in frame 8 (Figure 1a). The two videos were shown bilaterally at a distance of 27°, and played continuously, in a loop. At a distance of 30 cm from the screen face A was 24° wide and 33° high, and face B was 24.8° wide and 33° high. The diameter of the iris was 1.9° for both faces A and B. The Luminance, contrast, and hue, as well as saturation, were kept constant between the frames and the stimuli.

Procedure

Newborns were seated on the lap of an undergraduate student at a distance of 30 cm from the stimulus presentation monitor (27” screen size, 1920 X 1080 pixel resolution, 60 Hz). A camera was placed above the monitor that recorded newborn’s gaze thus allowing an experimenter to code online newborns’ visual behavior. The baby holder could check if newborns’ position was aligned to the center of the screen on a monitor that displayed his/her face. Total looking times and number of orienting responses were measured within a preferential looking paradigm with an infant-control procedure (Farroni et al., 2006; Rigato et al., 2007). Newborns were presented with two trials, in which a happy and a disgusted face were displayed simultaneously and bilaterally on the screen. Each trial began as soon as the newborns looked at a red flickering circle appearing in the center of the monitor. The left/right position of the stimuli was reversed between the first and the second trial and in the first trial half of the newborns saw the happy face on their right and the other half on their left. Each trial ended when the newborns watched each stimulus at least once, and shifted their gaze away for more than 10 s. Half of the video-recordings of eye movements were coded offline by an observer, blind to the stimuli shown. Inter-coder agreement (Pearson correlation) was .97 for total fixation time and .82 for number of orientations. The Intra-Class Correlation (ICC) coefficient was .98 for total fixation time and .85 for number of orientations, revealing an excellent absolute agreement between coders.

4.2.3 Results

A repeated-measures Analysis of Variance ($_{rm}$ANOVAs) was performed with *trial* (first vs. second) and *emotion* (happy vs. disgust) as within-subjects factors. The analysis showed no significant main effects or interaction both in total looking times (all ps > .295) and eye orienting responses (all ps > .187) (Figure 1b). Thus, contrary to our hypothesis, results of Experiment 1 show that newborns didn’t prefer a facial expression over the other.

4.3 EXPERIMENT 2

The lack of preference shown by newborns in Experiment 1 doesn’t directly imply that they weren’t able to differentiate between happy and disgusted faces. The presence within a visual preference task of dynamic stimuli could have distracted infants from detecting the differences between the two facial expressions. Moreover, it is also possible that newborns simply didn’t prefer an emotion over the other, even if they were able to discriminate between them. To ensure whether newborns were capable to distinguish between the two dynamic expressions despite the null result revealed during the preference task, we conducted a second experiment in which we used a visual habituation paradigm. If newborns are able to discriminate between the two facial expressions, then we expect them to look longer and orient more frequently towards the novel facial expression in test phase.

4.3.1 Method

Participants

Eighteen healthy full-term newborns recruited at the maternal unit of the San Gerardo Hospital of monza (11 girls; mean age: 46 h, range: 24-82, mean birth weight: 3349g, Apgar score: at least 8 after 5 minutes from birth) were tested when they were in an awake and alert state. We have tested other 6 newborns but they were then excluded from the final sample due to fussiness or being not cooperative (n=4) or to a position biased (i.e. looking towards the right or the left position for more than the 85% of the total looking time) (n=2). Parents signed a written informed consent before testing began. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethics Committee of the X Hospital.

Stimuli

Stimuli were the same as in Experiment 1.

Procedure

The experimental setting was the same as in Experiment 1. Newborns were tested using a visual habituation paradigm and total looking times and orienting responses were measured (Farroni et al., 2006). The habituation phase started when the newborn oriented towards a red flickering circle appearing in the center of the screen. During the habituation phase, newborns viewed bilaterally two videos of a face expressing the same emotion (happy or disgust). Half of the newborns were habituated to the happy faces, the other half to the disgusted faces. As in Experiment 1, two identities were used (face A and face B). In each habituation trial, videos were cycled continuously in a loop until the newborn shifted his/her gaze away for more than 2 s. Habituation phase ended when the newborn reached the habituation criterion which was a 50% decline in looking time on the last three consecutive trials, relative to the looking time on the first three trials. Following habituation, newborns were presented with two test trials in which a novel facial expression (happy for newborns habituated to disgusted facial expressions and vice-versa) and a familiar facial expression were displayed simultaneously and bilaterally on the screen. Each test trial began as soon as the newborns looked at a red flickering circle appearing in the center of the monitor. The left/right position of the stimuli was reversed between the first and the second trial and in the first trial half of the newborns saw the happy face on their right and the other half on their left. Each test trial continued in a loop until newborn looked for a minimum of 20 s and ended when each stimulus was watched at least once and when their gaze was shifted away for more than 500 ms. Half of the video-recordings of eye movements were coded offline by an observer, blind to the stimuli shown. Inter-coder agreement (Pearson correlation) was .91 for total fixation time and .84 for number of orientations. The Intra-Class Correlation (ICC) coefficient was .96 for total fixation time and .87 for number of orientations.

4.3.2 Results

All newborns reached the habituation criterion. In fact, a repeated-measures Analyses of Variance ($_{rm}$ANOVAs) on total looking times with *habituation condition* (happy vs. disgust) as the between-subjects factor, and *habituation trials* (first three vs. last three) as the within-subjects factor revealed a significant effect of habituation trials, $F(1,16) = 31.3$, $p < .001$, $\eta_p^2 = .662$. The average looking time on the first three habituation trials (M= 54.8 s, SD= 29.7) was significantly longer than the average looking time on the last three habituation trials (M = 17.2 s, SD= 9.9). No other effect was significant. On average, newborns required 6.75 trials to habituate to the happy face, and 7 trials to habituate to the face expressing disgust.

A repeated-measures Analyses of Variance (rmANOVAs) on total looking times with *trial* (first vs. second) and *novelty* (novel vs. familiar) as within-subjects factors and *habituation condition* (happy vs. disgust) as between-subjects factor, revealed a main effect of novelty, $F(1,16) = 4.87$, $p = .042$, $\eta_p{}^2 = .233$. Infants looked significantly longer at the novel ($M = 30.7$ s, $SD = 8.5$) than the familiar ($M = 22.8$ s, $SD = 9.2$) facial expression during test phase (Figure 1c). The other factors and interactions were not significant (all ps > .154).

A repeated-measures Analyses of Variance (rmANOVAs) on eye orienting responses with *trial* (first vs. second) and *novelty* (novel vs. familiar) as within-subjects factors and *habituation condition* (happy vs. disgust) as between-subjects factor revealed a main effect of novelty, $F(1,16) = 7.31$, $p = .016$, $\eta_p^2 = .314$. Infants oriented more frequently towards the novel ($M = 10.1$ s, $SD = 4.7$) than towards the familiar ($M = 7.5$ s, $SD = 3.7$) facial expression during test phase (Figure 1c). No other significant effect or interaction emerged from this analysis (All ps > .071). These results show that newborns are able to discriminate between a dynamic happy and disgusted expression, as indicated by longer looking times and more frequent eye orientations towards the novel stimulus in test phase.

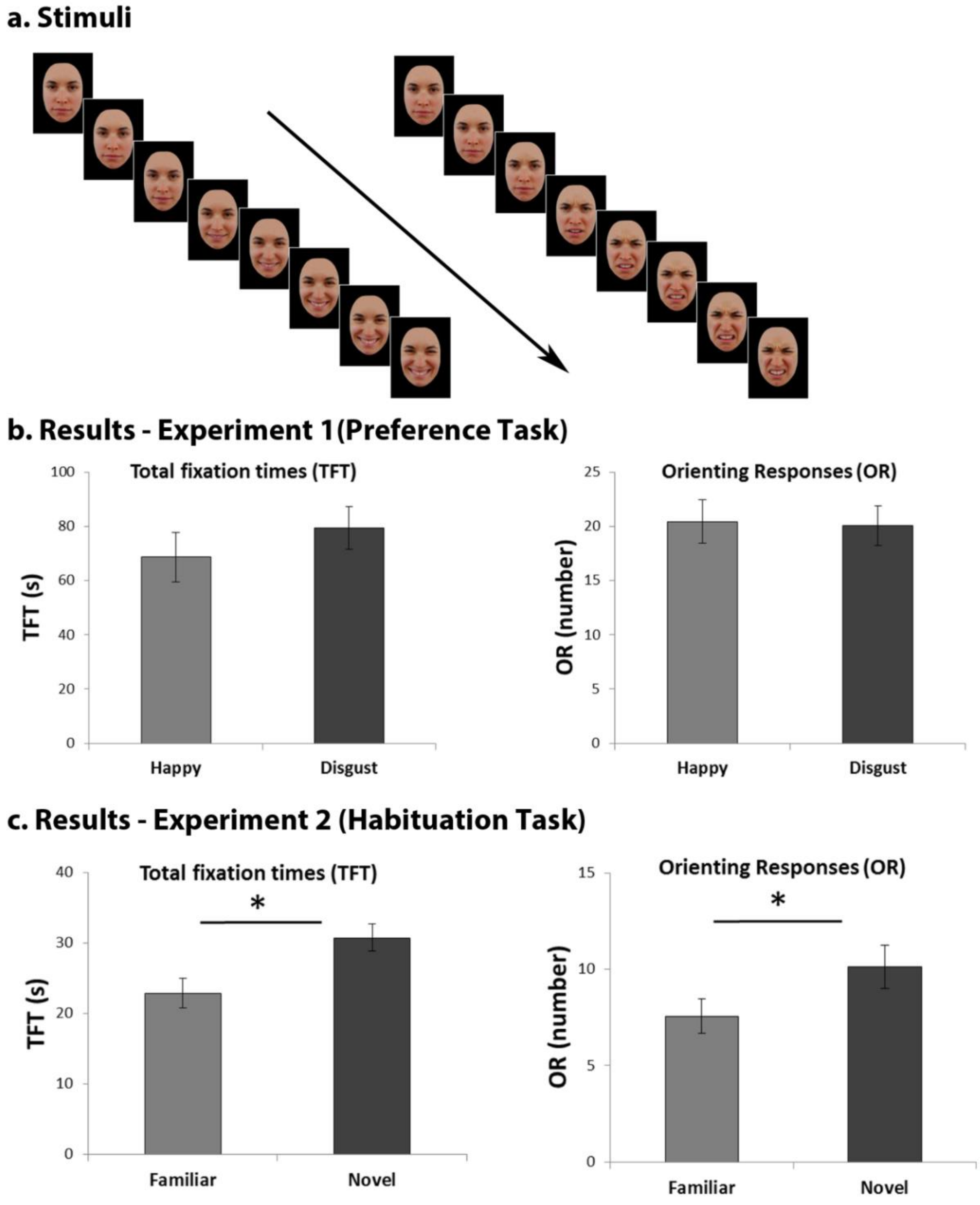


Figure 1. a) Frames composing the videos depicting the expression of disgust (left) and happiness (right) of one of the two face identities shown. Frames are presented in succession in the direction indicated by the arrow. **b)** Newborns' total looking times (left) and orienting responses (right) towards the expression of happiness and disgust in Experiment 1 (Preference task). **c)** Newborns' total looking times (left) and orienting responses (right) towards the novel and the familiar expression in the test phase of Experiment 2 (Habituation task). Error bars refer to the standard errors of the mean. *=p<.05

4.4 Discussion

In the present study we have tackled the intriguing question of whether in the first stages of postnatal life newborns are able to discriminate between dynamic displays of emotional facial expressions. This is the first study addressing the issue of newborns' ability to visually discriminate emotional expressions using ecologically valid dynamic face stimuli. We have found that when 2-day-old newborns were presented simultaneously with moving faces expressing happiness and disgust, they didn't manifest a spontaneous visual preference toward one of the two stimuli (Experiment 1). We have also demonstrated that the absence of a preference response in Experiment 1 wasn't explained by a general inability at birth to discriminate between the two dynamic emotional expressions. Newborns were actually able to differentiate between a happy and a disgusted moving face as testified by an overall preference in looking times and orienting responses towards the novel facial expression in the visual habituation task (Experiment 2).

Young infants typically prefer to look at happy faces at least until they reach the age in which they start to locomote (i.e. 7 months of age) and this interest for happy faces has been interpreted as the result of infants' early experience with smiling faces (Hoehl, 2013). Preferences for happy faces were also found at birth, at least in some limited circumstances (Farroni et al., 2007; Rigato et al., 2011). However, it is possible that, in these studies (Farroni et al., 2007; Rigato et al., 2011), newborns were responding to a very salient feature like the toothy smile.

Here we have shown that newborns didn't manifest any preference when highly salient facial features were controlled and when emotions were presented in a dynamic fashion. How can we explain newborns' lack of preference found in Experiment 1? One possible explanation can be related to the fact that dynamic displays represent extremely salient stimuli for newborns and the effect of kinematics on their discriminative abilities could be beneficial but also distracting. For example, recent findings (Guellaï, Coulon, & Streri, 2011) have found that the dynamics of talking faces were so attractive for newborn infants to the point of interfering with their ability to process properly the differences between face identities. Accordingly, emotional facial dynamics in the present study might have been too engaging to reveal a visual preference response towards one of the two emotional expressions. A second alternative explanation of the null result in Experiment 1 is that newborns were able to distinguish between the two stimuli during the preference task but they simply did not prefer a dynamic facial expression over the other. Both emotional dynamics might have attracted equally their attention.

Experiment 2 has confirmed that newborns were actually able to discriminate between two different dynamic faces expressing happiness and disgust. This indicates that 2-day-old newborns are already endowed with a mechanism that allow them to detect the changes that take place during the unfolding of an emotional facial expression. This is in line with several studies showing that the visual system is already sensitive to motion-based information at birth (Bulf et al., 2010; Longhi et al., 2015).

Our results imply that newborns' discriminative abilities are sophisticated enough to allow them to distinguish between two different, complex, and perceptually not very distant dynamic facial expressions. When Farroni and colleagues (2007) compared static facial expressions which differences weren't very marked (i.e. fearful vs. neutral faces), newborn failed to show a preference and even to discriminate between them. It's reasonable to think that small differences between static facial expressions might become undetectable by newborns' very immature visual system and that they might benefit by the presence of dynamics. Future studies should address this issue by directly comparing the effects of dynamic and static stimuli on newborns' discriminative abilities.

During the first year of life, significant postnatal experiences are required for the development of infants' ability to use the information coming from their caregiver's facial expression to understand, for instance, what is safe or harmful in their environment, and act accordingly (i.e., social referencing). Our results show that, despite the relatively little visual experience that newborns possess at birth, they are already able to distinguish between different dynamic facial expressions and this early ability might provide a fertile ground for the development of later and more complex cognitive skills. However, visual experiences accumulated in the first days of life may not be enough to boost newborns' preference towards one of the two dynamic emotional expressions. Significant postnatal experiences may be required to refine and enrich infants' understanding of others' emotional expressions.

Investigating how newborns process dynamic emotional facial expressions is fundamental to fully understand how they actually see and process the social world around them, which is constantly in motion. Our results show that at birth newborns are able to distinguish between different emotional facial dynamics and, thus, this study represents a crucial step towards a deeper comprehension of newborns' sensitivity to human emotional facial behavior.

Chapter 5. Binding emotion and action at 6 months of age: an EEG study.

5.1 Introduction

In Chapter 4, we have shown that as soon as birth, newborns are sensitive to dynamic, ecologically valid, facial expressions and we have highlighted the importance of studying infants processing of emotions as they naturally occur during our daily interactions. But we also know that in real life facial expressions mostly refer to someone or something in the environment. Facial expressions not only reveal the internal states of an interactor but, more importantly, tell us also about the intention behind his actions. If we observe a baby grasping a toy with an angry expression, we immediately predict that he is going to throw it away or hit it on the table. Differently, if the baby grasps it with a happy face, we know that he is willing to play with it. This example underlines the importance of linking the information coming from emotional expressions to intentions in our social interactions. The tight link between emotions and actions has been extensively investigated in adults and recent lines of research suggest that emotion and action processing influence each other in a bidirectional way (Enticott et al., 2012; Ferri et al., 2013; Mazzola et al., 2013). Indeed, emotion processing can both modulate (Hajcak, Molnar, George, Bolger, Koola, & Nahas, 2007; Enticott et al., 2012; Ferri et al., 2013; Mazzola et al., 2013) and, in turn, be modulated by motor activation (Enticott et al. 2008; Oberman et al., 2007). The link between emotional context and action processing is considered extremely adaptive for survival. For example, when we encounter an angry person we rapidly activate our attentional, motivational and our motor system as well in order to (re)act immediately and efficiently to potentially harmful situations (Gross, 1998; Bradley, Codispoti, Cuthbert, & Lang, 2001; Yuan, Lu, Yang, & Li, 2007; Mazzola et al., 2013).

A considerable amount of studies in infancy have investigated the neural underpinnings of emotion and action processing but separately, mainly employing Electroencephalography (EEG). To date it is clear that sensitivity to others' emotional expressions emerges very early in life and is indexed by three key event-related EEG components: the 290, the P400, and the Negative Central (Nc). The infant posterior P400 has been found to be larger in response to fearful compared to neutral and happy facial expressions (Leppanen, Moulson, Vogel-Farley, & Nelson, 2007) and to fearful compared to angry faces (Kobiella, Grossmann, Reid, & Striano, 2008). The posterior 290 was greater to angry compared to fearful expressions (Kobiella et al., 2008). The amplitude of the frontocentral midlatency Nc tends to be greater in response to fearful or angry faces compared to happy or neutral ones (DeHaan, Belsky, Reid, Volein, & Johnson, 2004; Leppanen, Moulson, Vogel-Farley, & Nelson, 2007; Peltola et al., 2009; Grossman, Striano, & Friederici, 2007; Hoehl & Striano, 2008). Interestingly, a more recent line of research on infants' sensitivity to emotions portrayed by bodies has shown that 8-month-olds but not 4-month-olds respond differently to dynamic point-light-displays (PLDs) emotional body expressions of happiness and fear as testified by a late (700-1100 ms) positive emotion-sensitive activation over temporal and parietal electrodes in the right hemisphere (Missana, Atkinson, & Grossmann, 2015). Further, 8-month-olds demonstrated to be sensitive to happy and fearful body expressions also when they were presented with static photographs (Missana, Rajhans, Atkinson, & Grossmann, 2014). These authors have found two emotion-sensitive ERP components namely the occipital 290 and a later fronto-central Nc. Both components were greater in amplitude for fearful compared to happy body expressions.

With respect to action processing in infancy, by 6 months of age infants start to predict, as measured by eye gaze, the goal of simple and familiar actions performed by others (Woodward, 1998; Hunnius & Bakkering, 2010, Kanagogi & Itakura, 2011; Kochukhova & Gredeback, 2010) and later, at 12 months, they start predicting more complex actions like moving an object in a container (Cannon, Woodward, Gredeback, von Hosten, & Turek, 2012; Falck-Ytter, Gredeback, & von Hosten, 2006). Neurophysiological evidence has also shown higher anticipatory ERP activations in response to goal-directed grasping actions compared to non-goal directed actions at 6 months of age (Nyström, 2008). Interestingly, by the end of the first year of life, infants can also benefit from the presence of contextual information to better understand others' intentions. For example 18-month-olds anticipate a joint or individual action on the basis of the level of engagement (socially engaged or not) with the actor who's performing the action (Fawcett et al., 2013). Further, 12-month-olds can use social cues like gaze and emotional expression to anticipate which object an actor is more likely to grasp (Phillips, Wellman, & Spelke, 2002). Thus, the context provides fundamental information to predict others' behavior. An attempt to link emotion and action was made by Hepach and colleagues (2013) who measured pupil dilation in response to happy and angry actors performing an action congruent or incongruent with their facial expression. Fourteen-month olds, but not 10-month-olds, showed greater pupil dilation for incongruent face/action pairings and this was taken as evidence of an early understanding of others' action on the basis of their emotional expressions. Thus at the end of the first year of life infants seem to be able to bind a type of action (i.e. patting or thumping a toy) to a specific facial expression and, thus, to the interactor internal state.

In sum, the first year of life is an amazing time of growth where babies learn to understand others’ emotions and to predict others’ actions. However, to date, no study has explored how these two big domains of cognition are integrated in infancy thus leaving open the question whether the emotional context can modulate how infants process an observed action. Since the very first months of life infants observe in many occasions actions embedded in an emotional context. Given that the integration of affective and action-related information is highly adaptive for humans, we expect that a rudimental mechanism that links emotion to action might be already present in the first year of life.

Here we hypothesize that the presence of an emotional context (both positive and negative) during action observation could lead to a facilitation in processing an action compared to a neutral context, indexed by an increase of efficiency at early stages of visual processing. Differently from the large amount of studies that investigated infants’ processing of facial expressions and actions in isolation, we will present 6-month-old infants with a realistic visual scene similar to what they usually see in real life: a woman who grasps an object with different facial expression (happiness, anger and neutral). Crucially, the kinematics of the actions in the three different emotional contexts are kept constant.

5.2 Experiment

5.2.1 Methods

Participants

The final sample included 14 healthy full-term 6-month-olds (11 females, mean age = 6 months 7 days, range = 185–199 days). An additional 26 infants were also tested, but were not included in the final sample due to fussiness (N=13) and no completion of an adequate number of trials to be considered for data analysis (i.e., 8 trials per condition; N = 8), or eye and body movements that resulted in excessive recording artifacts (N = 5). The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki (BMJ 1991; 302: 1194) and approved by the Ethical Committee of the University of Milano-Bicocca. Parents gave their written informed consent.

Stimuli, apparatus, and procedure

Testing took place in an electrically shielded and dimly illuminated room. Infants were seated on a parent's lap approximately 60 cm from a 24-inch screen used for stimulus presentation. A video camera installed above the screen recorded a video of the infant, which was synchronized with stimulus presentation for off-line coding of eye and body movements occurring during each trial. Infants were presented with 3 colored movies showing an actress (face, torso and an arm) performing an action in three different emotional contexts: 1) Happiness: the actress smiles and then grasps a ball 2) Anger: the actress displays anger and then grasps a ball 3) Neutral: the actress moves her mouth and then grasps a ball. The total duration of each movie was of 2000 ms; in the first 400 ms the actress dynamically expressed one of the three emotional facial expressions. Then, after further 400 ms, the grasping action started and the movement lasted approximately for 800 ms. Then the video remains still for further 400 ms (Figure 1). The kinematics of the actions in the three experimental conditions were identical. To obtain such identity, we applied the Blue Screen technique in order to superimpose on the same trunk different dynamic facial expressions. The dimension of the face from a distance of 60 cm from the screen was 5.75° of visual angle height and 9.5° in width. There was no restriction in number of trials shown, i.e., they could be played indefinitely. The three different conditions were presented in a pseudorandomized order by E-prime 2.0 software. Each trial started with a screensaver image (presented in the position where later will appear the face). The screensaver lasted 3000 ms and was dynamic for the first 2500 ms and then static for the remaining 500 ms. Then, one of the three movies was presented to the infant.

Figure 1. Example frames taken from the video clips in the Neutral, Happy and Angry condition.

ERP recording and analysis

Continuous scalp EEG was recorded from a 128-channel HydroCel Geodesic Sensor Net (Electrical Geodesic, Eugene, OR) that was connected to a NetAmps 300 amplifier (Electrical Geodesic, Eugene, OR) and referenced on-line to a single vertex electrode (Cz). Channel impedance was kept at or below 100 KΩ and signals were sampled at 500 Hz. EEG data were pre-processed off-line using NetStation 4.5 (Electrical Geodesic, Eugene, OR). The EEG signal was segmented to 1000 ms post-emotion and to 500 ms post-action onset, with a baseline period beginning 100 ms prior to the onsets. Data segments were filtered using a 1–30 Hz band-pass filter and baseline corrected using mean voltage during the 100 ms pre-stimulus period. Automated artifact detection was applied to the segmented data to detect individual sensors that showed > 200 uV voltage changes within the segment period. The entire trial was excluded if more than 18 sensors (15%) overall had been rejected. Data were then inspected manually to mark as bad segments containing drift and eye blinks. Of the remaining trials, individual channels containing artifacts were replaced using spherical spline interpolation. For each participant, average waveforms were generated within each experimental condition only if at least 8 artifact-free trials were overall available per condition collapsed across the target-side. Averaged data were then re-referenced to the average reference.

Inspection of the grand-averaged waveforms revealed that facial expressions elicited three expected deflections, namely the occipital N290, P400 and the frontocentral Nc. (Figure 2). Thereby, ERP responses to the emotion were analyzed over eight posterior electrodes (71-70(O1)-66-65 over the left and 76-83(O2)-84-90 over the right hemisphere) for the N290 and P400, and over twelve frontocentral electrodes (12-20-29-30-35-36 over the left and 104-105-110-111-112-118 over the right hemisphere) for the Nc (Figure 2). The grand-averaged waveforms elicited in response to the action, revealed an early negative ERP component at centroparietal sites within the 100-300 time window from action onset. Differences in ERP responses to the action were analyzed over six central electrodes (36-30-29 over the left and 104-105-11 over the right hemisphere) and over ten parietocentral electrodes (53-54-61-42-37 over the left and 78-79-86-87-93 over the right hemisphere) (Figure 3).

5.2. Results

Analysis of ERPs at the onset of the emotion.

For each ERP component (N290, P400 and Nc), a repeated-measures Analysis of Variance (ANOVA) was separately carried out on with Emotional context (Happy, Angry, Neutral) and Hemisphere (Right, Left) as within-subjects factor.

N290 (200-370 ms).

The ANOVA did not reveal any significant main effect or interaction for peak amplitude (all ps > .3). For the Latency of N290, the ANOVA revealed a main effect of Emotional context, F(2,26)= 8.71, p =.001, η_p^2 =.57, and of Hemisphere, F(1,13)= 7.73, p =.016, $\eta_p{}^2$ = .37, and a significant Emotional context by Hemisphere interaction, F(2,26)= 9.1, p < .001, $\eta_p{}^2$ = .86. Turkey’s HSD post-hoc test indicated that the N290 peaked earlier to the Angry face in the left hemisphere (227 ms) than to the Angry face in the right hemisphere (289 ms), p< .001. The latencies of the N290 were also shorter for the angry face in the left hemisphere compared to the Happy face in the left (283 ms), p< .001, and right hemisphere (278 ms), p <

.001 and also compared to the Neutral Emotion in the left (279 ms), p< .001, and right (273 ms) hemisphere, p < .001. No other comparison was significant (all ps > 08) (Figure 2).

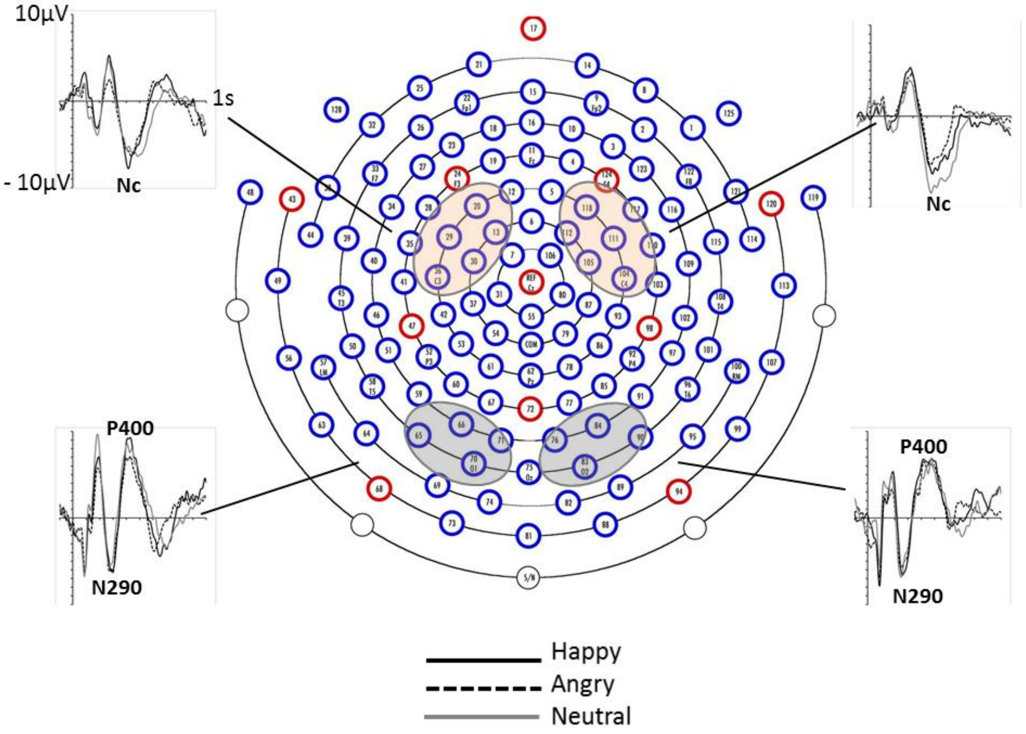


Figure 2. Average waveforms elicited at the onset of emotion over occipital and frontal clusters of the right and left hemisphere.

P400 (350-550 ms) and Nc (350-550 ms).

The ANOVA did not reveal any significant main effect or interaction for mean amplitude (all ps > .2).

Analysis of ERPs at the onset of the action.

A repeated-measures Analysis of Variance (ANOVA) was carried out on mean amplitudes with Site (Parietal, Central), Emotional context (Happy, Angry, Neutral) and Hemisphere (Right, Left) as within-subjects factor.

Centro-Parietal Negativity (100-300 ms).

The ANOVA for mean amplitude revealed a main effect of Hemisphere, F(1,13)= 10,8, p =.006, η_p^2 =.45. Overall, the amplitude of the negative component was larger in the left (-2.2 µV) than in the right (-0.9 µV) hemisphere. Further, there was a significant Site by Emotional context interaction, F(2,26)= 10.85, p < .021, η_p^2 =.45, that was further explored with Turkey's HSD post-hoc. The mean amplitude of the negative component was greater in parietal sites in response to Angry (-2.1 µV) than to Neutral Emotion (0.4 µV), p < .044, and to Happy (-2.4 µV) compared to Neutral Emotion (0.4 µV), p < .016. No other comparison was significant (all ps > .07) (Figure 3).

5.3

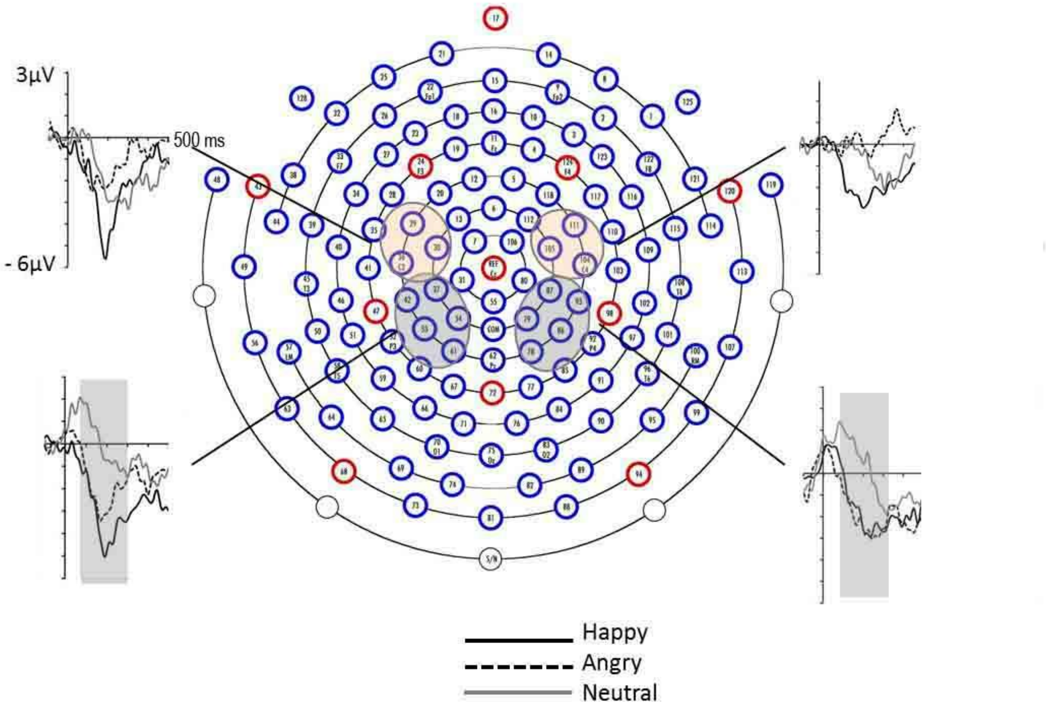


Figure 3. Average waveforms elicited at the onset of action over parietal and central clusters of the right and left hemisphere.

5.3 Discussion

In this study we examined ERPs modulations in 6-month-old infants in response to an object-directed grasping action performed in three different emotional contexts (Happy, Angry and Neutral). Interestingly, we did not find modulations in amplitude of emotion-sensitive ERP components (i.e., N290, P400 and Nc) in response to the emotional facial expressions presented in our stimuli. All these ERP components were clearly visible in the waveforms elicited by the facial expressions but the only significant difference between conditions was found on the latency of the N290, which was shorter in response to the angry emotion compared to the happy and neutral expressions. Our results are discordant with a previous finding that has found in 7-months-olds greater amplitudes of the Nc to the happy compared to the angry expression (Grossman et al., 2007). However, it has to be noted that our study differentiates from previous ones for two main aspects: 1) we presented dynamic facial expressions and not static faces; 2) in our study facial expressions were embedded in a context and weren't in isolation. Thus, discrepancies between studies can be explained by the different nature of the stimuli used. Possibly, 6-month-old infants were able to distinguish angry faces from happy or neutral expressions, even if this was not evident in EEG responses. This is confirmed by the fact that the three different emotional contexts affect the subsequent processing of goal-directed actions on the cortical level.

In fact our results revealed that the affective information conveyed by facial expressions can modulate ERP visual components at early stages of action visual processing. Specifically, the modulatory effect of emotion on action processing occurred at parietal sites between 100 and 300 ms from action onset, and was indexed by a negative deflection that was greater in amplitude during the observation of an action embedded in an emotional context (both angry and happy) compared to a neutral one. The emotional context might have increased the efficiency of processing of the subsequent goal-directed action. Interestingly, the effect of emotion was strongly evident over parietal electrodes. The parietal lobe is thought to have a crucial role in action understanding (Fogassi et al., 2005). EEG studies with adults (Ortigue et al., 2010; Proverbio et al., 2011; Wheaton et al., 2001) and children (Berchio et al., 2015) have found a parietal N2 response (150–280 ms) (whose neural generators includes regions involved in action processing) which seems to be modulated by the action's purpose (Ortigue et al., 2010; Proverbio et al., 2011). Scalp distribution and latency of the adult N2 seem to have some similarities with the negative deflection found in our study in response to a goal-directed action. Further, the adult N2 parietal activation has been found to be lateralized to the left hemisphere (Ortigue et al., 2010). Similarly, the negative action-related activation found in our study was more visible and greater in amplitude in the left parietocentral hemisphere than in the right one.

Alternatively, the negative activations that we have found in the parietal lobe could be related to overt shifts of attention towards the goal of the action (i.e., the ball). However, happy, angry and neutral facial expressions were equally salient to infants, given that they allocate the same amount of attention to all the three emotions, as shown by equal Nc responses. Second, attention-related ERPs in infants have been found mainly in occipital-temporal sites (Richards et al., 2000; Natale et al., 2016) rather than in parietal electrodes. Thus, it is unlikely that differences in the negative parietal ERP were due exclusively to attentive processes. However, further investigations are needed to identify the neural generators of the differential negative action-related activations found in our study.

To sum up, in the present study we have found that the emotional context can modulate the processing of a goal-directed action at 6 months of life. In everyday life, infants are constantly engaged in dynamic interactions with other people who display emotions together with their action or intention to act. Recent evidence has shown that the ability to link a type of action to the corresponding emotional state emerges at the end of the first year (Hepach et al., 2013) and this capability may need sophisticated social-cognitive skills acquired through experience to develop as, first of all, the ability to recognize emotions and to link this information to the other related events that occur in the environment. Here we have shown that the precursors of the ability to bind affective-related to action-related information might be traced at the neural level earlier in life. Specifically, we have shown that the emotional context (both positive and negative), enhance infants' processing of a human action. An action performed in an emotional context might acquire a certain relevance among other events in the world. For instance, an action performed in a negative context can be potentially threatening, while an action performed in a positive context can prompt people to engage in interactions (i.e., play, explore). Thus, emotion information might be a fundamental clue that allows infants, from very early in life, to respond efficiently to relevant behaviors that occur in their environment.

Conclusion

Humans are social creatures and, from birth, they spend a great deal of time attending to faces. Faces can convey emotional information during social interactions that can help infants to interpret the internal states and the intentions of others. So far, many of the studies that have explored infants' processing of facial expressions have reduced the stimulus situations to static and disembodied faces. In section 2, we aimed at exploring how newborns and infants process facial expressions as they usually take place during their daily social interactions: in motion and embedded in a more complex social and communicative scene.

In Chapter 4, we examined 2-day-old newborns' sensitivity to dynamic facial expressions of happiness and disgust. Our results have shown that newborns, shortly after birth, are able to discriminate between different dynamic emotions. This result supports a handful of studies that have found an early sensitivity to dynamic stimuli at birth by extending this finding also to dynamic facial expressions. Facial expressions are a combination of internal features of the face that change configuration over time. We have shown that newborns, despite the limits of their immature visual system, are already endowed with a mechanism that helps them to detect morphological changes of the face that occur while an interactor expresses emotion. This early discriminative ability could provide a solid basis on which infants can build up more sophisticated social abilities throughout infancy.

During social interactions, infants have to deal with different sources of information. For instance, when they interact with their caregiver, they need to process the information coming from their facial expressions and use it to predict her behavior. In Chapter 5, we explored whether affective-related information could affect infants' processing of a goal-directed action by using EEG. We presented 6-month-olds infants with a scene that resembles more closely what they usually encounter in their daily interactions: a woman that expresses an emotion and then performs an action towards an object in her environment. We have found that the action was processed at the neural level more efficiently when it was embedded in an emotional context than in a neutral one. We took this result as evidence of the presence in the first year of life of a rudimental mechanism that uses affective information to enhance and facilitate the processing of relevant events that occur in infants' environment, providing early precursors for the development of the ability to bind affective information to actions.

In conclusion, in Section 2 we deepened the understanding on how emotional information is processed in infancy. We have shown that at birth newborns are sensitive to dynamic emotional facial expressions and that the emotional context can modulate the processing of an action at 6 months of age. Overall, these main findings show the importance of studying how infants process the world as they actually see it, that is complex and rich of events coming from different sources of information.

General conclusion

The current thesis presents a series of experiments that aim at exploring the understanding of gestures and emotions in infancy.

The three studies described in Section 1 shed light on the developmental origins of the ability to recognize and understand human gestures, like grasping gestures, touching gestures and facial gestures. Specifically, results of the study described in Chapter 1 have shown that the neural basis of action priming can be traced already at 6-months of age. Grasping gestures are powerful means that can trigger shifts in attention of the infant. Importantly, results have shown that, at the neural level, action priming leads to a sensory facilitation in processing targets that appear in the cued location. Crucially, action priming effect occurs only in response to gestures that the infants could perform highlighting the fundamental role that sensorimotor experience plays in the understanding of grasping gestures.

The studies presented in Chapter 2 and 3 focused on the early sensitivity to gestures at birth. The main finding of the study described in Chapter 2 is that as soon as a few days after birth, newborns seem to be sensitive to gestures that comprise a tactile event. In this study newborns have shown to be able to discriminate between touching and non-touching gestures, manifesting a visual preference to the former than to the latter. Moreover, at 3 months of age spontaneous preferential responses change possibly as a result of further experiences accumulated during development. The last chapter of the first section, Chapter 3, has demonstrated that when newborns are presented with A-V information that is part of their sensorimotor repertoire, newborns can successfully associate a facial gesture to the matching sound even in absence of temporal synchrony. Although the two studies presented in Chapter 2 and 3 don’t test directly the role of pre- and postnatal sensorimotor experiences, we support the idea that they might have played a crucial role in the developing of the early sensitivity to human touching and facial gestures at birth.

Section 2 deepened our understanding about emotion processing in infancy. First of all, results of the study described in Chapter 4 have shown that already at birth newborns are able to differentiate between dynamic facial expressions of happiness and disgust, showing to already possess a perceptual system that allows them to process the complex dynamic world that surrounds them. Lastly, in Chapter 5, which is the last chapter of this section and of the present thesis, we have shown that facial expressions represent a source of information that can modulate on the cortical level infants’ processing of an observed action.

The role of sensorimotor experience in social understanding.

Converging evidence is consistent with the idea that motor experience has a unique and powerful contribution to action priming (Daum et al, 2011; Bakker et al., 2014; Melinder et al., 2015) and action understanding (e.g. Woodward, 2009, Stapel et al., 2016). In line with these findings, the study presented in Chapter 1, has shown that only a gesture that infants are able to perform can facilitate shifts of attention in their environment according to the direction of the hand grasp. Thus, we have shown that the infant brain responds selectively only to gestures that are part of their sensorimotor repertoire. One’s own sensorimotor experience with grasping gestures may have helped infants to discover which movements are possible or impossible to perform. Hence, the acquired sensorimotor experience might allow them to detect quickly whether an observed movement violates or not the biomechanical constrains of a human hand by replicating those movements onto their own motor system. A very intriguing finding is that even sensorimotor experiences acquired during prenatal life seem to play a key role in newborns' early perceptual abilities. The importance of fetal experiences has been for a long time ignored, due to a traditional view that wrongly considered fetal movements as unintentional and accidental (Ennouri & Boch,1996; van der Meer, van der Weel, & Lee, 1995). During prenatal life fetuses perform touching gestures directed toward their external world and their own body (Sparling et al., 1999). This means that they have the possibility to experience the sensory counterpart of their movements toward the external world as well as tactile sensations as result of touching their own body. The study presented in Chapter 2 suggests that somatosensory-motor associations accumulated in the womb and in the first hours after birth may allow newborns to link executed with observed touching gestures. Moreover, infants' spontaneous preferential responses to touching or no-touching gestures in the first 3 months of life seem to be modulated by somatosensory-motor, visual and affective/communicative experiences. Finally, the fact that newborns' early sensitivity to touch is specific to gestures with a higher social component, here a human body-to-body contact, suggests that the social nature of the gesture plays a key role in triggering newborns' visual attention from the very early stages of postnatal life. In this regard, it is also noteworthy the evidence in human adults indicating that the human tactile mirror system, in particular some of its cortical regions, may be best tuned to represent human body-part interactions, than the contact between inanimate objects, or between human body-parts and objects (Blakemore, Bristow, Bird, Frith, & Ward, 2005; Rossetti, Miniussi, Maravita, & Bolognini, 2012; Keysers, Kaas, & Gazzola, 2010), and it differentially responds to the sight of an intentional touch as compared to an accidental touch (Ebisch et al., 2008).

The hypothesis of the presence of early action-observation couplings at birth is corroborated by results in Chapter 3. In this study we have demonstrated that newborns can improve their multisensory capabilities in the presence of audio-visual information that is part of their sensorimotor experience. Interestingly, multimodal experiences accumulated during fetal life and in the first days after birth with yawns and hiccups derive primarily from sensorimotor stimulations. This study is in line with a growing line of evidence suggests that audiovisual speech information could be mapped using sensorimotor information coming from orofacial movements (see Guellai, Streri & Yeung, 2014 for a review). This view is supported by evidence showing that, since birth, infants produce more imitative responses when they are presented with congruent, compared to incongruent, audiovisual speech information from face and voice (i.e., mouth openings and corresponding produced vowels) (Coulon et al., 2013; Lagerstee, 1990; Kuhl & Metzoff, 1996; Patterson & Werker, 1999). Recently, a more direct evidence of the influence of sensorimotor information on intersensory speech perception was provided by Yeung and Werker (2013). In this study, 4, 5-month-olds were able to make audiovisual matching when presented with side-by-side talking faces producing a vowel ([i] and [u]). Crucially, this preference for sound-matching gestures disappeared when, during audiovisual presentation, infants were producing lip movements similar to the ones needed to produce the heard vowel (i.e., chewing a teeth ring while presented with audiovisual [i] or sucking on a pacifier while presented with audiovisual [u]). Differently, they could still match audiovisual information when they were producing mismatching-lip movements.

In addition, evidence has shown that during prenatal life fetuses also display facial muscle configurations that can be associated to the expression of distress and positive states (Reissland et al., 2011). Thus, these prenatal sensorimotor experiences with facial movements could have boosted newborns' early discriminative abilities of dynamic facial expressions (Chapter 4). However, this doesn't mean that newborns shared the same emotion of the observed actress, but rather a common motor representation of the facial movement. Through the first year of life infants' understanding of emotions is increasingly refined. In Chapter 5, we have shown that at 6 months of life infants' are sensitive to observed actions embedded in different emotional contexts. Studies in adults suggest the presence of a direct neural pathway through which emotional contexts may drive the neural motor system (Mazzola et al. 2013). Our finding shows that cortical connections between emotion-related and action-related areas might be already in place, although in an immature form, at 6 months of age. In fact, infants showed at the neural level to process differently the goal of an action performed in an emotional compared to a neutral context. Investigating how infants integrate and combine information coming from different domains of social cognition is essential to understand how they process more complex situations similar to the ones that they daily encounter during their life. During interactions we have to process and integrate information coming from different sources, like facial expressions, eye gaze, action goals, gestures. Future research might explore the effect of emotional context on mu rhythm desynchronization during action observation as a direct measure of motor cortex modulation.

Toward an embodied view of infants' social understanding.

The present investigation pulled together different but interrelated areas of early social cognition: social attention, sensitivity to touching and facial gestures, perception of emotions and actions. Our findings have shown that these different aspects of social understanding seem to be all shaped by infants' own sensorimotor experiences which allow them to make sense of the social world since the very first stages of life. It is starting to be increasingly recognized in both developmental and adult social cognition the need to consider, besides the mind and the environment, the key role of the body in the construction of the meaning of social information.

For instance, the "Like-Me" framework (Metzoff 2007; 2013) posits that infants start to make sense of the social world through a felt similarity between their own bodily acts and the bodily acts of others. The perception of this similarity depends on a primitive body representation that is influenced by one's own body experiences which are used by the infants to link behaviors of others to behaviors of the self through shared representations (Meltzoff, 2007). Interestingly, studies in adults suggest that we don't only share the action of others, but also others' tactile sensations and emotions (Keysers & Gazzola, 2009). For example, observing others' being touched triggers vicarious activity in the somatosensory cortex (i.e. Keysers et al., 2004; Blakemore et al., 2005; Ebisch et al., 2008), while observing others' emotions like disgust, pain and happiness activates the insula and frontal operculum involved in experiencing similar emotions (Jabby, Swart, & Keysers, 2007; Botvinick et al., 2005). Thus, mirror activity might be considered a mechanism that can be extended to other domains of social cognition beyond action understanding and might not be considered a peculiar property of the motor cortex. Results of the present investigations represent a starting point from which to explore how infants use information coming from their own sensory experience in domains of social cognition other than action understanding, like social attention, emotion, multisensory integration and somatosensory processing. To this end, novel applications of methods for recording infant brain activity can foster our understanding of how we connect self and other through body representations during development.

Future directions.

Increasing attention is being devoted to the role of neural representations of the body in adult social cognition. As discussed in the previous paragraph, there is evidence that one's own perceptual and sensory experience is involved in the processing of others' social and emotional signals (Keysers & Gazzola, 2009). However, developmental research in the field is lacking. Evidence from magnetoencephalography (MEG) (Bosseler, 2013; Kuhl, 2014), functional near-infrared spectroscopy (Filippetti, 2014; Aslin, 2015) and, as well, electroencephalography (EEG) can shed light on how human infants develop the link between the self and the other in the first days and months of life.

A growing body of EEG studies on the mu rhythm have shown infants perceive others' actions on the basis of their own sensorimotor capabilities (Marshall & Meltzoff, 2011; 2014). However, the mu rhythm is not solely an index of motor activation but is also related to somatosensory processing (Arnstein, 2011; Ritter, 2009). The study of the activation of somatosensory cortices might be of great interest to explore of how infants use own body representations to make sense of the social world. In fact, somatosensory cortices are involved not only in the processing of tactile and pain stimulations but, in concert with other structures such as the amygdala and right visual cortices, seem to be fundamental in retrieving socially relevant information from faces, like emotional expressions (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000). Thus, the study of mu rhythm as an index of somatosensory activation would shed light on how infants develop the sense of the self as distinct but, at the same time, similar to others by investigating self-other links in a different domains of social cognition like in the processing of touching gestures and emotion.

References

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., & Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. Journal of Neuroscience, 20(7), 2683-2690.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. Nature Reviews Neuroscience, 4(3), 165-178.

Adolphs, R. (2006). Perception and emotion: How we recognize facial expressions. Current directions in psychological science, 15(5), 222-226.

Aldridge, M. A., Braga, E. S., Walton, G. E., & Bower, T. G. R. (1999). The intermodal representation of speech in newborns. *Developmental Science*, *2*(1), 42-46.

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., & Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS One*, *8*(7), e67916.

Anisfeld, M. (1991). Neonatal imitation. *Developmental review*, *11*(1), 60-97.

Anisfeld, M. (1996). Only tongue protrusion modeling is matched by neonates. *Developmental Review*, *16*(2), 149-161.

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain topography*, 24(2), 149-163.

Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. Annual review of psychology, 66, 349-379.

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., & Gazzola, V. (2011). μ-suppression during action observation and execution correlates with BOLD in dorsal premotor, inferior parietal, and SI cortices. Journal of Neuroscience, 31(40), 14243-14249.

Atkinson, J., Braddick, O., & French, J. (1979). Contrast sensitivity of the human neonate measured by the visual evoked potential. *Investigative ophthalmology & visual science*, 18(2), 210-213.

Bahrick, L. E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. *Advances in child development and behavior*, *30*, 153-189.

Bahrick LE. Body Perception: Intersensory Origins of Self and Other Perception in Newborns. Current Biology. 2013; 23: R1039-R1041.

Bakker, M., Daum, M. M., Handl, A., & Gredebäck, G. (2015). Grasping the difference through experience! - Neurophysiological correlates of action perception during the onset of grasping. Social Cognitive and Affective Neuroscience 10, 769-776. doi: 10.1093/scan/nsu119.

Bardi, L., Di Giorgio, E., Lunghi, M., Troje, N.F., & Simion, F. (2015). Walking direction triggers visuo-spatial orienting in 6-month-old infants and adults: An eye tracking study. Cognition, 141, 112-120. doi: 10.1016/j.cognition.2015.04.014.

Barrera, M. E., & Maurer, D. (1981). The perception of facial expressions by the three-month-old. *Child development*, 203-206.

Blakemore SJ, Bristow D, Bird G, Frith C, Ward J. Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. Brain. 2005; 128: 1571-1583.

Biro S, Leslie AM. Infants' perception of goal-directed actions: development through cue-based bootstrapping. Developmental science. 2007; 10: 379-398.

Berchio, C., Rihs, T. A., Michel, C. M., Brunet, D., Apicella, F., Muratori, F., ... & Umilta, M. A. (2014). Parieto-frontal circuits during observation of hidden and visible motor acts in children. A high-density EEG source imaging study. *Brain topography*, *27*(2), 258-270.

Bertenthal, B. I., Boyer, T. W., & Harding, S. (2014). When do infants begin to follow a point? Developmental Psychology, 50, 2036–2048. doi:10.1037/a0037152.

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*(1), 3-15.

Bornstein, M. H., & Arterberry, M. E. (2003). Recognition, discrimination and categorization of smiling by 5-month-old infants. *Developmental Science*, *6*(5), 585-599.

Bosseler, A. N., Taulu, S., Pihko, E., Mäkelä, J. P., Imada, T., Ahonen, A., & Kuhl, P. K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception.

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. Neuroimage, 25(1), 312-319.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: defensive and appetitive reactions in picture processing. *Emotion*, *1*(3), 276.

Brenna, V., Proietti, V., Montirosso, R., & Turati, C. (2013). Positive, but not negative, facial expressions facilitate 3-month-olds' recognition of an individual face. *International Journal of Behavioral Development*, 37(2), 137-142.

Bulf, H., & Turati, C. (2010). The role of rigid motion in newborns' face recognition. *Visual Cognition*, 18(4), 504-512.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European journal of neuroscience, 13(2), 400-404.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., ... & Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: An fMRI study. Journal of cognitive neuroscience, 16(1), 114-126.

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental science*, *15*(1), 35-42.

Campos, J. J., & Stenberg, C. (1981). Perception, appraisal, and emotion: The onset of social referencing. *Infant social cognition: Empirical and theoretical considerations*, *273*, 314.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proceedings of the national Academy of Sciences, 100(9), 5497-5502.

Castiello U, Becchio C, Zoia S, Nelini C, Sartori L, Blason L, et al. Wired to be social: the ontogeny of human interaction. Plos One. 2010; 5: e13199.

Chen, X., Striano, T., & Rakoczy, H. (2004). Auditory–oral matching behavior in newborns. *Developmental science*, *7*(1), 42-47.

Cohen LB. Attention-getting and attention-holding processes of infant visual preferences. Child Development. 1972; 869-879.

Cohen LB. A two process model of infant visual attention. Merrill-Palmer Quarterly of Behavior and Development. 1973; 157-180.

Coulon, M., Hemimou, C., & Streri, A. (2013). Effects of seeing and hearing vowels on neonatal facial imitation. *Infancy*, *18*(5), 782-796.

Craighero, L., Leo, I., Umiltà, C., & Simion, F. (2011). Newborns' preference for goal directed actions. Cognition, 120, 26–32. doi:10.1016/j.cognition.2011.02.011.

Csibra, G., Kushnerenko, E., & Grossman, T. (2008). Electrophysiological methods in studying infant cognitive development. In C. A. Nelson & M. Luciana (Eds.), Handbook of Developmental Cognitive Neuroscience (pp. 247–262). Cambridge, MA: MIT Press.

Csibra, G., & Gergely, G. Csibra, G., & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. Acta psychologica, 124(1), 60-78. (2007).

Daum, M. M., Ulber, J., & Gredebäck, G. (2013). The development of pointing perception in infancy: Effects of communicative verbal cues on covert shifts of attention. Developmental Psychology, 49, 1898–1908. doi: 10.1037/a0031111.

Daum, M. M., & Gredebäck, G. (2011). The development of grasping comprehension in infancy: covert shifts of attention caused by referential actions. Experimental Brain Research, 208, 297-307. doi: 10.1007/s00221.

De Vries, J. I. P., & Fong, B. F. (2006). Normal fetal motility: an overview. *Ultrasound in obstetrics & gynecology*, *27*(6), 701-711.

de Haan, M., & Nelson, C. A. (1998). Discrimination and categorization of facial expressions of emotion during infancy. *Perceptual development: Visual auditory, and speech perception in infancy*, 247-286.

de Haan, M., Johnson, M. H., & Halit, J. (2003). Development of face-sensitive event- related potentials during infancy: A review. International Journal of Psychophysiology, 51, 45–58. doi:10.1016/S0167-8760(03)00152-1.

De Haan, M., Belsky, J., Reid, V., Volein, A., & Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *Journal of Child Psychology and Psychiatry*, *45*(7), 1209-1218.

Ebisch SJ, Perrucci MG, Ferretti A, Del Gratta C, Romani GL, Gallese V. The sense of touch: embodied simulation in a visuotactile mirroring mechanism for observed animate or inanimate touch. Journal of cognitive neuroscience. 2008; 20: 1611-1623.

Eimer, M. (1996). ERP modulations indicate the selective processing of visual stimuli as a result of transient and sustained spatial attention. *Psychophysiology*, *33*(1), 13-21.

Eimer, M. (1998). Mechanisms of visuospatial attention: Evidence from event-related brain potentials. *Visual Cognition*, *5*(1-2), 257-286.

Enticott, P. G., Johnston, P. J., Herring, S. E., Hoy, K. E., & Fitzgerald, P. B. (2008). Mirror neuron activation is associated with facial emotion processing. *Neuropsychologia*, *46*(11), 2851-2854.

Enticott, P. G., Harrison, B. A., Arnold, S. L., Nibaldi, K., Segrave, R. A., Fitzgibbon, B. M., ... & Fitzgerald, P. B. (2012). Emotional valence modulates putative mirror neuron activity. *Neuroscience letters*, *508*(1), 56-59.

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature neuroscience*, *9*(7), 878-879.

Fantz, R. L. (1961). The origin of form perception. *Scientific American*.

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, 4(1), 2-13.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. Proceedings of the National Academy of Sciences, 99, 9602–9605. doi: 10.1073/pnas.152159999.

Farroni, T., Mansfield, E.M., Lai, C., & Johnson, M.H. (2003). Motion and mutual gaze in directing infants’ spatial attention. Journal of Experimental Child Psychology, 85, 199–212. doi:10.1016/S0022-0965(03)00022-5.

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants’ use of gaze direction to cue attention: The importance of perceived motion. Visual Cognition, 7, 705–718. doi:10.1080/13506280050144399.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. Infancy, 5(1), 39-60.

Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental science*, *16*(6), 841-849.

Ferri, F., Ebisch, S. J., Costantini, M., Salone, A., Arciero, G., Mazzola, V., ... & Gallese, V. (2013). Binding action and emotion in social understanding. *PloS one*, *8*(1), e54091.

Field, T. M., Woodson, R., Greenberg, R., & Cohen, D. (1982). Discrimination and imitation of facial expressions by neonates. *Science,* 218, 179-181.

Field, T. M., Cohen, D., Garcia, R., & Greenberg, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior and development*, *7*(1), 19-25.

Field T. Touch. MIT press; 2003.

Filippetti ML, Johnson MH, Lloyd-Fox S, Dragovic D, Farroni T. Body perception in newborns. Current Biology. 2013; 23: 2413-2416.

Filippetti, M. L., Orioli, G., Johnson, M. H., & Farroni, T. (2015). Newborn body perception: sensitivity to spatial congruency. *Infancy*, *20*(4), 455-465.

Filippetti, M. L., Lloyd-Fox, S., Longo, M. R., Farroni, T., & Johnson, M. H. (2014). Neural mechanisms of body awareness in infants. Cerebral Cortex, bhu261.

Finan, D. S., & Barlow, S. M. (1998). Intrinsic dynamics and mechanosensory modulation of non-nutritive sucking in human infants. *Early human development*, 52(2), 181-197.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, *308*(5722), 662-667.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. Current opinion in neurobiology, 11(2), 219-224.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. Neuroimage, 35(4), 1674-1684.

Geangu, E., Senna, I., Croci, E., & Turati, C. (2015). The effect of biomechanical properties of motion on infants' perception of goal-directed grasping actions. Journal of Experimental Child Psychology, 129, 55-67. doi:10.1016/j.jecp.2014.08.005.

Gredebäck, G., & Daum, M. M. (2015).The microstructure of action perception in infancy: decomposing the temporal structure of social information processing. Child Development Perspectives, 9, 79–83. doi: 10.1111/cdep.12109.

Gredebäck, G., & Kochukhova, O. (2010). Goal anticipation during action observation is influenced by synonymous action capabilities, a puzzling developmental study. Experimental Brain Research, 202, 493–497. doi: 10.1007/s00221-009-2138-1.

Gredebäck, G., Melinder, A. M. D., & Daum, M. M. (2010). The neural basis and development of pointing comprehension. Social Neuroscience, 5, 441–450. doi: 10.1080/17470910903523327.

Gredebäck, G., Theuring, C., Hauf, P., & Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. Infancy, 13, 533–543. doi: 10.1080/15250000802329529.

Green, J. R., & Wilson, E. M. (2006). Spontaneous facial motility in infancy: A 3D kinematic analysis. *Developmental psychobiology*, 48(1), 16.

Gross, J. J. (1998). The emerging field of emotion regulation: an integrative review. *Review of general psychology*, *2*(3), 271.

Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: A neurobehavioral study. *Brain and cognition*, *64*(1), 30-41.

Guellaï, B., Coulon, M., & Streri, A. (2011). The role of motion and speech in face recognition at birth. *Visual Cognition*, 19(9), 1212-1233.

Guellaï, B., Streri, A., & Yeung, H. H. (2014). The development of sensorimotor influences in the audiovisual speech domain: some critical questions. *Frontiers in psychology*, *5*.

Guellaï, B., Streri, A., Chopin, A., Rider, D., & Kitamura, C. (2016). Newborns' Sensitivity to the Visual Aspects of Infant-Directed Speech: Evidence From Point-Line Displays of Talking Faces.

Hajcak, G., Molnar, C., George, M. S., Bolger, K., Koola, J., & Nahas, Z. (2007). Emotion facilitates action: a transcranial magnetic stimulation study of motor cortex excitability during picture viewing. *Psychophysiology*, *44*(1), 91-97.

Hari, R. (2006). Action–perception connection and the cortical mu rhythm. Progress in brain research, 159, 253-260.

Heyes, C. (2010). Where do mirror neurons come from?. Neuroscience & Biobehavioral Reviews, 34(4), 575-583.

Heinze, H.J., Mangun, G.R., Burchert, W., Hinrichs, H., Scholz, M., Münte, T.F., Gös, A., Scherg, M., Johannes, S., Hundeshagen, H., Gazzaniga, M.S., & Hillyard, S.A. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. Nature, 372, 543–546. doi:10.1038/372543a0.

Hepach, R., & Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *Journal of experimental child psychology*, *115*(1), 16-29.

Hertenstein MJ. Touch: Its communicative functions in infancy. Human Development. 2002; 45: 70-94.

Hillyard, S. A., Luck, S. J., & Mangun, G. R. (1994). The cuing of attention to visual field locations: Analysis with ERP recordings. In *Cognitive electrophysiology* (pp. 1-25). Birkhäuser Boston.

Hillyard, S.A., & Anllo-Vento, L. (1998) Event-related brain potentials in the study of visual selective attention. Proceedings of the National Academy of Sciences, 95, 781–787. Retrieved from http://dx.doi.org/10.1073/pnas.95.3.781.

Hoehl, S., & Striano, T. (2008). Neural Processing of Eye Gaze and Threat-Related Emotional Facial Expressions in Infancy. *Child development*, *79*(6), 1752-1760.

Hoehl, S., Wiese, L., & Striano, T. (2008). Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS One*, *3*(6), e2389

Hoehl, S., & Striano, T. (2010). The development of emotional face and eye gaze processing. *Developmental science*, *13*(6), 813-825.

Hoehl, S., & Pauen, S. (2011). Do surprised faces affect infants' attention toward novel objects?. *NeuroReport*, *22*(17), 906-910.

Hoehl, S. (2013). Emotion processing in infancy. In K.H. Lagattuta (Ed.), Children and Emotion. New Insights into Developmental Affective Sciences. Contributions to Human Development. Basel, Karger, 26, 1-12.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. Psychological Science, 9, 131–134. doi:10.1111/14679280.00024.

Hunnius, S., & Geuze, R. H. (2004). Developmental changes in visual scanning of dynamic faces and abstract stimuli in infants: A longitudinal study. *Infancy*, 6(2), 231-255.

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental psychology*, *46*(2), 446.

Isomura, T., & Nakano, T. (2016, December). Automatic facial mimicry in response to dynamic emotional stimuli in five-month-old infants. In Proc. R. Soc. B (Vol. 283, No. 1844, p. 20161948). The Royal Society.

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition*. Blackwell.

Johnson, M. H. (2001). Functional brain development in humans. Nature Reviews Neuroscience, 2(7), 475-483.

Johnson, M. H., Grossmann, T., & Kadosh, K. C. (2009). Mapping functional brain development: Building a social brain through interactive specialization. Developmental psychology, 45(1), 151.

Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: modifications based on two decades of data from infants and adults. Neuroscience & Biobehavioral Reviews, 50, 169-179.

Kahana-Kalman, R., & Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child development*, *72*(2), 352-369.

Kaitz, M., Meschulach-Sarfaty, O., Auerbach, J., & Eidelman, A. (1988). A reexamination of newborns' ability to imitate facial expressions. *Developmental Psychology*, 24(1), 3.

Kätsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, 66(4), 233-242.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nature communications*, *2*, 341.

Kestenbaum, R., & Nelson, C. A. (1990). The recognition and categorization of upright and inverted emotional expressions by 7-month-old infants. *Infant Behavior and Development*, 13(4), 497-511.

Keysers, C., & Gazzola, V. (2009). Expanding the mirror: vicarious activity for actions, emotions, and sensations. Current opinion in neurobiology, 19(6), 666-671.

Keysers, C., & Gazzola, V. (2010). Social neuroscience: mirror neurons recorded in humans. Current biology, 20(8), R353-R354.

Keysers C, Wicker B, Gazzola V, Anton J L, Fogassi L, Gallese V. A touching sight: SII/PV activation during the observation and experience of touch. Neuron. 2004; 42: 335-346.

Keysers C, Kaas JH, Gazzola V. Somatosensation in social perception. Nature Reviews Neuroscience. 2010; 11: 417-428.

Kobiella, A., Grossmann, T., Reid, V. M., & Striano, T. (2008). The discrimination of angry and fearful facial expressions in 7-month-old infants: An event-related potential study. *Cognition and Emotion*, *22*(1), 134-146.

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115-1125.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: a review. *Emotion Review*, 5(1), 41-46.

Kuehn E, Mueller K, Turner R, Schütz-Bosbach S. The functional architecture of S1 during touch observation described with 7 T fMRI. Brain Structure and Function. 2014; 219: 119-140.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant behavior and development*, *7*(3), 361-381.

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. American Association for the Advancement of Science.

Kuhl, P. K., & Meltzoff, A. N. (1988). Speech as an intermodal object of perception. In *Perceptual development in infancy: The Minnesota symposia on child psychology* (Vol. 20, pp. 235-266).

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, *100*(4), 2425-2438.

Kuhl, P. K., Ramírez, R. R., Bosseler, A., Lin, J. F. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. Proceedings of the National Academy of Sciences, 111(31), 11238-11245.

Kurjak A, Stanojević M, Andonotopo W, Scazzocchio-Duenas E, Azumendi G, Carrera JM. Fetal behavior assessed in all three trimesters of normal pregnancy by four-dimensional ultrasonography. Croatian medical journal. 2005; 46: 772-780.

Kochukhova, O., & Gredebäck, G. (2010). Preverbal infants anticipate that food will be brought to the mouth: An eye tracking study of manual feeding and flying spoons. *Child development*, *81*(6), 1729-1738.

LaBarbera, J. D., Izard, C. E., Vietze, P., & Parisi, S. A. (1976). Four-and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child development*, 535-538.

Legerstee, M. (1990). Infants use multimodal information to imitate speech sounds. *Infant behavior and development*, *13*(3), 343-354.

Leppänen, J. M., Richmond, J., Vogel-Farley, V. K., Moulson, M. C., & Nelson, C. A. (2009). Categorical representation of facial expressions in the infant brain. *Infancy*, *14*(3), 346-362.

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child development*, *78*(1), 232-245.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. Nature Reviews Neuroscience, 10(1), 37-47.

Leppänen, J. M. (2011). Neural and developmental bases of the ability to recognize social signals of emotions. Emotion Review, 3(2), 179-188.

Lewkowicz, D. J. (1986). Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior and Development*, *9*(3), 335-353.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychological bulletin*, *126*(2), 281.

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: newborns match nonhuman primate faces and voices. *Infancy*, *15*(1), 46-60.

Lewkowicz, D. J., & Pons, F. (2013). Recognition of amodal language identity emerges in infancy. *International journal of behavioral development*, *37*(2), 90-94.

Lewkowicz, D. J., Minar, N. J., Tift, A. H., & Brandon, M. (2015). Perception of the multisensory coherence of fluent audiovisual speech in infancy: Its emergence and the role of experience. *Journal of experimental child psychology*, *130*, 147-162.

Lew A & Butterworth GE. Hand-mouth contact in newborn babies before and after feeding. Developmental Psychology. 1995; 31: 456–63.

Lickliter, R. (2011). The integrated development of sensory organization. *Clinics in perinatology*, *38*(4), 591-603.

Lickliter, R. (2000). The role of sensory stimulation in perinatal development: insights from comparative research for care of the high-risk infant. *Journal of Developmental & Behavioral Pediatrics*, *21*(6), 437-447.

Lickliter, R., & Bahrick, L. E. (2001). The salience of multimodal sensory stimulation in early development: Implications for the issue of ecological validity. *Infancy*, *2*(4), 451-467.

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye tracking in 4- to 9-month-olds. *Journal of experimental Child Psychology*, 108(1), 180–189.

Longhi E, Senna I, Bolognini N, Bulf H, Tagliabue P, Macchi Cassia V, et al. Discrimination of Biomechanically Possible and Impossible Hand Movements at Birth. Child Dev. 2014; doi: 10.1111/cdev.12329.

Ludemann, P. M., & Nelson, C. A. (1988). Categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology*, *24*(4), 492.

Macchi, C. V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference?. *Psychological Science*, *15*(6), 379-383.

Macchi CV, Turati C, Simion F. Can a nonspecific bias toward top-heavy patterns explain newborns' face preference?. Psychological Science. 2004; 15: 379-383.

Macchi Cassia V, Simion F, Umiltà C. Face preference at birth: The role of an orienting mechanism. Developmental Science. 2001; 4: 101-108.

Machado, C. J., & Bachevalier, J. (2003). Non-human primate models of childhood psychopathology: the promise and the limitations. Journal of Child Psychology and Psychiatry, 44(1), 64-87.

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human perception and performance*, *17*(4), 1057.

Marshall, P. J., & Meltzoff, A. N. (2011). Neural mirroring systems: Exploring the EEG mu rhythm in human infancy. Developmental Cognitive Neuroscience, 1(2), 110-123.

Marshall, P. J., & Meltzoff, A. N. (2014). Neural mirroring mechanisms and imitation in human infants. Phil. Trans. R. Soc. B, 369(1644), 20130620.

Markham, R. G., Shimizu, T., & Lickliter, R. (2008). Extrinsic embryonic sensory stimulation alters multimodal behavior and cellular activation. *Developmental neurobiology*, *68*(13), 1463-1473.

Marx, V., & Nagy, E. (2015). Fetal behavioural responses to maternal voice and touch. *PloS one*, *10*(6), e0129118.

Mazzola, V., Vuilleumier, P., Latorre, V., Petito, A., Gallese, V., Popolizio, T., ... & Bondolfi, G. (2013). Effects of emotional contexts on cerebello-thalamo-cortical activity during action observation. *PloS one*, *8*(9), e75912.

Meltzoff, A. N., & Moore, M. K. (1983). The origins of imitation in infancy: Paradigm, phenomena, and theories. *Advances in infancy research*.

Meltzoff AN, Moore MK. Explaining facial imitation: A theoretical model. Early development and parenting. 1997; 6: 179-192.

Melinder, A. M. D., Konijnenberg, C., Hermansen, T., Daum, M. M. & Gredebäck, G., (2015). The developmental trajectory of pointing perception in the first year of life. Experimental Brain Research, 233, 641-47. doi: 10.1007/s00221-014-4143-2.

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. Developmental science, 10(1), 126-134.

Meltzoff, A. N. (2013). Origins of Social Cognition Bidirectional Self-Other Mapping and the "Like-Me" Hypothesis.

Meyer K, Kaplan JT, Essex R, Damasio H, Damasio A. Seeing touch is correlated with content-specific activity in primary somatosensory cortex. Cerebral Cortex. 2011; 21: 2113-2121.

Montagu A. Touching: The human significance of the skin. Columbia University Press; 1971.

Missana, M., Rajhans, P., Atkinson, A. P., & Grossmann, T. (2014). Discrimination of fearful and happy body postures in 8-month-old infants: an event-related potential study. *Frontiers in human neuroscience*, *8*, 531.

Missana, M., Atkinson, A. P., & Grossmann, T. (2015). Tuning the developing brain to emotional body expressions. *Developmental science*, *18*(2), 243-253.

Myowa-Yamakoshi M, Takeshita H. Do Human Fetuses Anticipate Self-Oriented Actions? A Study by Four-Dimensional (4D) Ultrasonography. Infancy. 2006; 10: 289-301.

Nagy E, Pal A, Orvos H. Learning to imitate individual finger movements by the human neonate. Developmental science. 2014; doi: 10.1111/desc.12163.

Negri, G. A., Rumiati, R. I., Zadini, A., Ukmar, M., Mahon, B. Z., & Caramazza, A. (2007). What is the role of motor simulation in action and object recognition? Evidence from apraxia. Cognitive Neuropsychology, 24(8), 795-816.

Nelson, C.A., Thomas, K.M., & de Haan, M. (2006). Neural bases of cognitive development. In W. Damon, R. M. Lerner, D. Kuhn, & R. S. Siegler (Eds.), Handbook of Child Psychology, Cognition, Perception, and Language (pp. 3–57). New York: Wiley.

Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child development*, 1239-1242.

Nelson, C. A. (2014). Neural measures of social attention across the first years of life: characterizing typical development and markers of autism risk. Developmental Cognitive Neuroscience, 8, 131-43. doi:10.1016/j.dcn.2013.09.006.

Nyström, P. (2008). The infant mirror neuron system studied with high density EEG. *Social neuroscience*, *3*(3-4), 334-347.

Oberman, L. M., Pineda, J. A., & Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Social cognitive and affective neuroscience*, *2*(1), 62-66.

Ortigue, S., Sinigaglia, C., Rizzolatti, G., & Grafton, S. T. (2010). Understanding actions of others: the electrodynamics of the left and right hemispheres. A high-density EEG neuroimaging study. *PloS one*, *5*(8), e12160.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. Neuroimage, 48(3), 564-584

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, *18*(1), 79-85.

Patterson, M. L., & Werker, J. F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development*, *22*(2), 237-247.

Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of experimental child psychology*, *81*(1), 93-115.

Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, *6*(2), 191-196.

Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, nsn046.

Pihko, E., Nevalainen, P., Stephen, J., Okada, Y., Lauronen, L., 2009. Maturation of somatosensory cortical processing from birth to adulthood revealed by magnetoencephalography. Clin. Neurophysiol. 120, 1552–1561

Piontelli, A. (2014). *Development of normal fetal movements*. Springer.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*(1), 53-78.

Posner, M. I. (1980). Orienting of attention. The Quarterly Journal of Experimental Psychology, 32, 3–25. doi:10.1080/00335558008248231.

Proverbio, A. M., Riva, F., Paganelli, L., Cappa, S. F., Canessa, N., Perani, D., & Zani, A. (2011). Neural coding of cooperative vs. affective human interactions: 150 ms to code the action's purpose. *PloS one*, *6*(7), e22026.

Quadrelli, E., & Turati, C. (2016). Origins and development of mirroring mechanisms: A neuroconstructivist framework. British Journal of Developmental Psychology, 34(1), 6-23.

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental science*, *14*(1), 92-105.

Reid, V.M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cuing facilitates neural processing of objects in 4-month-old infants. NeuroReport, 15, 2553-5. doi:10.1097/00001756-200411150-00025

Reissland N, Francis B, Aydin E, Mason J, Schaal B. The development of anticipation in the fetus: A longitudinal account of human fetal mouth movements in reaction to and anticipation of touch. Developmental Psychobiology. 2013; 56: 955-963.

Reissland, N., Francis, B., Mason, J., & Lincoln, K. (2011). Do facial expressions develop before birth?. *PLoS One*, 6(8), e24081.

Rayson, H., Bonaiuto, J. J., Ferrari, P. F., & Murray, L. (2016). Mu desynchronization during observation and execution of facial expressions in 30-month-old children. Developmental cognitive neuroscience, 19, 279-287.

Richards, E. J. (2000). Localizing the development of covert attention in infants with scalp event-related potentials. Developmental Psychology, 36, 91-108. doi:10.1037//00121649.36.1.91

Richards, J. E. (2003). Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Developmental Science*, *6*(3), 312-328.

Ritter, P., Moosmann, M., & Villringer, A. (2009). Rolandic alpha and beta EEG rhythms' strengths are inversely related to fMRI-BOLD signal in primary somatosensory and motor cortex. Human brain mapping, 30(4), 1168-1187.

Rigato, S., Menon, E., Johnson, M. H., Faraguna, D., & Farroni, T. (2011). Direct gaze may modulate face recognition in newborns. *Infant and Child Development*, 20(1), 20-34.

Rosenstein, D., & Oster, H. (1988). Differential facial responses to four basic tastes in newborns. *Child development*, 59, 1555-1568.

Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. Developmental Science, 15, 426–35. doi: 10.1111/j.1467-7687.2012.01139.x.

Rossetti A, Miniussi C, Maravita A, Bolognini N. Visual perception of bodily interactions in the primary somatosensory cortex. European Journal of Neuroscience. (2012); 36: 2317-2323.

Rochat P, Hespos SJ. Differential rooting response by neonates: Evidence for an early sense of self. Early development and Parenting. 1997; 6: 105-112.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience, 2(9), 661-670.

Sai, F. Z. (2005). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development*, *14*(1), 29-50.

Sato, M., Kanenishi, K., Hanaoka, U., Noguchi, J., Marumo, G., & Hata, T. (2014). 4D ultrasound study of fetal facial expressions at 20–24weeks of gestation. *International Journal of Gynecology & Obstetrics*, 126(3), 275-279.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Cognitive Brain Research*, 20(1), 81-91.

Saxe R, Tzelnic T, Carey S. Knowing who dunnit: Infants identify the causal agent in an unseen causal interaction. Developmental Psychology. 2007; 43: 149-158.

Schaal, B., Lecanuet, J. P., & Granier-Deferre, C. (1999). Sensory and integrative development in the human fetus and perinate: The usefulness of animal models.

Schaefer M, Xu B, Flor H, Cohen LG. Effects of different viewing perspectives on somatosensory activations during observation of touch. Human brain mapping. 2009; 30: 2722-2730.

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. Current Biology, 18, 668–671. Doi:10.1016/j.cub.2008.03.059.

Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. Social Neuroscience, 1, 220–234. doi: 10.1080/17470910600989797.

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4-to 6-month-old infants. *Developmental Psychobiology*, *25*(6), 411-425.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. Proceedings of the National Academy of Sciences, 105(2), 809-813.

Slater A. (1995). Visual perception and memory at birth. In C. Rovee-Collier & L.P. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 107–162). Norwood, NJ: Ablex.

Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory– visual pairings. *Developmental Science*, *2*(3), 333-338.

Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, *96*(1), B1-B11.

Sparling JW, Van Tol, Chescheir NC. Fetal and neonatal hand movement. PhysicalTherapy. 1999; 79: 24-39.

Sparling JW, Wilhelm IJ. Quantitative Measurement of Fetal Movement: Fetal-Post and Movement Assessment (F-PAM). Physical & Occupational Therapy in Pediatrics. 1993; 1: 97-114.

Stapel, J. C., Hunnius, S., Meyer, M., & Bekkering, H. (2016). Motor system contribution to action prediction: temporal accuracy depends on motor experience. *Cognition*, *148*, 71-78.

Stein, B. E., Perrault Jr, T. J., Stanford, T. R., & Rowland, B. A. (2009). Postnatal experiences influence how the brain integrates information from different senses. *Frontiers in integrative neuroscience*, *3*, 21.

Streri, A., & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory & motor research*, *20*(1), 13-18.

Streri, A., & Gentaz, E. (2004). Cross-modal recognition of shape from hand to eyes and handedness in human newborns. *Neuropsychologia*, *42*(10), 1365-1369.

Szufnarowska, J., Rohlfing, K. J., Fawcett, C., & Gredebäck, G. (2014). Is ostension any more than attention? Scientific Reports, 4, 5304. doi:10.1038/srep05304.

Jabbi, M., Swart, M., & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. Neuroimage, 34(4), 1744-1753.

Turati, C., Montirosso, R., Brenna, V., Ferrara, V., & Borgatti, R. (2011). A Smile Enhances 3-Month-Olds' Recognition of an Individual Face. *Infancy*, 16(3), 306-317.

Vaillant-Molina, M., Bahrick, L. E., & Flom, R. (2013). Young infants match facial and vocal emotional expressions of other infants. *Infancy*, *18*(s1), E97-E111.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological bulletin*, *134*(3), 383.

Valenza E, Simion F, Cassia VM, Umiltà C. Face preference at birth. Journal of Experimental Psychology: Human Perception and Performance. 1996; 22: 892-903.

Valenza, E., & Bulf, H. (2007). The role of kinetic information in newborns' perception of illusory contours. *Developmental science*, 10(4), 492-501.

Vecera, S.P., Johnson, M.H. (1995). Eye gaze detection and the cortical processing of faces: Evidence from infants and adults. Visual Cognition, 2, 101–129. doi:10.1080/13506289508401722.

Võ, M. L. H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of vision*, 12(13), 3-3.

Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice?. *Developmental psychology*, *22*(3), 373.

Wheaton, K. J., Pipingas, A., Silberstein, R. B., & Puce, A. (2001). Human neural responses elicited to observing the actions of others. *Visual neuroscience*, *18*(3), 401-406.

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian Infants Scan Own- and Other-Race Faces Differently. *PLoS One*, 6(4), e18621.

Wilcox, B. M., & Clayton, F. L. (1968). Infant visual fixation on motion pictures of the human face. *Journal of experimental child psychology*, *6*(1), 22-32.

Wilcox, T., Stubbs, J. A., Wheeler, L., & Alexander, G. M. (2013). Infants' scanning of dynamic faces during the first year. *Infant Behavior and Development*, 36(4), 513-516.

Woodward AL. Infants selectively encode the goal object of an actor's reach. Cognition. 1998; 69: 1-34.

Woodward AL. Infants' ability to distinguish between purposeful and non-purposeful behaviors. Infant Behavior and Development. 1999; 22: 145-160.

Xiao, N. G., Quinn, P. C., Liu, S., Ge, L., Pascalis, O., & Lee, K. (2015). Eye tracking reveals a crucial role for facial motion in recognition of faces by infants. *Developmental psychology*, 51(6), 744.

Yuan, J., Lu, H., Yang, J., & Li, H. (2011). Do not neglect small troubles: moderately negative stimuli affect target processing more intensely than highly negative stimuli. *Brain research*, *1415*, 84-95.

Yeung, H. H., & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science*, 0956797612458802.

Young-Browne, G., Rosenfeld, H. M., & Horowitz, F. D. (1977). Infant discrimination of facial expressions. *Child Development*, 555-562.

Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). The development of intermodal emotion perception from bodies and voices. *Journal of experimental child psychology*, *126*, 68-79.

Zoia S, Blason L, D’Ottavio G, Bulgheroni M, Pezzetta E, Scabar A, et al. Evidence of early development of action planning in the human foetus: a kinematic study. Experimental Brain Research. 2007; 176: 217-226.

Zmyj N, Jank J, Schütz-Bosbach S, Daum MM. Detection of visual–tactile contingency in the first year after birth. Cognition. 2011; 120: 82-89.