Cognitive capacities of the infant mind — a neuroimaging perspective

Mohinish Shukla, Vivian Ciaramitaro
Department of Psychology
University of Massachusetts, Boston

Corresponding author:
Mohinish Shukla
DBS, Department of Psychology
McCormack Hall, University of Massachusetts, Boston
Boston, MA - 02125, USA
Phone: +1 617 287 6302
E-mail: mohinish.shukla@umb.edu

Author Note

Substantial parts of this chapter are reproduced with permission from Aslin, Shukla, and Emberson (2015).
Abstract

Although the mental life of infants remains a mystery, behavioral and neurophysiological studies have amassed a great deal of data demonstrating similarities and differences between infants and adults in the processing of complex stimuli in several cognitive domains. Here we concentrate on several such studies in diverse domains including the perception of material objects and events, and the processing of socially relevant stimuli such as faces and language. These findings suggest cognitive continuity between infants and adults, and support the view that, in several respects, their mental lives might be comparably intricate.
Cognitive capacities of the infant mind — a neuroimaging perspective

**Introduction**

The average layperson, if asked whether humans are conscious, would probably reply in the affirmative. But if, instead of “humans” we asked if the “human infant” is conscious, it’s not clear what the average lay person would say. As recently as 1890, William James proposed that the sensory world of the infant was an undifferentiated mess, which he colorfully described as a “blooming, buzzing confusion” (James, 1890/1981). Indeed, in the behaviorist tradition, and its modern incarnations, the infant is proposed to have very little actual mental content, and that content is slowly acquired from the environment using (sophisticated) general learning mechanisms. By mental content, we mean, among other things, constructs, such as concepts, thoughts, beliefs, or desires, that are intentional, and to some extent represent knowledge about the world. If the mind of an infant is indeed a blooming, buzzing confusion, then it’s not clear what substantial stuff they could be conscious of. For example, it wasn’t until the 1980s that infants were given nothing more than token anesthetics when undergoing a surgical intervention. The assumption was that infants could not be conscious of pain. It took empirical studies documenting the negative consequences of operating on infants without anesthetics to make pediatric anesthetic use commonplace (Anand & Hickey, 1987).

Clearly, the idea that infants are conscious has profound impacts in how we treat them. While the use of pediatric anesthesia is an extreme example, it is still important to consider how we see infants in terms of their cognitive capacities. Take the case of the language we speak. In the 1964 movie classic *My Fair Lady*, Professor Henry Higgins suggests that it’s the flower-girl Eliza’s Cockney accent that marks her low social class and that if the Colonel Pickering were to speak as she did, why, he might be selling flowers too. Indeed, social psychologists have documented that the language one
speaks has a profound influence on how one is viewed by others (e.g., Gluszek & Dovidio, 2010). How, then, should we view infants, who cross-culturally do not appear to have any language at all? Can we safely assume that they have no conception of language at all (as is common in the empiricist tradition), or should we assume that they have a broad, and general understanding of language, that is subsequently narrowed to the language in their environment? The same question can be asked of other cognitive capacities that we believe human adults possess, including knowledge of objects and events, of numbers, rationality, and even basic social skills.

In this chapter, we will examine studies that have looked at cognitive capacities in infants in several such domains, with a special focus on converging evidence provided by behavioral and neuroimaging studies. The neuroimaging studies we review utilize electrophysiology techniques (electroencephalography, EEG, and event-related potentials, ERPs; see Srinivasan & Nunez, 2012, for an overview), and hemodynamic measurements (including functional magnetic resonance imaging, fMRI, and functional near-infrared spectroscopy, fNIRS; see Aslin et al., 2015, for a recent review of the use of hemodynamic measures in studying infant cognition). Our coverage of these studies is not intended to be complete; we have instead selected studies that parallel and reinforce each other. We cannot say with any certainty whether or not infants are conscious of objects, social relations etc., but we can at least say that they are capable of certain relevant computations in these domains. We will primarily highlight the processing of objects in the physical domain, and of faces and language in the social domain.

Object representation and individuation

A fundamental distinction in how we see the world, is as objects and actions. That is, our perceptual systems encode some algorithms for parsing and grouping the perceptual world into discrete entities, and then assign these entities properties that can be static (like their shape or color), or dynamic (like hopping or slithering). Several
-lines of evidence suggest that infants start out with impoverished representations of objects, which they subsequently refine through experience (e.g., Baillargeon et al., 2012), although some classes of stimuli, like human faces, appear to be have more detailed initial representations (see next section).

Behavioral studies have been largely concerned with understanding early cognitive capacities and constraints on such object individuation. However, methodologically speaking, it is difficult to understand how an infant is parsing a scene that she is watching, so researchers have primarily relied on paradigms that require infants to parse a scene into relevant object representations and hold these representations in memory, for subsequent recall. Using such encoding-and-recall behavioral paradigms, researchers have found evidence for object permanence — the ability to maintain a memory trace of a previously-seen object, in infants as young as 2–3 months of age (e.g., Aguiar & Baillargeon, 1999). In a typical study, infants are exposed to small objects that appear and disappear from either side of an opaque screen. Researchers vary properties of the appearing and disappearing objects to probe the nature of what features of an object infants encode. For example, a red ball might emerge from the left edge of an opaque occluder, and then turn back and disappear behind it. Subsequently, if a red cube similarly emerges and disappears from the right edge of the opaque occluder, infants around 7 months of age are surprised (look longer) if the opaque occluder drops to reveal only the red cube, suggesting that they were presumably expecting two objects. A change in color alone (e.g., using a green ball instead of the red cube from the right edge) does not lead to such a surprise reaction, suggesting that infants’ early object representations preferentially encode shape over color. Recent work however suggests that the demands for parsing and those for encoding in memory might be in conflict, specially for infants with their limited resources, such that infants might need to prioritize encoding objects versus encoding the features of the objects (Kibbe & Leslie, 2013).
Turning to neuroimaging, EEG measures in 7- to 12-month-olds have been shown to correlate with better behavioral performance on object permanence tasks (Bell & Fox, 1992). In this study, resistance to increasing delays between the disappearance of the final object and the lowering of the opaque screen (indicating a stronger memory trace of the object) were accompanied by stronger neuronal activity or bigger frontal EEG power. fNIRS has also been used to study object permanence in infants (Baird et al., 2002). In a series of linked studies, Wilcox and colleagues have used fNIRS to investigate hemodynamic responses in ~7-month-old infants. Using paradigms similar to the ones described above, these authors found that infants react to changes in shape and texture, but not in color; and that these behavioral findings are reflected in increased hemodynamic responses for shape and texture changes compared to color changes, in fNIRS channels placed over the temporal cortex (see Wilcox & Biondi, 2015, for a recent review).

Related to infants’ early capacity for object individuation, is their capacity to compute how many objects are present in an array. A wealth of experimental and theoretical work has revealed that there are two core systems for numerical thinking – one system (Core 1) for representing the approximate magnitude of large collections of objects, and a second system (Core 2) for representing the precise cardinality of sets of up to three objects (e.g., Feigenson, Dehaene, & Spelke, 2004).

Studies have carefully controlled that behavioral results implicating a representation of numerosity do not arise from confounds introduced by changes in low-level stimulus features that can co-vary with the number of objects displayed, such as size, surface area, or density (Lipton & Spelke, 2003; Xu & Spelke, 2000). Furthermore, studies have shown that numerical discriminations are independent of the sensory modality of the input; infants can discriminate quantities presented by either visual or auditory objects (Izard, Sann, Spelke, & Streri, 2009; Jordan & Brannon, 2006; Lipton & Spelke, 2003).
Recently, non-invasive neuroimaging methodologies have been used to investigate the neural correlates of the number sense, revealing many of the same underlying brain networks for numerosity in adulthood and in the developing brain. In adults, the intraparietal sulcus in the posterior parietal region has been shown to represent approximate numerical information, irrespective of the sensory modality of the input, or the use of symbolic or non-symbolic representations (for review see Dehaene, Molko, Cohen, & Wilson, 2004). Furthermore, small numerosities were encoded closer to the midline, while larger quantities were encoded more laterally, suggesting that the behavioral difference between the two Core systems is accompanied by a corresponding anatomical difference in localization.

Just as in adults, infants also show activity in parietal cortex when performing non-verbal numerical tasks. A recent fNIRS study by Hyde and colleagues finds that the parietal cortex is active in 6-month-old infants, an age before language development and before experience with symbolic numerical representations (Hyde, Boas, Blair, & Carey, 2010). Turning to ERP studies in infants, Hyde and Spelke (2011) found that 6–7.5-month-olds showed distinct ERP components sensitive to precise, small numbers versus approximate, large numbers. ERP evidence also suggests selective processing in parietal cortex for numerosity as opposed to other low-level sensory features (Izard, Dehaene-Lambertz, & Dehaene, 2008). Thus, specialization for number in the infant brain can be seen independently of any formal mathematical experience or education or even before language.

**Social Cognition**

Turning away from the material world of objects and their numerosities, let us now consider the human as a social animal. As such, we would expect human infants to display early social capacities. That is, in addition to bringing some primitive computations about their physical world, we would expect infants to also bring some
primitive computations relevant to their social world. Indeed, newborn infants show preferences for socially relevant stimuli, in particular faces, described below, and speech and language, described in the next section.

A face is an important, socially relevant stimulus category not just because it signals the identity of a person, but also because it conveys information about mental states, for example those expressed through emotions. The standard view of the development of face recognition had been that the ability to discriminate different faces was not present until 10 years of age and that such an ability was driven primarily by experience (Carey & Diamond, 1977; Carey, Diamond, & Woods, 1980). Classic studies in newborns have long since discredited this view, providing evidence that even newborns have the machinery for face recognition. In one such study, Goren, Sarty, and Wu (1975) found that newborn infants preferentially oriented towards face-like stimuli with primitive “eye” and “mouth”-like features. Infants did not prefer such stimuli when they were presented upside-down, suggesting a particular sensitivity to specific feature arrangements that might be indicative of faces in their visual input. Subsequent experiments have refined this notion of an early face-like template, to include features like the contrast polarity of the eyes (black pupils on a white background; Farroni et al., 2005).

In adults, evidence for a specialized mechanism for face processing comes both from studies of prosopagnosia — a neuropsychological condition with a specific deficit in the visual processing of faces (e.g., McNeil & Warrington, 2013) — and imaging studies that have pinpointed a specific brain region, the fusiform face area (FFA: Kanwisher, McDermott, & Chun, 1997). Unfortunately, the FFA is inaccessible using fNIRS owing to its distance from the scalp, and there have not been many studies using fMRI with very young children (but see Cantlon, Pinel, Dehaene, & Pelphrey, 2011). Researchers have used fNIRS to examine additional neural areas that might be involved in processing face stimuli in 5–8-month-old infants. These studies examine activations
to faces compared to control stimuli, including inverted faces or pictures of vegetables (Otsuka et al., 2007), and typically find that faces preferentially engage lateral (temporal) cortices. Although Otsuka et al. (2007) found a right-hemisphere (RH) advantage for faces, subsequent studies found engagement of bilateral temporal areas, and additionally showed that these areas show higher activations for blocks of several different faces, compared to blocks with a single, repeated face, and also show size and viewpoint invariance (Kobayashi et al., 2011; Kobayashi, Otsuka, Kanazawa, Yamaguchi, & Kakigi, 2012).

ERP studies find that infants as young as 3 months of age show enhanced amplitude of a particular component, the N290, for upright relative to inverted faces; such an inversion effect is not found for upright versus inverted cars (Peykarjou & Hoehl, 2013). Further, different ERP components have been found to be active in 7-month-old infants when viewing happy versus fearful or angry faces (Nelson & De Haan, 1996; Leppänen, Moulson, Vogel-Farley, & Nelson, 2007). Jessen and Grossmann (2015) relied on previous behavioral and ERP studies in adults and ERP studies in 5- and 7-month-old infants (Kouider et al., 2013) to define supra- and subliminal face stimuli, and found that neuronal responses in 7-month-old infants can not only discriminate between happy versus fearful facial expressions but that these neuronal responses also depend on whether the information about the face is being processing subliminally (subconsciously) or supraliminally (consciously).

Put together, the imaging studies provide converging evidence that in their first year of life infants are particularly tuned to various aspects of a critical, socially relevant visual stimulus, a human face. In the next section we will turn to language, probably the most important human feature, and of critical importance in all social relations.
Language

Language is considered to be a defining feature of our species (Maynard Smith & Szathmary, 1995; Szathmáry, 2015). Language plays a particularly crucial role in the study of the mind and consciousness, as it allows the communication of mental content, including knowledge and beliefs. As noted in the introduction, a lack of language could potentially be interpreted as a lack of mental content, which makes the study of language capacities in infants specially important.

The preponderance of studies in infant language research primarily document changes in behavioral and neurophysiological responses to speech-like and non-speech like stimuli, in order to construct theories for the development of various linguistic competencies and their brain bases (e.g., Minagawa-Kawai et al., 2011). Indeed, behavioral studies have revealed that speech is perceived as a special kind of input by young infants. Infants show a preference for speech over other sounds, including backward speech, and speech engages special processing algorithms, including categorization of sets of visual objects at 3 to 4 months of age (e.g., Ferry, Hespos, & Waxman, 2010) or cueing referential communication at 6 months (Senju & Csibra, 2008).

Infant imaging studies have revealed that speech shows asymmetric patterns of activity across the two brain hemispheres, indicating cortical specialization for speech stimuli, even at birth (Peña et al., 2003; Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002). In particular, compared to non-speech, speech preferentially activates the left hemisphere. In one such study, (Peña et al., 2003) studied neonates using fNIRS to examine patterns of cortical activations to normal, Forward speech, compared to Backward speech (in which the same recorded speech samples are played backwards); and to Silence, when no stimulus is presented. These authors showed that the neonate brain already showed cortical specialization: while channels over the right temporal
cortex showed similar activations to both *Forward* and *Backward* speech, channels over the left temporal cortex showed preferential activation for *Forward* over *Backward* speech.

Similar studies with fMRI have not only found left-lateralized responses specifically to speech, but have also identified other, frontal areas that are asymmetrically active for speech as compared to non-speech in older infants (2- to 3-month-olds in Dehaene-Lambertz et al., 2002; 21-month-olds in Redcay, Haist, and Courchesne, 2008).

Likewise, studies using EEG have found greater left hemisphere activity when infants listen to speech and greater right hemisphere activity when infants listen to non-speech sounds, such as noises and piano chords in one-week to 10-month-old infants (Molfese, Freeman, and Palermo, 1975; also see Molfese and Molfese, 1979; 1980; 1985). Furthermore, the left-lateralization of responses is present not only for spoken, but also for signed languages (Petitto et al., 2000). Importantly, ERP measures to speech in newborns are predictive of later language proficiency (Molfese & Molfese, 1985, 1997).

Not all aspects of speech perception are left lateralized. In particular, it has been proposed that the left hemisphere is primarily responsible for computations over shorter time-scales, while the right hemisphere is responsible for computations over longer time-scales (see reviews in Poeppel, 2003; Hickok & Poeppel, 2007). This division of labour in terms of the temporal aspect of speech corresponds with the proposed division of the phonological organization of speech sounds at the segmental (e.g., consonants and vowels) level and the suprasegmental, prosodic level, including changes in pitch, intensity, and duration. For example, a sentence like “It is raining” can be spoken as a statement or a question – in both cases, the segmental level is constant, while the prosody varies. Neurophysiological studies with infants have shown preferential activation for prosody both at the level of the sentence (e.g., Homae, Watanabe, Nakano, Asakawa, & Taga, 2006, with Japanese 3-month-olds) and at the level of the
Language is much more than speech. Speech acts are but the linear externalizations of abstract, hierarchically organized grammatical structures. Although most imaging studies have investigated the processing of speech, few have addressed the question – can the mind/brain of the young infant extract abstract structure from the speech input? Starting with the now-classic study by Marcus and colleagues (Marcus, Vijayan, Bandi Rao, & Vishton, 1999), it has been shown that infants can extract simple structures from short exposures to carefully controlled artificial speech stimuli. For example, when trained with trisyllabic sequences such as ‘de-li-li,’ 7-month-old infants subsequently discriminated novel sequences that followed the $A-B-B$ pattern (such as ‘wo-fe-fe’) from ones that did not (such as ‘wo-fe-wo’). Recently, the impressive power of 7-month-olds’ abstract processing capacity was revealed in a study that asked if infants could represent an $A-B-B$ pattern if the elements of the pattern were not just single syllables, as in the original Marcus study, but were themselves sequences that followed one rule or another (Kovács & Endress, 2014). For example, an $A$ element could be ‘wo-fe-fe’ or ‘de-li-li,’ while a $B$ element could be ‘wo-fe-ga’ or ‘de-li-me.’ Remarkably, 7-month-olds were able to learn these rules-over-rules, suggesting that infants possess some ability to represent their perceptual input as hierarchically organized structures.

Imaging studies have examined neural responses that accompany the extraction of structure from organized, artificial speech stimuli. Using fNIRS, Gervain, Macagno, Cogoi, Peña, and Mehler (2008) found that trisyllable blocks (that is, a series of trisyllables separated by silence) that had an $A-B-B$ structure showed greater activation in left cortical areas compared to random syllable triplets (an $A-B-C$ structure). Further, the left-lateralized cortical response over frontal areas showed a progressive enhancement over the experimental period, suggesting that frontal areas might have been responding to increased familiarity with the structure, indicating not
just analysis, but possibly also learning (see also Gervain, Berent, & Werker, 2012; Wagner, Fox, Tager-Flusberg, & Nelson, 2011).

Taken together, the behavioral and neuroimaging data provide converging evidence that the newborn baby’s brain is prepared to rapidly identify, analyze, and learn from socially important stimuli in its perceptual input, whether this be faces (previous section) or language.

Conclusions

Imagine a boulder hurtling down a mountain path that forks just ahead of the boulder\(^1\). When we see the boulder take the path on its left instead of the one on its right, we don’t infer that the boulder made a “choice,” but that small variations in the physical properties of the path just before the fork led to the boulder following one path over the other, in a purely deterministic fashion. Now imagine a runner sprinting down the path and also taking the path on her left. In this case, we are likely to attribute the selection of the path to some internal mental state like a belief or a desire. That is, purely behavioral observations are not sufficient to infer mental states. Similarly, researchers who study consciousness are acutely aware of the seemingly intractable philosophical problem of trying to study something as intangible as mental states through video recordings or measurements of voltage distributions across the scalp (e.g., Kouider et al., 2013).

Instead of addressing this complex question, we can ask – is the mental life and the accompanying brain physiology of infants significantly different from that of adults? In this chapter we have tried to summarize behavioral and imaging evidence that suggests that, right from birth, infants display evidence for possessing sophisticated cognitive capacities in several cognitive domains. Moreover, these capacities are equivalent to similar capacities in adults, either in their behavioral outcomes, in their

\(^1\)This analogy is loosely borrowed from Roeper (2007), Ch. 12
physiology, or both.

Given these substantial similarities, and given that the baby human will grow into an adult human, it is reasonable to ask, is an infant merely an adult without the years of experience? The results summarized here suggest that this might indeed be the case. That is, the similarities between infants and adults ought to compel us to recognize that the mental life of an infant might be substantially similar to that of an adult, however the mental life of an adult is constructed. For example, if a linguistic theory of the adult language faculty requires constructs such as nouns or verbs, then there is no a priori reason to suspect that these same constructs are not part of the mental life of infants. In his book *The Prism of Grammar*, Tom Roeper gathers together a wealth of evidence to show that the earliest utterances of infants reflect deep, abstract structures, and creative processes that form the basis of human thought (Roeper, 2007). We hope that our survey of the cognitive abilities of infants, as assessed through behavioral of physiological methods, provide further impetus to the idea that our infants are humans much like us, and thereby afford them the same dignity that is extended to all other members of our species.
References


