Infant Joint Attention, Neural Networks and Social Cognition

Peter Mundy and William Jarrold
School of Education, Department of Psychiatry and the M.I.N.D. Institute, University of California at Davis, 2315 Stockton Blvd., Sacramento, CA, 95817-9956, 916-702-0310

Abstract

Neural network models of attention can provide a unifying approach to the study of human cognitive and emotional development (Posner & Rothbart, 2007). This paper we argue that a neural networks approach to the infant development of joint attention can inform our understanding of the nature of human social learning, symbolic thought process and social cognition. At its most basic, joint attention involves the capacity to coordinate one’s own visual attention with that of another person. We propose that joint attention development involves increments in the capacity to engage in simultaneous or parallel processing of information about one’s own attention and the attention of other people. Infant practice with joint attention is both a consequence and organizer of the development of a distributed and integrated brain network involving frontal and parietal cortical systems. This executive distributed network first serves to regulate the capacity of infants to respond to and direct the overt behavior of other people in order to share experience with others through the social coordination of visual attention. In this paper we describe this parallel and distributed neural network model of joint attention development and discuss two hypotheses that stem from this model. One is that activation of this distributed network during coordinated attention enhances to depth of information processing and encoding beginning in the first year of life. We also propose that with development joint attention becomes internalized as the capacity to socially coordinate mental attention to internal representations. As this occurs the executive joint attention network makes vital contributions to the development of human symbolic thinking and social cognition.

In this paper we provide an overview of theory and research on how a neural networks outlook on joint attention assists in our understanding of the mechanisms that support human social learning, symbolic thinking and social cognition. This theory and research has stemmed in large part from the contemporary pursuit of the question; how do people become capable of sharing information with one another? One immediate answer is that it has a lot to do with the phylogeny and ontogeny of bio-behavioral systems specific to language. However, in the later part of the 20th century it became clear that vital elements of the cognitive foundations for information sharing develop before the onset of language in human infancy (Bates, et al. 1979; Bruner, 1975; Werner & Kaplan, 1963). One especially pertinent observation was that by 6 to 9 months infants become increasingly capable of sharing experience about objects and events by directing or following the visual gaze of social partners (Bakeman & Adamson, 1984; Bates et al., 1979; Scaife & Bruner, 1975, see Figure 1). Bruner (1975) referred to this type of preverbal referential behavior as joint attention. He suggested that joint attention...
behaviors reflect the early onset of psychological processes that are specific to social reference but are distinct from, albeit necessary to the ontogeny of language.

By and large research over the past 30 years has supported Bruner’s hypothesis about the distinct nature of communication versus language development and the primacy of infant joint attention development with regard to the former. For example, Willems, Boerl, Ruiter Noordzig et al. (2009) have recently reported evidence for two distinct but interacting neurocognitive systems in adults, one unique to communication processes and one unique to language processes. Studies and theory also suggest that joint attention is pivotal to the development of the human neurocognitive communication functions (Mundy 1995; 2003; Mundy, Sullivan & Mastergeorge, 2009) and that the neurocognitive joint attention functions are integral to the subsequent development of social cognition, as well as language (e.g. Baldwin, 1995; Charman et al. 2000; Hirotani et al. 2009; Kasari et al. 2008; Kwisthout et al. 2009; Meltzoff & Brooks, 2008; Mundy, Fox & Card et al. 2003). More to the point of this article recent theory suggests that neurocognitive joint attention development is supported by a distributed neural network that synthesizes multiple sources of information in real time in social interactions and, though not necessarily species specific, contributes to a unique constellation of cognitive capacities in human beings (Mundy et al. 2009).

Bruner (1995) characterized a principal approach to research on joint attention as one that is focused on the “epistemological question”. That is to say, much of the research has been concerned with describing what the development of joint attention behaviors reveals about change in the stages of knowledge infants possess about other people’s minds, or reveals about the nature of cognitive modules that may be dedicated to knowledge about people’s minds (e.g., Baron-Cohen, 1995; Tomasello et al., 2005). This “social cognitive” paradigm emphasizes the comparison of the types of joint attention infants are capable of at different ages. These comparisons have provided grounds for numerous experimentally controlled inferences and insights about age-related changes in infant’s knowledge about the intentions of other people (e.g., Baron-Cohen, 1995; Carpenter, Nagell & Tomasello, 1998; Meltzoff & Brooks, 2008). The emphasis on knowledge has become so strong in this paradigm that in some opinions joint attention, per se, is not thought to be achievable until a critical module or stage of social cognitive development is reached. Social cognitive theory argues that this point in development is heralded by signs of “intentional” joint attention inferred from the emergence of infants’ tendency to acknowledge shared experience with the use of spontaneous gaze alternation between objects and people by about 12 to 15 months of age (Baron-Cohen, 1995; Tomasello et al. 2005; see Fig. 1).

A complimentary but alternative approach to the social cognitive paradigm is exemplified by the question of; does joint attention lead to social knowledge acquisition (learning) and the capacity to share knowledge with others, and if so, how? This question was part of Bruner’s (1975) original motivation for the study of joint attention. He thought of joint attention as an early onset interactive key to social learning and, therefore, a potentially informative natural example to emulate in curriculum design (Bruner, 1995).

A critical difference between these two views that joint attention ostensibly follows from and is defined by social knowledge development in the social cognitive paradigm, but is conceived of as leading to knowledge development and sharing information in the alternative paradigm. Our research and theory on joint attention development has been guided by this alternative approach for quite some time. A maxim of this approach is that cognitive and psychological development is not only effectively modeled in terms of relatively discontinuous changes in knowledge or modules, but also in terms of changes and continuities in the speed, efficiency, and combinations of information processing that give rise to knowledge or psychological phenomenon (Hunt, 1999).
The approach also emphasizes a constructivist view of cognitive and neurocognitive development in which infants learn as much about the minds of other people from their own actions in episodes of joint attention, as they do from the perceptions of others’ behavior (Mareschal, Johnson, Sirois, Spratling et al. 2007; Mundy, 2003; Mundy & Newell, 2007; Piaget, 1952). Indeed, this paradigm is associated with a model that describes joint attention development in terms of advances in the parallel processing of three sources of information illustrated in Figure 2. When people interact in joint attention they actively engage in: 1) self-referenced processing of information about actions they generate (e.g. movement and orientation of one’s own eyes or control of visual attention), and about bodily state (e.g. affect and spatial position relative to a referent), 2) processing information about another person’s attention and behavior, as well as, 3) integrating those strands of input with processing information about a commonly referenced object or event (e.g., Mundy, Kasari & Sigman, 1993; Mundy & Newell, 2007; Mundy, Sullivan & Mastergeorge, 2009).

We assume the early development of this type of parallel processing involves a transactional or dynamic interplay between neural, psychological and behavioral development. That is to say, practice with parallel self-referenced, other-referenced, and object/event-referenced information processing during joint attention in infancy is both a consequence and an epigenetic organizational impetus related to the establishment of distributed neural processing network that involves the flow of information across distal frontal, temporal and parietal cortical systems (provided below). With maturation and experience, this distributed neural network comes to serve a social-executive function that enables infants to engage in increasingly effortless social-coordination of attention to internal information, external social, and external objects/event information in social interactions. With development the overt operations of socially coordinating visual attention to external references become internalized as cognitive operations that allow for the joint coordination of mental attention to common representations. This internalization of joint attention process contributes to the neurocognitive foundation for social cognition, symbolic thought, and self awareness (Mundy et al. 2009). We refer to this as the parallel and distributed processing model (PDPM) of joint attention development. The following provides a more details related to this synopsis of this model and a review of research on the distributed neural networks involved in joint attention development.

THE DEVELOPMENT AND DISSOCIATION OF JOINT ATTENTION

There are two main functional categories of joint-attention behaviors in infancy. Responding to joint attention (RJA) refers to measures of infants’ ability to follow the direction of the gaze and gestures of others in order to share a common point of reference. RJA functions as an automatic reaction to the potential that others’ gaze signifies an important source of information in the environment. Alternatively, initiating joint attention (IJA) involves infants’ generation of gestures and eye contact to direct others’ attention to objects, to events, and to themselves. The prototypical function of IJA is to “show” or spontaneously seek to share interests or pleasurable experience with others (Fig. 2). Just as speech comprehension and expression dissociate in early language development (Bates, 1993) one important characteristic of the ontogeny of human joint attention is that IJA and RJA also dissociate in development.

Several independent studies indicate that the cross dimension correlations between the frequency or consistency of IJA and RJA (i.e. internal consistency of measures of a common construct) are weak to non-significant concurrently and predictively in samples of 9, 12, 15 and 18-month-old infants (e.g. Meltzoff & Brooks, 2008; Mundy et al., 2007; Sheinkopf et al., 2004; Slaughter & McDonald, 2003). In contrast, within dimension correlations of the frequency of infant IJA and RJA across age (i.e., test-retest reliability) has been shown to be significant in a large sample of infants from 9 to 12, 12 to 15, and 15 to 18 months, and even 9 to 18 months for IJA (Mundy et al. 2007). Moreover, although both infant IJA and RJA
measures correlate with or predict learning and cognitive development, these measures also exhibit different patterns of correlations with childhood IQ (Ulvund & Smith, 1998), frontal brain activity (Caplan, Chugani, Messa, Guthrie, et al., 1993; Mundy et al., 2000); reward based behavioral goal-inhibition and self-monitoring behaviors (Nichols et al., 2005), attention related self-regulation (Morales et al., 2005), and attachment (Claussen et al., 2002).

Results from clinical and comparative research also provided evidence for the dissociation of IJA and RJA development. Indeed, this dissociation may have first been noted in research on autism. Both RJA and IJA are useful in the early identification and diagnosis of autism childhood (e.g. Lord, et al., 2000; Stone, Counrod & Ousley, 1997). RJA impairments, however, become less evident among children with autism after they achieve a cognitive developmental level of approximately 30 to 36 months (Mundy et al., 1994). Subsequent research has indicated that there is inconsistent evidence of a robust syndrome specific impairment in the ability to process gaze or respond to joint attention in school age children or older people with autism (Nation & Penny, 2008). On the other hand, IJA deficits are observed in children with autism from preschool through adolescence (e.g. Charman, 2004; Dawson et al., 2004; Hobson & Hobson, 2007; Mundy et al., 1986; Sigman & Ruskin, 1999).

The correlates of IJA and RJA also diverge as much as they converge in studies of autism. Both IJA and RJA are related to executive inhibition and language development in autism, although RJA may be the more robust correlate of language development (Bono et al., 2004; Dawson et al., 2002, 2004; Griffith et al., 1999; Sigman & McGovern, 2005). However, to our knowledge, only IJA has been observed to be significantly associated with differences in social and affective symptom presentation in autism (Charman 2004; Kasari et al., 1990; Kasari et al., 2007; Lord et al., 2003; Mundy et al., 1994; Naber et al., 2008; Sigman & Ruskin, 1999).

Comparative research also suggests that there are important distinctions to be understood between IJA and RJA development. Chimpanzees reportedly show the capacity for RJA. Indeed, gaze tracking is common to a variety of animals and even observed in the neonatal period of development (Jaime, Lopez, & Lickliter, 2009). However, animals and primates other than humans display little evidence of IJA or spontaneous attempts to convey information singularly for the social sharing or experiences with members of their own species (Tomasello & Carpenter, 2005). This suggests that basic information processing systems that enable adaptive responding to the perception of gaze direction as manifested in RJA might be common to many animals. However, a second processing system involved the spontaneous generation of a joint attention to initiate experience sharing with others may be a unique and defining facet of human nature (Tomasello & Carpenter, 2005).

Why do different types of joint attention behaviors appear to be dissociated in typical development, in neurodevelopmental disorders such as autism, and phylogenetically across human and non-human primates? One plausible hypothesis that was first raised several years ago is that IJA and RJA reflect overlapping but substantially different types of social information processing that occur in both shared and unique neural systems (Mundy, 1995;
The Two Neural Systems of Joint Attention and Social Cognition

Electroencephalography and Positron Tomography data from studies of infants have indicated that during early development IJA is associated with frontal-cortical activity (Caplan et al., 1993; Henderson et al., 2002; Mundy et al., 2000; Torkildsen et al., 2008), while RJA and related behaviors are more closely tied to parietal and temporal cortical processes (e.g. Frieschen, Bayliss, & Tipper, 2007; Materna, Dicke & Thern, 2008; Mundy et al., 2000). One interpretation of these data (Mundy et al. 2000) is that early in development IJA and RJA respectively involve functions of the anterior and posterior cortical attention networks that have been described by Michael Posner and others (Posner & Peterson, 1990; Posner & Rothbart, 2007). The functions of the posterior attention network are common to many primates, but the anterior attention network is not well represented in primates other than humans (Astafiev, Shulman, Stanley & Snyder et al., 2003; Emery, 2000; Jellema, Baker, Wicker, & Perrett, 2000).

RJA appears to be most closely associated with the posterior system, which regulates relatively rapid, involuntary attention orienting to perceived objects/events. The posterior attention system begins to develop in the first 3 months of life, and prioritizes orienting to biologically meaningful stimuli (Posner & Rothbart; 2007). It is supported by neural networks of the parietal/precuneous and superior temporal cortices (Figure 3). These neural networks are active in the perception of the eye and head orientations of others, as well as the perception of spatial relations between self, other and the environment (Frieschen, Bayliss, & Tipper, 2007). In particular, the recent work of Materna, Dieke, & Their (2009) suggests that the posterior part of the STS region and adjacent regions of the precuneous are specifically involved in extracting and using detailed directional information from the eyes of another person to redirect one’s own gaze and establish joint attention. On the other hand regions of the intraparietal sulcus may contribute more generic encoding of spatial direction and mediate shifts of spatial attention independent of the type of cue that triggers this process (Materna et al. 2009). The posterior system is also involved in the development of cognitive representations about the world built from information acquired through external senses (Dosenbach et al., 2007; Fuster, 2006; Cavana & Trimble, 2006).

Processes associated with initiating joint attention may reasonably be expected to be supported by the later developing anterior network involved in the regulation of more volitional, self-initiated, goal directed attention deployment. The anterior network includes elements of the anterior cingulate, medial superior frontal cortex including the frontal eye fields, anterior prefrontal cortex and orbital frontal cortex (e.g. Dosenbach et al., 2007; Fuster, 2006). The development of the intentional control of visual attention begins at about 3 to 4 months of age, when a pathway from the frontal eye fields (BA 8/9) that releases the superior colliculus from inhibition begins to be actively involved in the prospective control of saccades and visual attention (Canfield & Kirkham, 2001; Johnson, 1990). The function of this pathway may underlie four-month-old infants’ ability to suppress automatic visual saccades in order to respond to a second, more attractive stimulus (Johnson, 1995), and six month olds’ ability to respond to a peripheral target when central, competing stimuli are present (Atkinson, Hood, Wattam-Bell, & Braddick, 1992). We assume that the functions of this pathway also enable the type of IJA behavior that involves intentional gaze alternation between interesting events and social partners (Mundy, 2003, see Fig. 1).

In summary, volitional or intentional control of attention is one of the primary features thought to distinguish IJA and RJA development. We are relatively compelled to engage RJA to look where others look, but we choose to spontaneously engage in IJA to affect others attention to
share experience. This conceptualization lead to a corollary assumption that variability in the
tendency to engage in IJA behaviors among individuals or groups may involve motivation
processes and sensitivity to the reward value of social engagement (Dawson et al. 2004; Mundy,
1995). One specific suggestion is that frontal networks involved the Behavioral Activation
System (BAS) and reward sensitive activation of appetitive, approach behavior may be
involved to a greater degree in IJA and RJA (Mundy, 1995). The BAS is thought to involve
networks of the ventromedial frontal cortex, striatum, and amygdala (Davidson & Irwin,
1999). In a related vein, Tomasello et al. (2005) suggested that constitutional processes
associated with a species-unique motivation to share emotions, experience, and activities with
other persons may critical to IJA and part of what differentiate human and other primate
communicative and social cognitive development. A third idea is that frontal mediation of
motivation is thought to be a vital component of resolving conflict between goal related
behavior alternative. Research on the neural networks involved in the latter suggest that
functions of the anterior cingulate may play a vital role IJA development and individual
differences (Mundy, 2003).

Theoretically, the differences in the functions, developmental timing, and the neural networks
involved in the anterior and posterior attention systems outlined in the proceeding paragraphs
contribute to the developmental dissociation of IJA and RJA in typical, atypical and primate
development (Mundy et al., 2000, 2007; Mundy & Newell, 2007). However, although IJA and
RJA appear to follow distinct bio-behavioral paths of development it is also likely that they
integrate in development. For example, although predominantly associated with frontal activity
EEG data also indicates that activation of a distributed anterior and posterior cortical system
at 14 months predicts IJA development at 18 months in infancy (Henderson et al., 2002). Hence,
what we observe as the neural correlates of joint attention in infants may be only a part of the
picture of the mature, fully integrated mature human joint attention network. Two recent fMRI
studies have significantly advanced our understanding of the distributed anterior-posterior
cortical network that is associated with joint attention processes in adults.

Williams, Waiter, Perra, Perrett, et al. (2005) presented 12 right handed male participants with
congruent and incongruent responding to joint attention stimuli during fMRI data acquisition.
The task for the participants was to track a red ball that appeared and then disappeared in
different regions of the perimeter of a video screen. A video image of an adult male face was
visible in the middle of the screen. On Congruent RJA trials the visual gaze of the male stimulus
face shifted to the region of the screen prior to and predicting the correct location of the
appearance of the red ball. In the Incongruent RJA trials the visual gaze of the male stimulus
face shifted to a region of the screen that did not predict the correct location of the appearance
of the ball. In the former gaze tracking benefited correct and rapid responding and in the second
condition inhibition of gaze tracking benefited optimal responding. The methods and results
of this study are depicted in Figure 4. Contrasts associated with task performance in the
Congruent and Incongruent RJA revealed neural activation across a neural network that
included the caudate, medial frontal cortex (BA 9,10), cingulate and anterior cingulate and
parietal- precuneous cortex (Williams et al. 2005).

In another study Schilbach, Wilms, Eickhoff, Romanzetti et al. (in press) employed an
innovative paradigm to test the hypothesis that initiating and responding to joint attention
reflected activation of partially dissociated or independent neural networks. In particular, they
tested the hypothesis that initiating of joint attention recruited activation of neural systems
involved in motivation and reward to a greater extent than did responding to joint attention
(Mundy, 1995; Mundy & Newell, 2007). To do so Schilbach’s research team collected 3 Telsa
fMRI data from 21 participants who were presented with a task with four conditions.
During neuroimaging trials participants were instructed to look at one of three objects that appeared on a video monitor. A face created by virtual reality methods also appeared at the center of the video stimuli. On some trials the gaze of the virtual face was made to be responsive to the participant’s gaze. The gaze of the virtual face either followed the participants correctly in the joint attention condition (JA) emulating the experience of successfully initiating joint attention, or looked elsewhere in the non-joint attention condition (NOJA) emulating the experience of unsuccessfully initiating joint attention. On other trials participants experienced following the gaze of the virtual other person to one of three objects to emulate the experience of Responding to Joint Attention (RJA). On some trials the gaze of the virtual other person predicted the correct region to look for the object (joint attention condition, JA) and on some trials the gaze of the virtual other person failed to predict the correct region to look for objects (non-joint attention condition, NOJA). This resulted in a 2×2 factorial design comparing IJA and RJA and JA and NOJA conditions.

The a-priori comparison of IJA-JA + RJA-JA > IJA-NOJA + RJA-NOJA indicated a unique pattern of activation for joint attention in a distributed neural network involving the regions of the dorsal and ventral medial prefrontal cortex, medial orbital frontal cortex, posterior cingulate cortex, subgenual cingulate and ventral striatum, right hippocampus and anterior temporal cortex (Figure 5). The interaction of IJA, RJA, and condition such that (IJA-JA > RJA-JA) > (IJA-NOJA > RJA-NOJA) revealed activity in a network of regions in the inferior and superior parietal cortex, middle cingulate cortex and insular cortex specific to IJA-JA versus all other behaviors. RJA-JA versus all other behaviors was associated with specific recruitment of the left superior frontal lobe. Finally, IJA-JA versus RJA-JA revealed unique activation for IJA in the ventral striatum bilaterally for Initiating Joint Attention and activation of the ventral medial prefrontal cortex for Responding to Joint Attention.

Schilbach et al. (in press) provide a cogent interpretation of their seminal data. Among the points they innumerate three stand out in terms of their relevance to hypotheses related to the PDP model of joint attention. First their data are consistent with the hypothesis that joint attention involves a distally distributed neural network of frontal temporal and parietal cortical and subcortical neural activity (Mundy, 2003). Second, RJA and IJA reflect functions of partially independent sub-units of this neural network, and these may be differentiated on the basis of the latter’s association with motivation, striatal, dopaminergic functions (Mundy, 1995; 2003). Finally, joint attention may involve self referenced processing and the cortical systems that support proprioceptive and interoceptive, such as the insular cortex (Mundy, Gwaltney & Sullivan, in press). We now turn to a more detailed consideration of this idea.

Self-Referenced Processing and Joint Attention

Recall that joint attention involves a tripartite deployment of attention and triadic information processing (Fig. 2). When we engage in joint attention we attend to and process information about: 1) an object or event, 2) another person’s attention and behavior related to the object, and 3) self-reference information about our own attention to, and experience of the object and the situation.

Self-referenced processing refers to implicit, subjective, and pre-reflective processing and integration of information from one’s own body (e.g. heart rate, volitional muscle movement) with perceptual and cognitive activity such as representations in working or long term memory (Northoff et al., 2006). It involves at least two types of information. Interoception describes sensitivity to physiological information originating within the body, such as heart rate, respiration, autonomic arousal, respiration, and emotional states. Proprioception involves sensitivity to the position, location, orientation, and movement of the body. Research suggests that specific, distributed neural systems may be involved in processing interoceptive information (anterior insula) and proprioceptive information about movement and orientation.
(anterior cingulate and parietal cortices). Robust connectivity between these associative cortical systems also likely allows for a fundamental integration of these two dimensions of self-referenced information (Craig, 2009; Balslev & Miall, 2008; Uddin & Menon, 2009). Indeed, comparative research suggests that interoception and its integration with proprioception may be more elaborated in humans via Von Economo neurons of the anterior cingulate than in other primates and mammals (Allman, Watson, Tetreault, & Hakeem, 2005; Craig, 2003).

In the past discussions, the first author naively described self-referenced information processing in joint attention singularly in terms of the proprioceptive functions that were attributed to functions of medial frontal structures, such as the anterior cingulate (Mundy, 2003). This was an important oversight because it conflated interoceptive self-referenced processing with goal related proprioceptive processing. Recently we have attempted to bring greater clarity to putative nature of self-referenced processing in joint attention (Mundy, et al. in press). The dual nature of self-referenced information processing in joint attention may be better described in terms of: a) proprioception, such as feedback from ocular muscle control and the vestibular system related to the spatial direction of one’s own visual attention and head posture (see Butterworth & Jarrett, 1991 for related discussion), and b) interoception including but not limited to information about arousal and the positive (rewarding), neutral or negative valence of self perception of the object or event, as well as the valence of sharing attention with a social partner. The latter conjecture now finds some substantiation in the observations of insular cortical activity associated with adult’s experience of initiating joint attention (Schilbach et al. In press).

The capacity to meet the demands of multiple target, self-other-object attention deployment and information processing in joint attention is only gradually acquired. Facility with the demands of social attention coordination emerges in infancy as a function of the interaction between frequent, incremental practice with joint attention behaviors (e.g. gaze following or gaze alternating, Fig. 1) and neurocognitive maturation. Daily if not hourly practice with joint attention typically begins by three months of age and continues through at least the second year (D’Entremont et al., 1997; Carpenter et al., 1998).

Hypothetically practice with management of multiple sources of information processing in joint attention contributes to the developmental articulation of self- and other-referenced neural processing networks, which aid the psychological differentiation of self from other in early development (c.f. Decety & Sommerville, 2003; Northoff et al., 2006). This hypothesis is consistent with a school of thought that human self-awareness does not simply rest upon advances in self-reference proprioception and interoception. Rather, self-awareness is an incrementally achieved property of social relational processing that involves the analogous, parallel processing of self- and other-referenced information (Chen, Boucher, & Parker Tapias, 2006; Piaget, 1952; Vygotsky, 1962). Typical human conscious self-awareness may not arise in social isolation because its development requires a deep and prolonged synthesis of self-referenced information processing (proprioceptive/interoceptive) and other referenced information processing (e.g. Piaget, 1952). Neurocognitive research suggests that this synthesis occurs by way of the repeated social experiences of infants interacting with the maturation of a distributed frontal, temporal and parietal cortical system (Decety & Sommerville, 2003; Keysers & Perret, 2004).

Keysers and Perret (2004), for example, propose that the relational processing of self-referenced and other-referenced information in infancy may be thought of in terms of a form of Hebbian learning that facilitates social-cognitive development. Hebb (1949) proposed that learning is supported by a basic neural function. Neural networks that are repeatedly active at the same time become associated, such that specific activity [e.g. re-presentations] in one
network triggers activity in the other (Hebb, 1949). Keysers and Perrett suggest that co-activation of neural networks for processing self-generated information and information about people is fundamental to the development of representations and knowledge about self and others. We are inclined to agree with Hebbian model proposed by Keysers & Perrett, but would add that rostral-medial-frontal processing of information about self-produced visual attention and integrated with posterior processing of the attention of others is primary to the Hebbian mapping involved in social-cognitive development.

In this regard we conceptualize the development of joint attention in terms of an executive function. One definition of executive functions is that they involve the transmission of bias signals throughout neural networks to selectively inhibit comparatively automatic behavioral responses in favor of more volitional, planned and goal-directed ideation and action in problem solving contexts (Miller & Cohen, 2001). The aggregate effect of these bias signals is to guide the flow of neural activity along pathways that establish the proper mappings between inputs, internal states, and outputs needed to perform a given task more efficiently (Miller & Cohen, 2001).

With this definition in mind we propose that the incremental practice of joint attention in approximately the 3- to 15-month period is both a cause and an experience-expectant organizer of the emergence of frontal bias signals. These signals facilitate adaptive mappings across outside-in posterior cortical {temporal-precuneous} processing of exteroceptive information about the attention behaviors of other people, and the inside-out medial-frontal and insular-cortical proprioceptive and interoceptive processing of internal states, as well as information related to the intentional control of visual attention. This mapping results in the integrated development of a distributed anterior and posterior cortical joint attention system (Fig. 3; Mundy, 2003; Mundy et al., 2009).

In our view, the early establishment of this cortical mapping of a joint processing system is formative with respect to the development of the shared neural network of representations of self and others that have been described as driving the developmental progression of self-awareness social-cognition in infancy through adulthood (e.g. Decety & Grezes, 2006; Keysers & Perrett, 2004; Vanderwal, Hunyadi, Grupe, Conners, & Schultz, 2008). That is to say the Hebbian mapping associated with social-cognition which Keysers and Perrett(2004) have so lucidly described begins, with the integrated anterior cortical processing of information about self-produced visual attention and posterior cortical processing of the visual attention of others (Mundy et al., 2009). With sufficient practice and experience, the capacity for the joint processing of self-other attention information begins to consolidate and require less mental effort. As basic joint attention processes are mastered and require less effort it becomes a social executive function that contributes to the development and efficiency of social learning, symbolic thinking and social-cognitive problem solving and development. How this model of neural network mapping helps us to better understand the connection between and joint attention and development each of these domains will next be considered in turn.

Joint Attention and Social Learning

During language learning how do infants correctly associate their parents’ vocal labels to the correct distal object or event in the midst of an environment flush with a myriad of potential referents? Baldwin (1995) suggested that they use RJA and the direction of gaze of their parent to guide them to the correct area of the distal environment, thereby reducing the potential for ‘referential’ mapping errors. Infants’ can also use IJA to denote something of immediate interest. This assists parents to provide new information in context when the child’s interest and attention is optimal for learning and the reduction of referential mapping errors learning (Tomasello & Farrar, 1986). Hence, joint attention may be conceived of as a self-organizing system that facilitates information processing in support of social-learning (Mundy, 2003).
This ‘learning function’ is fundamental to the nature of joint attention (Bruner, 1975) and continues to operate throughout our lives (e.g., Bayliss et al., 2006; Nathan et al., 2007). For example, imagine the school readiness problems of a five-year-old who enters kindergarten but is not facile with, and/or sufficiently motivated to coordinate attention with the teacher.

Another viewpoint on the links between joint attention and early learning is provided by the recent research of Striano and her colleagues (Hirotani, Stets, Striano, & Friedman, 2009; Striano, Chen, Cleveland, & Bradshaw, 2006a; Striano, Reid & Hoel, 2006b). Striano et al. (2006a,b) compared the visual recognition memory of 9-month-old infants who studied pictures in one of two conditions. In the active joint attention condition each infant’s mother made eye contact with their child and then turned to look at the pictures their infant studied. In passive joint attention condition the mother’s did not make eye contact but simply looked at pictures in parallel with their infants. The results indicated that the 9-month-olds in the active joint attention condition displayed EEG evidence of enhanced neural activity associated with greater depth of processing and behavioral evidence of better recognition memory. More recently Hirotani et al. 2009 report EEG evidence of widespread frontal, central, parietal neural network activity and better depth of processing among 18 month olds during a word learning task in joint attention versus non-joint attention conditions. The authors note that Joint attention may not only serve to referentially map new word object associations but somehow also enforce the relation between the two.

One mechanism of enhanced depth of processing during joint attention may involve a role of neural networks in learning. Parallel and distributed cognitive theory suggests that depth of encoding is optimized by the simultaneous activation of multiple neural networks during information processing (e.g., Munakata & McClelland, 2003; Otten et al., 2001). Joint attention, we assume, engages the activation of a distributed neural network comprised of nodes involved in the frontal processing of self-referenced attention related information and posterior cortical processing of information related to the attention of another person (see Figure 6). Research suggests that activation of such a distributed network occurs in infants (Henderson et al. 2002) just as it does in adults (e.g. Schilbach et al. in press). Therefore, it is plausible that processing information during joint attention enhances learning by embedding semantic or lexical encoding in association with the parallel activation of a distributed, neural network engaged in processing of information pertaining to the attention of self and the attention of others. This suggests that neural network activation during joint attention may be one principal mechanism of our advanced human capacity for social-learning.

**Joint Attention and Symbolic Thinking**

Tomasello et al., (2005) have provided perhaps the most well articulated model of joint attention development in terms of three stages of what infants know about other people. In the first Understanding Animate Action stage, 3- to 8-month-old infants can perceive contingencies between their own animate actions and emotions relative to the animate actions and affect of others. However, they cannot represent the internal mental goals of others that are associated with these actions. In the next Understanding of Pursuit of Goals stage 9-month-olds become capable of shared action and attention on objects (e.g. building a block tower with parents). Tomasello et al., 2005 suggest this involves joint perception, rather than joint attention, because the social-cognitive capacity to represent others internal mental representations necessary for true joint attention is not yet available. The latter emerges between 12-15 months in the Understanding Choice of Plans stage. This stage is heralded when infants become truly active in initiating episodes of joint engagement by alternating their eye-contact between interesting sights and caregivers (Tomasello et al., 2005, also see Fig. 1c1,2,3). This shift to active alternating gaze indicates infants’ appreciation that others make mental choices about alternative actions that affect their attention. Infants also now know themselves as agents that
initiate collaborative activity based on their own goals. Hence, the development of “true” joint attention at this stage is revealed in the capacity to adopt two perspectives analogous to speaker-listener.

The capacity to adopt two perspectives is also assumed to be an intrinsic characteristic of symbolic representations. In this regard, Tomasello et al., (2005) raise a seminal hypothesis that symbolic thought is a transformation of joint attention. They argue that symbols themselves serve to socially coordinate attention so that the intentions of the listener align with those of the speaker. In other words, linguistic symbols both lead to and are dependent upon the efficient social coordination of covert mental attention to common abstract representations among people. This hypothesis fits well with the parallel and distributed processing model of joint attention, but the perspective of the PDP MODEL places it in a substantially different developmental and neural networks framework.

As previously noted the notion that true joint attention does not emerge until requisite social-cognitive awareness or modules emerges in the latter part of the first year (Baron-Cohen, 1995; Tomasello et al., 2005) is not germane to the PDP MODEL. Rather, the PDP Model holds that joint attention is defined in terms of the multiple level attention deployment and social information processing which begins to be practiced by infants by three to four months of age (D’Entremont, Hains, & Muir, 1997; Farroni, Masssuccessi, & Francesca, 2002; Hood, Willen, & Driver, 1998, Morales et al., 1998; Striano & Reed, 2006; Striano & Stahl, 2005). Indeed, even the types of active alternating gaze thought to mark the onset of true joint at 12-15 months (Tomasello et al., 2005) develops no later than at 8-9 months of life, and quite possibly earlier (Mundy et al., 2007; Venezia et al., 2004).

The PDP Model also includes the proposition that with practice, the operations of joint attention to external objects or events become internalized to form cognitive operations that enable us to socially coordinate mental attention to representations (Mundy et al. 2009). This ability to socially coordinate mental attention across people to common cognitive representations is essential to the development and use of symbolic thinking throughout the life span (Werner & Kaplan, 1963). Thus, in addition to conceiving of linguistic symbols as enabling the social coordination of covert attention to common mental representations across people (Tomasello, et al. 2005), it is equally plausible to conceive of internalized joint attention processes as the basis of social coordination of covert attention to common mental representations that enables the symbolic development. Thus, the activation of the distributed joint attention neural network always occurs in the course of initial symbolic development, and in subsequent symbolic thinking throughout the life span. It is possible to have joint attention without symbolic thinking, but the converse seems unlikely.

Only indirect support exists for this hypothesis at this time. Theory and data suggest that symbolic representations are often initially encoded during the joint processing of information about the overt attention of self and of others directed toward some third object or event (Adamson et al., 2004; Baldwin, 1995; Kasari et al., 2007; Wener & Kaplan, 1963). Experimental studies of intervention with autism suggest that improvement of joint attention development has a positive effect on symbolic behavior development and visa versa (Kasari et al. 2007).

A most useful test of the joint attention hypothesis of symbolic thought would be to use imagining methods to examine the expectation that there is the expected overlap between network brain activity associated with symbol acquisition or symbolic thought process, and the network brain activity associated with joint attention. Although such a comparison has yet to be conducted two studies do bear on this prediction. German et al. (2004) and Whitehead et al. (2009) have reported that observations of the pretend rather than conventional object (e.g.
tennis racket as guitar) elicits activation in a distributed neural network involving systems of
the medial prefrontal areas (Brodmann’s areas [BA] 9/632, 9, and 10), inferior frontal gyrus
(BA 44, 47), temporo-parietal regions (BA 21 and 22), and parahippocampal areas, including
the amygdale (German et al. 2004; Whitehead et al. 2009). These studies were interpreted to
suggest that there is a strong association between neural networks that support symbolic pretend
cognitive functions and “mentalizing” social cognitive processes (e.g. Whitehead et al.
2009). We would suggest though that it is equally possible that any such association is at least
partially mediated by the role the neural network for internalized joint attention processes plays
in both symbolic processes and social cognition (Mundy et al. 2009; Tomasello et al. 2005).
Let us finally consider the latter in more detail.

The Joint Attention Foundations of Social Cognition

In considering how the PDP model represents the contribution of joint attention to social
cognitive development it is useful to recall that PDP perspective gives equal footing to the
significance of infants’ development of their own intentional visual behavior in modeling joint
attention and social cognitive development (Mundy et al., 1993). The assumption here is that
neonates and young infants receive greater quantities and fidelity of information about self-
intended actions, such as active control of visual attention, through proprioception and
interception than they receive about other’s intended actions through exteroceptive
information processing. Thus, infants have the opportunity to learn as much or more about
intentionality from their own actions as from observing the actions of others. The development
of joint attention and social cognition are constructivist process that involves self-perception
as a foundation for the attribution of meaning to the perception of others behaviors. We have
referred to this as the “inside out” processing assumption of the PDP model (Mundy &
Vaughan, 2008).

Accordingly, the PDP Model suggests that, while IJA and RJA reflect distinct functions in
infancy, the interaction of information processing associated with each type of joint attention
plays is vital to social-cognitive development. As previously noted the dissociative but
interactive development of IJA and RJA may be similar to the dissociative but interactive paths
of early receptive and expressive language development. The former is supported in part by
posterior cortical systems of the temporal and (Wernicke’s area) as well as parietal cortex, and
the latter by dorsolateral frontal systems of Broca’s area. The distinct nature of these neural
systems may help to explain why receptive and expressive language expression is not
necessarily highly correlated in the early period of language development (Bates, 1993).
However, the distributed neural networks for language interact via massive connectivity
supported by associative fibers (arcuate fasciculus). Theoretically, more optimal language
related development is associated with greater coherent activity across the distributed neural
network systems of expressive and receptive language development (Mundy et al. 2003;
Salmelin & Kujala, 2006).

Similarly the manifest effect of joint attention on social cognition can be conceptualized in
terms of a synthesis of information processing output from the distributed neural networks for
RJA and IJA (Mundy & Newell, 2007). With regard to the former Jellema, Baker, Wicker &
Perrett (2000) epitomized the goal related neurocognitive output for the RJA system in primates
terms of the “concept” that that “where others’ eyes go, their behavior follows”. Mundy and
Newell (2007) went on to suggest that, if part of the cognitive output from the posterior RJA
system can be characterized as “where others’ eyes go, their behavior follows”, then a parallel
part of the cognitive output of the anterior system IJA could be characterized as “where my
eye’s go, my [own] behaviors follows.”

The interaction and juxtaposition of information and concept formation from these two systems
is posited to serve as an important engine of cognitive development. It enhances the

Neural Netw. Author manuscript; available in PMC 2011 October 1.
developmental differentiation of awareness of self-agency with respect to the control of
attention versus others’ agency and attention control. Moreover, the same integration of the
anterior and posterior attention system that enables the comparative monitoring of overt
attention (looking behavior) ultimately plays the same role when individuals begin to be able
to attend to internal representations about the behavior of self and representations of the
behavior of others. They (we) can begin to impute the conditional proposition that if self
referenced-control of attention relates to one’s own goal-directed behavior representations then
goal related behavior in others must follow from their self-control of attention (see top “thought
balloon” in Fig. 2). As this integrative anterior–posterior capacity to attend to overt or covert
information about self and others matures with advances in representational abilities, a fully
functional, adaptive, human social-cognitive system emerges.

Another perspective on the developmental outcome of the integration of IJA and RJA is
suggested by the description Keysers and Perrett (2004) provide of the role of Hebbian learning
in social-cognitive development. Neural networks that are repeatedly active at the same time
become associated, such that activity [e.g. “re-presentations”] in one network triggers activity
in the other (Hebb, 1949). Keysers and Perrett suggest that frequent early experience with the
common activation of neural networks for processing self-generated information while
processing information about conspecifics leads to an automatic reciprocal activation of self
referenced information when thinking about others later in development. They suggest this
reciprocal Hebbian activation of information learned about self and other is fundamental to the
capacity to understand the actions and intentions of others. It seems likely that the longstanding
notions of Hebbian learning process invoked by Keysers and Perrettis have much in common
the new labels for related constructs such as simulation (Gordon, 1986) and mirror neurons
(i.e., Decety & Summerville, 2003; Williams, 2008) that are commonly invoked in current
models of social-cognitive development.

Recall that the PDP MODEL accommodates these interrelated ideas/constructs and suggests
that one fundamental domain of relevant Hebbian mapping for social-cognitive development
begins with integrated frontal processing of information about self produced visual attention
and posterior processing of the attention of others. The PDP also suggests that once well
practiced the joint processing of attention information requires less mental effort. Once the
mental effort to engage in joint attention decreases, it becomes possible for infants to learn
from combined self and other referenced information processing in new ways. This is a critical
assumption of the PDP assumes and referred to as the “learning to do” and “learning from”
phases of joint attention (Mundy & Vaughan, 2008).

A useful analogy to consider here may be learning to ride a bicycle. At first the demands of
coordinating the processing of motor movement information, vestibular information and visual
information are so high that it is difficult to attend to one’s surrounding when first learning to
ride. However, once some mastery of coordinating information processing and motor output
is achieved we begin to experience the mental capacity to focus our attention on our
surroundings and enjoy, if not learn from the information about the world that flows by as we
cycle. We imagine it’s much the same course of development for infants with respect to joint
attention. At first the process is difficult to manage, but with practice effort decreases and most
infants begin to be able to process and learn from the rich array of social comparative
information available during joint attention (Mundy et al. 1993).

Recall the assertion that the co-active system of joint attention begins synergize as frontal
executive bias signals (Miller & Cohen, 2001) that increasingly enable attending to multiple
sources of information during infancy. Accordingly, the “learning to” phase of joint attention
development may be thought of as reflecting the emergence of frontal bias signals that establish
the proper mappings across: 1) outside-in posterior cortical {temporal-precuneous} processing
of inputs about the attention behaviors of other people and, 2) inside-out rostral-medial frontal cortical \{dorsal frontal medial cortex (BA 8-9), anterior cingulate, insular cortex\} processing of internal states and outputs related to active vision. We imagine this mapping is influence by Hebbian learning processes (Keysers & Perrett, 2004), genetic and especially activity dependent genetic processes (Mundy et al. 2009) and environmental processes (e.g. Claussen et al. 2002). The mapping ultimately results in the [more or less] functional, integrated development of a distributed anterior and posterior cortical joint attention system.

As basic joint attention process is mastered and becomes an executive function its ‘effort to engage’ goes down. As effort decreases joint attention development has the potential to shift to a “learning from” phase of development. In the learning from phase the capacity to attend to and learn from the integration of multiple sources of information in “triadic attention” deployment becomes more common (Scaife & Bruner, 1975). Triadic attention contexts provide infants with rich opportunities to compare information gleaned through processing internal states associated with volitional visual attention deployment and the processing of the visual attention of others in reference to a common third object or event. Through something akin to cognitive simulation (Gordon, 1986) infants may begin to impute that others have intentional control over their looking behavior that is similar to their own.

The role of simulation in the learning from phase of joint attention development is powerfully illustrated by a recent sequence of elegant experimental studies (e.g. Brooks & Meltzoff, 2002, Meltzoff & Brooks, 2008). Twelve-month-olds often follow the head turn and “gaze direction” of testers, even if their eyes are closed. After 12 months, though, infants discriminate and only follow the head turn and gaze of testers whose eyes are open. This suggests that infants’ understanding of the meaning of the eye gaze of others may improve after about 12 months leading many older infants to inhibit looking in the “eyes closed” condition (Brooks & Meltzoff, 2002).

To examine this interpretation Meltzoff and Brooks (2008) manipulated the information processing of infants in joint attention. They provided 12-month-old infants with the experience of blindfolds that occluded their own looking behavior. After gaining that experience, 12-month-olds acted like older infants in so far as they did not follow the head-turn of blindfolded testers, but did follow the head-turn and gaze of non-blindfolded testers. Meltzoff and Brooks also cleverly provided 18-month-olds with experience with blindfolds that looked opaque, but were actually transparent when worn. After this experience in this condition the 18-month-olds reverted to following the gaze of blindfolded social partners. These data strongly suggest that the infants demonstrated inside-out learning and constructed social-cognitive awareness about others’ gaze based on the experience of the effects of blindfolds on their own active vision.

Lest we fall into the constraints of a stage model we would like to emphasize that the learning to and learning from phases of joint attention development are conceptualized as overlapping rather than sequential. Although more of one than the other may occur at a particular age from 3 months on, examples of both are likely to be observed across many ages. For example, although Meltzoff and Brooks (2008) have provided a seminal example of learning from joint attention in the second year of life the PDP Model holds that some learning from joint attention likely has its onset early in the first year.

One of the first and most vitally informative types of constructivist actions infants take involves the self-control of their looking behaviors, or “active vision”. The science of vision has moved away from the study of “seeing” or passive visual perception, to the study of “looking” or intentional, active-vision and attention deployment (Findlay & Gilchrist, 2003). Active vision in infancy begins to develop at 3-4 months (e.g. Canfield & Kirkham, 2001; Johnson, 1990; 1995).
The human eye is thought to have unique spatially signifying characteristics, such as highlighted contrast between the dark coloration of the pupil and iris versus the light to white coloration of the sclera. This has lead to the suggests that human eyes that provide a potent source of information to infants about other people’s attention that promotes social cognitive development (e.g., Baron-Cohen, 1995; Tomasello, Hare, Lehman, & Call, 2006). However, the same spatial characteristics of eye that provide information about other people to infants also likely allows the saccades of infants to be readily observed by other people and, consequently, act as elicitors of contingent social feedback. When infants shift attention to an object their parents may pick-up and show them the object. When infants shift attention to their parents’ eyes they may also receive a vocal or physical parental response. Thus, just as the characteristics of eyes make it easier for infants to perceive the attention of others, the signal value of eyes makes the active control of vision a likely nexus of information about infants’ own developing intentional agency. Active vision is one of the first domains that enable infants to experience goal directed selection of information to attend to, and a behavior that elicit contingent social behavior responses from other people, such as parental smiles, vocalizations or gaze shifts. It also is one of the first types of volitional actions that infants use to control stimulation in order to self-regulate arousal and affect (Posner & Rothbart, 2007).

The contemporary literature on social-cognitive development often validly but, singularly notes the potential importance of the information infants gather from processing the visual attention of others (e.g. Johnson et al., 2005). It neglects the potential importance of the information infants’ process about their own active vision, and socially contingent responses. The PDP MODEL alternatively also notes the potential for inside-out processing, constructivism and the role of active vision in the development of joint attention and social cognition. Indeed, this aspect of the PDP MODEL offers one explanation for why activation of the frontal eye-fields (a cortical area involved in volitional saccadic control) is a consistent component of the brain activation network observed to correlate with social-cognition in imaging studies (Mundy, 2003). This is because the volitional control of active vision, via the frontal eye fields, may be central to developing an integrated sense of the relations between self-attention and other-attention, which is fundamental to joint attention and social cognition.

The PDP MODEL emphasizes inside-out processing, constructivism and the role of active vision in the development of joint attention. However, it does not maintain that the inside-out processing of self-attention is more important for social-cognitive development than out-side in processing of other’s-attention. This is because, as previously noted the PDP MODEL holds that social meaning, and even conscious self-awareness, cannot be derived from processing either self-attention or other’s-attention in isolation (cf. Decety & Sommerville, 2003; Keyser & Perret, 2004; Vygotsky, 1962). Ontogeny and phylogeny may be best viewed as a dynamic system that, through interactions of multiple factors over time and experience, coalesce into higher order integrations, structures and skills (e.g. Smith & Thelen, 2003). Joint attention, or the joint processing of the attention of self and other, is such a dynamic system. Its pertinence for human development derives from the unique synthesis of the rapid parallel processing of self attention and other attention across distributed neural networks. Consequently, it is not possible to comprehensively account for the role of joint attention in typical or atypical development with research or theory that focuses on only one of its elements in isolation.

Summary and Conclusions

Posner and Rothbart (2007) have cogently revived the notion that neural network models of attention can provide a common, unifying approach to theory and research on many aspects of human cognitive and emotional development. We resonate with this perspective and suggest that the study of joint attention from a neural networks PDP model helps us understand commonalities among developmental processes associated social learning, symbolic thinking,
social cognition and social motivation. In this view joint attention is conceptualized as social executive function of a frontal-parietal neural network, which constitutes a unique and discrete domain of human cognition that has a vital functional impact on social-communicative competence throughout the life span. The joint attention executive function serves to enable the rapid, parallel and comparative processing of information about one’s own attention and the attention of other people in social interactions. This executive function emerges as with active behavioral practice in the first year of life and is both a cause and consequence of the development of a parallel and distributed neural network of primarily frontal and parietal cortical and sub-cortical systems. Once codified this parallel and distributed processing system contributes to the human capacity for increased depth of processing in social learning situations, mental coordination of attention to socially common reference of abstract representations (symbolic thought) and the capacity to mentally represent the attention of self and others. The latter is integral to social cognition. This perspective suggests that we may have as much to learn about human cognition and development through the study of joint attention as an executive attention neutral network, as we do from conceptualizing joint attention as a vital byproduct of development of social-cognitive knowledge.

REFERENCES


Kasari, C.; Freeman, S.; Paparella, T. The UCLA RCT on Play and Joint Attention; Paper presented at the Biennial Conference of the Society for Research on Child Development; Boston, MA. 2007, April;


Neural Netw. Author manuscript; available in PMC 2011 October 1.


Figure 1.
Illustrations of different types of infant social attention coordination behaviors: a) Responding to Joint Attention-RJA involving following another person’s gaze and pointing gesture; b) Initiating Joint Attention-IJA involving a conventional gesture ‘pointing’ to share attention regarding a poster on the wall, c1,2,3) IJA involving alternating eye contact to share attention with respect to a toy, d) Initiating Behavior Request involving pointing to elicit aid in obtaining an out of reach object, and e) Responding to Behavior Requests involving following an adult’s open-palm “give it to me” gesture.
Figure 2.
An illustration of integrated triadic information processing during joint attention. In the lower part of the figure the arrows depict “person B’s” integration of: 1) processing her own visual information about an object (arrow from person B to apple); 2) processing information about the visual attention of person A (arrow from person B to person A), and 3) processing self referenced information about her attention to the apple and person B, including feedback from eye muscles, affect or arousal (e.g. heart rate) associated with information processing of the event, reward value of self generated goal directed behavior in the context of the event, etc., (also see Figure 1c1,2,3). The top portion of the figure illustrates the developmental internalization of operations of joint attention. The behavioral operations once first involved in the social coordination of attention and information processing with respect to and external referent becomes re-described as a social executive system that contributes to the human capacity to socially coordinate mental attention to a representation of shared experience.
Figure 3.
Illustration from Mundy & Newell (2007) depicting the lateral (top) and medial (bottom) illustrations of the distributed system of Brodmann’s cytoarchitectonic areas of the cerebral cortex associated with Initiating Joint Attention and the anterior attention system, as well as RJA and the posterior attention systems. The former include areas 8 (frontal eye fields) 9 (prefrontal association cortex), 24 (anterior cingulate), 47 and 43 (orbital prefrontal and insula associative cortices). The latter include areas 7 (precuneous, posterior parietal association area), 22, 41, and 42 (superior temporal cortex) and 39 and 40 (parietal, temporal, occipital association cortex).
Figure 4. Illustration of the face/gaze direction stimuli that preceded the appearance of the ball stimuli for congruent (joint attention) and non-congruent (non-joint attention) conditions in Williams et al. (2005). Illustration of regions of cortical activation uniquely associated with the joint attention condition (red), the non-joint attention condition (green) and both (green on red).
Figure 5.
Illustrations of facial stimuli observed by participants in the self initiated joint attention condition and other initiated joint attention conditions in the study of Schilbach et al. (In press). Composite brain images can also be seen for frontal-parietal activation associated with the joint attention condition; parietal central activation associated with the non joint attention condition, and ventral striatum activation associated with self initiated joint attention versus prefrontal activation associated with other initiated joint attention condition.
Figure 6. An estimation of the distributed neural network activated in joint attention, and its commonality with the network of brain regions active only in self processing (self > other) and co-active in both self and other information processing. Estimation of the former was derived from data available in Schilbach et al. (In press) and Williams et al. 2005. Estimation of the latter was based on data available from Northoff et al. (2006) and Lombardo et al. (2007). Note activation associated with the insular cortex is not illustrated in this figure.