Infant Behavior & Development 40 (2015) 64–72

Contents lists available at ScienceDirect

Infant Behavior and Development

Full length article

How do infants recognize joint attention?

Erik Gustafsson[a], Julie Brisson[b], Christelle Beaulieu[a], Marc Mainville[a], Dominique Mailloux[a], Sylvain Sirois[a,*]

[a] Université du Québec à Trois-Rivières, Chaire de Recherche du Canada en Neuroscience du Développement Cognitif, Trois Rivières, QC, Canada
[b] Université de Rouen, Laboratoire Psychologie et Neurosciences de la Cognition et de l'Affectivité (EA4700), Département de Psychologie, Rouen, France

ARTICLE INFO

Article history:
Received 12 September 2014
Received in revised form 6 April 2015
Accepted 12 April 2015

Keywords:
Joint attention
Pupillometry
Social cognition
Violation of expectation

ABSTRACT

The emergence of joint attention is still a matter of vigorous debate. It involves diverse hypotheses ranging from innate modules dedicated to intention reading to more neuro-constructivist approaches. The aim of this study was to assess whether 12-month-old infants are able to recognize a "joint attention" situation when observing such a social interaction. Using a violation-of-expectation paradigm, we habituated infants to a "joint attention" video and then compared their looking time durations between "divergent attention" videos and "joint attention" ones using a 2 (familiar or novel perceptual component) × 2 (familiar or novel conceptual component) factorial design. These results were enriched with measures of pupil dilation, which are considered to be reliable measures of cognitive load. Infants looked longer at test events that involved novel speaker and divergent attention but no changes in infants' pupil dilation were observed in any conditions. Although looking time data suggest that infants may appreciate discrepancies from expectations related to joint attention behavior, in the absence of clear evidence from pupillometry, the results show no demonstration of understanding of joint attention, even at a tacit level. Our results suggest that infants may be sensitive to relevant perceptual variables in joint attention situations, which would help scaffold social cognitive development. This study supports a gradual, learning interpretation of how infants come to recognize, understand, and participate in joint attention.

1. Introduction

At around 6 months of age, infants begin to follow shift of gaze or head turn of adults, and by 12 months of age are able to actively coordinate focus on an object with a second person (Butterworth & Jarrett, 1991; Corkum & Moore, 1998; Morales et al., 2000; Scaife & Bruner, 1975). Labeled as "joint attention", this ability for a triadic interaction between two people and an object has been described as a key component of our social cognition, allowing the sharing of experience and knowledge (Heal, 2005). In fact, infants' skills in initiating and responding to joint attention predict their linguistic, social, and emotional abilities in later life (Morales et al., 2000; Mundy et al., 2007; Mundy & Gomes, 1998; Mundy & Newell, 2007; Parlade et al., 2009; Vaughan Van Hecke et al., 2007). For this reason, elucidating phylogenetic and ontogenetic factors trigger the emergence of joint attention is crucial for understanding the development of human social cognition. However, this is

* Corresponding author at: Département de Psychologie, Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières, QC, Canada G9A 5H7.
E-mail address: sylvain.sirois@uqtr.ca (S. Sirois).

http://dx.doi.org/10.1016/j.infbeh.2015.04.007
0163-6383/

still a matter of vigorous debate (see Tomasello, Carpenter, Call, Behne, & Moll, 2005a). Several non-exclusive hypotheses have been proposed, often based on the relationship between "joint attention" and the understanding of other persons as intentional agents (Tomasello, 1995). One line of hypothesis proposes hardwired modules dedicated to intention reading, such as an action-interpretation system that perceives action as goal directed (Gergely & Csibra, 2003), or a "shared-attention mechanism" involving innate modules perceiving goals and eye gaze direction (Baron-Cohen, 1997). Such modules may be involved in joint attention by helping the infant either to monitor the adult reaction when pointing to an object, or to become aware when the adult wants to share the perception of an object. Other studies focused on mirror-neuron systems believed to mediate action understanding in humans (Buccino, Binkofski, & Riggio, 2004; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Kohler et al., 2002), but whether such neurons are innate, or shaped by experience, is still debated (Ferrari, Tramacere, Simpson, & Iriki, 2013) as is their role in action understanding (Hickok, 2009). Presumably, the developmental pathway for understanding and responding to joint attention depends on early social interactions.

An alternative approach addresses the identification of such ontogenetic processes and their role. Dominant contemporary perspectives under the social-cognitive model suggest that infants begin to understand particular kinds of intentional states in others only after they have experienced them first in their own activity (Falck-Ytter, Gredebäck, & von Hofsten, 2006; Sommerville & Woodward, 2005; Tomasello et al., 2005a). According to this model, initiating and responding to joint attention should be highly correlated processes, insofar as they depend upon the same association between intention and goal-related behaviors in oneself and in others. The neuronal systems supporting such identification with others would flourish through certain typical sensorimotor and social contingencies such as protoconversation or mirroring behavior of parents toward their infant (Heyes, 2010; Trevarthen, Kokkinaki, & Fiamenghi, 1999). Epigenetic effects may also be involved, although the specific socio-environmental stimuli likely to trigger these molecular changes are yet to be identified (Ferrari et al., 2013). Moreover, one problem in addressing the potential ontogenetic processes implicated in understanding of joint attention in infants concerns the wide variety of child-rearing practices across human cultures. In fact, it is not known whether children across cultures are similarly exposed to supposedly typical social interactions such as protoconversation and mirroring (Bhavnagri, 1986; Bhavnagri & Gonzalez-Mena, 1997; Kagan & Klein, 1973; Ochs & Schieffelin, 2001). Considering the diversity of human ontogenetic niches, the mechanisms involved in ability to respond to communicative signals, such as those implicated in joint attention, would have to be extremely flexible.

Previous studies have largely focused on how infants manage to respond to joint attention. However, it is important to note that initiating joint attention and responding to it are abilities that should be considered separately (Mundy & Newell, 2007). Indeed, Mundy (2003) demonstrated that children diagnosed with autism, a disorder characterized by social interaction impairment, show capacities for responding adequately to joint attention but rarely for initiating joint attention behavior. A similar pattern of results was found in a study investigating social cognition in chimpanzees (Tomasello, Carpenter, & Hobson, 2005b). Moreover, responding to joint attention can be measured as early as 6 months of age; that is, at least 3 months before infants can initiate a joint attention situation (Mundy & Newell, 2007). Taken together, these findings suggest that distinct mechanisms could underlie different features of joint attention abilities, suggesting that alternative models are required to fully explain the development of joint attention in infancy.

Mundy, Card, and Fox (2000) and Mundy, Sullivan, and Mastergeorge (2009) proposed such a model based on the fact that responding to joint attention activates the posterior neuronal systems dedicated to orientation and perceptual attention while initiating joint attention is related to the late maturing anterior attention systems in infants. Once adult, imaging data suggest an integrated activity of anterior and posterior systems during joint attention situations (Henderson, Yoder, Yale, & McDuffie, 2002; Williams, Waiter, Perra, Perrett, & Whiten, 2005). This attention-system model suggests that learning about and from self-control attention is the first step to joint attention. At a later stage, maturation of cortical networks, in conjunction with experiencing joint attention situations, would result in a sufficiently organized, accurate, and fast integrated processing of information between anterior and posterior neuronal networks thought to support intention reading and goal understanding (Mundy & Newell, 2007). Such a model represents a more parsimonious hypothesis compared to the social-cognitive one. Indeed, it suggests that before gaining an accurate comprehension of the conceptual features of joint attention supported by intention reading abilities, infants may just initiate and respond to joint attention based on orienting and attention systems elicited by gaze direction.

The aim of this study was to assess whether 12-month-old infants are able to recognize a "joint attention" situation between two people (one pointing out a toy to the other) as a specific social interaction. For that purpose, using a violation-of-expectation paradigm, we compared looking time durations between "divergent attention" and "joint attention" situations. Infants are expected to look longer in a "divergent" condition since the novelty of the situation should elicit stronger reactions from infants. However, preferential looking at a conceptually novel event is often confounded with that event's perceptual novelty (Hunter & Ames, 1988; Jackson & Sirois, 2009; Roder, Bushnell, & Sasseville, 2000). In order to explore independently and jointly the effects of conceptual novelty and simple perceptual novelty, we used a 2 (familiar or novel component) $\times$ 2 (joint or divergent attention) factorial design (Bogartz, Shinskey, & Speaker, 1997). However, conceptually distinct events are, by necessity, perceptually distinct (Sirois & Mareschal, 2002), making looking time data potentially equivocal. For this reason, the substantial and relevant information provided by looking time durations will be enriched with measures of pupil dilation. Although primarily a function of luminance, pupil dilation is also an unbiased indicator of information processing load (Beatty, 1982). Pupil diameter is positively associated with cognitive arousal and has already been used to investigate the detection of impossible events and irrational social interactions by infants (Gredebäck & Melinder, 2010; Jackson & Sirois, 2009; Karatekin, 2004, 2007; Porter, Troscianko, & Gilchrist, 2007; Sirois & Jackson, 2012). Rather than contrast cumulative