VARIABILITY AND INDIVIDUAL DIFFERENCES IN EARLY SOCIAL PERCEPTION AND SOCIAL COGNITION
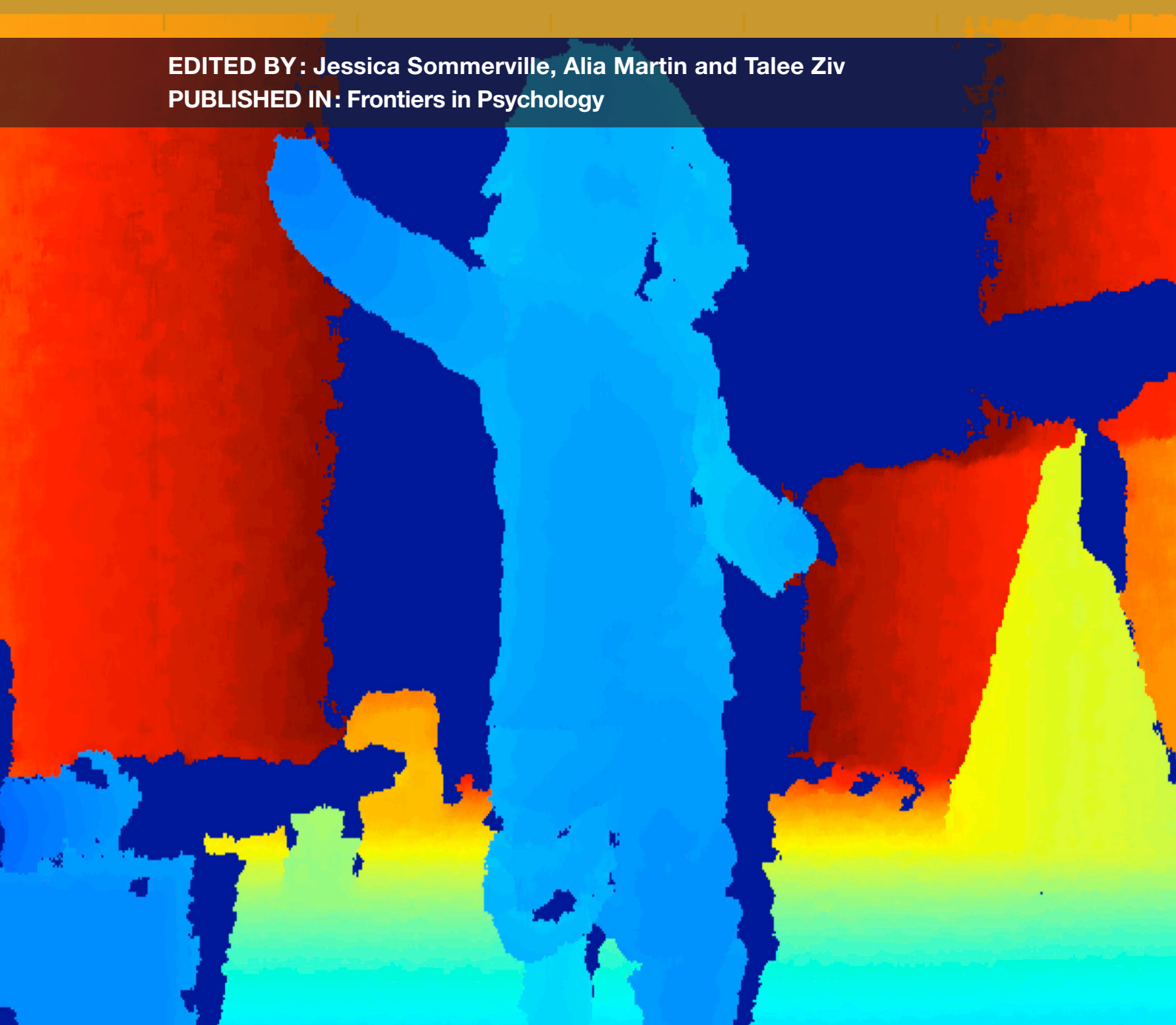
EDITED BY : Jessica Sommerville, Alia Martin and Talee Ziv

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: **researchtopics@frontiersin.org**

VARIABILITY AND INDIVIDUAL DIFFERENCES IN EARLY SOCIAL PERCEPTION AND SOCIAL COGNITION

Topic Editors:
Jessica Sommerville, University of Washington, USA
Alia Martin, Harvard University, USA
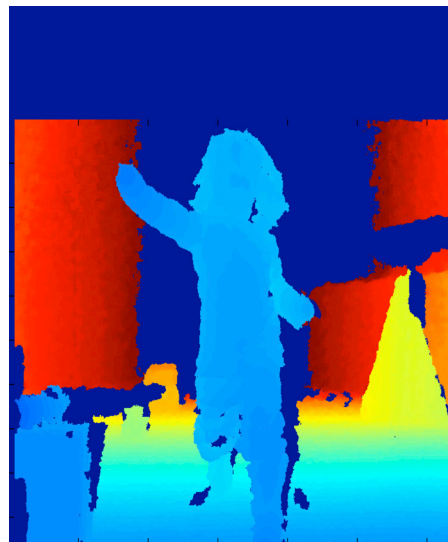**Talee Ziv**, University of Washington, USA

Image by Robert Hepach

Over the past three decades mounting evidence has suggested that infants' social perceptual and social cognitive abilities are considerably richer than was once thought. By the end of the second year of life, infants discriminate faces along various social dimensions, attend to and understand others' goals and intentions, use the emotions of others to guide their learning and behavior, attribute dispositional characteristics to other agents, and make basic social evaluations. What has also become clear is that there is a great deal of variability in infants' social perception and cognition. A critical, outstanding question concerns the nature and meaning of such variability.

The proposed research topic welcomes papers addressing cutting-edge questions regarding variability and individual differences in early social perception and social cognition. The goal of these papers is to investigate overarching questions in this domain, which are necessary to move the field forward.

Variability in early social perception and social cognition (among other domains) in infancy and early childhood is often attributed to noise, or overlooked in favor of focusing on age-related changes. Yet, recent work suggests that variability in social perceptual and social cognitive tasks reliably inter-relates, and predicts real-world social behaviors. For example, infants' everyday experience with different face categories predicts individual differences in face processing, infants' production of goal-directed actions predicts their simultaneous understanding of these actions, and variability in social attention during the second year of life is related to theory of mind during the preschool years. These findings suggest that variability in performance on social perception and social cognition tasks is not merely a nuisance variable, but, rather, may provide the key

to addressing significant questions regarding the nature of infants' social perception and social cognition, and the processes that underlie developmental change. Acknowledging and closely examining and investigating variability in early social perceptual and social cognitive abilities may represent a powerful approach for understanding development in (at least) two ways. First, variability can signal transitional points in the developmental onset of a given ability. Thus, such variability, and the extent to which variability relates to experience and/or other abilities, can be used to test hypotheses regarding mechanisms that underlie developmental changes. Second, variability can represent more enduring individual differences between infants. In this case, critical questions arise regarding the source of individual differences (that is, what factors shape the emergence of individual differences?) and whether such early individual differences contribute to the development of more advanced and sophisticated forms of social cognition and behavior.

The goal of this research topic will be to encourage researchers to take variability in early social perception and cognition seriously. Papers that give variability center stage, and are aimed at addressing the value of variability for identifying developmental mechanisms, as well as investigating the existence, source, and antecedents of early individual differences in social perception and social cognition are welcomed. Taken together, the contributed papers will provide integral new information to the study of social perception and social cognition over the first three years of life.

Table of Contents

Editorial: Variability and Individual Differences in Early Social Perception and Social Cognition

Alia Martin[1]*, Talee Ziv[2]* and Jessica A. Sommerville[2]*

[1] *Department of Psychology, Harvard University, Cambridge, MA, USA,* [2] *Department of Psychology, University of Washington, Seattle, WA, USA*

The Editorial on the Research Topic

Variability and Individual Differences in Early Social Perception and Social Cognition

In this research topic, we showcase state-of-the-art research on the sources and meaning of variability and individual differences in early social perception and cognition. These papers demonstrate that such variability contributes to our understanding of early development, requires specific methodological toolboxes and skill sets to expand our developmental inventory, and relates to important abilities later in life.

The first set of papers (Examining individual variability to elucidate developmental mechanisms and pathways) highlights the importance of focusing on individual variability (above and beyond group variability) for understanding the underlying mechanisms of early social cognition and how meaningful early differences relate to abilities later in childhood. Hepach et al. propose that studies of social cognition go beyond using a single behavioral measure (i.e., looking time) and offer two novel measures, pupil dilation and postural changes. They argue that these measures may provide insight into the mechanisms and motivations underlying behavioral responses (e.g., prosocial behaviors) in which there may be individual differences even if behavioral indices are similar at a group level.

The other papers in this section use longitudinal research to investigate the meaning of measures of infant social cognitive development in the context of broader developmental trajectories. Gampe et al. found that infants' action perception and action production in a contralateral reach task improve consistently across the second half of the first year of life at a group level; however, infants' individual action production abilities were not correlated across time, and there was no consistent causal relation between action perception and production across age. Thus, this study of individual differences uncovers a dynamic link between action perception and production in the first year of life. Brink et al. investigate how infant social attention relates to preschool social cognition. They found that infants' habituation to a social display of intentional action was related to measures of infant social-interactive experiences and temperament and, later, preschool theory of mind. These studies highlight the role of both experience and temperament in predicting social cognitive abilities, as each factor was an independent predictor of preschool theory of mind ability. Adding to this work on the downstream consequences of social cognition in infancy, Kristen-Antonow et al. examine effects of early individual differences in social cognition on the developing sense of self. Infants' social responsiveness in a still-face paradigm predicted their later mirror-self-recognition, and their imitation abilities predicted later delayed-self-recognition. This work emphasizes the importance of measuring variability to expose the underlying mechanisms of development.

The next three sets of papers focus on variability and individual differences in early social cognition across several other topics. For example, while it is known that infants attend to and encode others' goal-directed actions from an early age, the papers by Bakker et al., Gerson et al., and Dunfield and Johnson illustrate that there are subtle nuances in how different individuals might process goal-related events (Variability in the encoding of actions and goals). Bakker et al. report an EEG study with 9-month-old infants. They found differences in the P400 component across their sample when infants observed a typical instance of intentional non-verbal communication, a "give-me" gesture, vs. a similar non-intentional gesture, a rotated hand. However, the P400 differences between the two types of gesture were significantly larger in female infants than male infants, suggesting a deeper encoding of communicative intent in females and raising questions about what mechanisms might lead to these differences.

Gerson et al. also show evidence of individual differences in 8-month-old infants' goal processing. Infants who became more planful in their goal-directed actions after training in a means-end task were more able to process another person's goal in a means-end task than infants who received the training but did not become as planful afterward. These studies suggest that infants' own goal-directed action production not only influences their ability to parse goal-directed behavior, but that this link varies based on infants' success in learning to achieve their goals.

Dunfield and Johnson ask whether instrumental and social goals differ, and consider whether the study of adult social cognition can inform the study of infant social cognition. When goals were unambiguous, they found no individual differences in adults' ability to identify instrumental goals or social goals. When goals were ambiguous, however, insecurely attached adults were significantly less likely to attribute social goals than securely attached adults. All of these papers suggest that there is meaningful variability in how we perceive and interpret goals across development.

The next set of papers focuses on variability in how infants perceive the facial and affective information in their environment (Variability in the factors moderating children's encoding of facial and affective information). Kim et al. report a preference for female faces over male faces in 3- and 10-month-old infants, but only for white faces (not black or Asian). This effect did not differ based on the race of infants' own primary caregiver. Interestingly, Liu et al.'s study conducted in China with Chinese infants revealed a preference for female over male Asian faces at 3 and 6 months, but not 9 months of age. Furthermore, this gender preference was not observed for Caucasian faces, highlighting the importance of examining variability across cultural contexts. Liu et al. also found an effect of experience on the own-race female face preference in their sample: infants who had more experience with male caregiving showed a decreased preference for female faces.

Ravicz et al. ask whether temperament differences in infants are related to affective perception in the form of responses to emotional facial expressions. Using fNIRS, they tested infants' neural responses to happy female faces, and found that infants with lower negative emotionality showed more preferential left hemisphere prefrontal cortex activation to the happy facial

expressions. Ben-Israel et al. examine how innate factors may influence affective processing. They investigated whether a dopamine D4 receptor polymorphism is related to sex differences in affective knowledge in young children. In a longitudinal twin study, they found that for carriers of the 7-repeat allele in both their 3- and 5-year-old samples, boys had higher affective knowledge scores than girls, but there were no sex differences in non-carriers. These differences were in contrast to sex differences found in adults in which females with the polymorphism show increased empathy compared to males. These results illustrate the importance of considering that variability (for example, in genetic effects on cognition) may differ in type and meaning at different developmental timepoints.

In the following section (Effects of variability in linguistic experience and ability on early cognition), two papers examine the effects of variation in early linguistic experience and skills on children's social cognitive abilities. Zimmermann et al. found the typical video transfer deficit in a puzzle-solving task: 2-year-olds solved the puzzle more effectively following imitation of a live model compared to a non-interactive video model. Yet individual differences in children's ability to spontaneously generate a semantic label for the puzzle (e.g., "a sailboat") were associated with a reduction in the video transfer deficit, suggesting that children's ability to solve the problem using a verbal cue helped them to learn from a non-interactive model.

Henderson and Scott focus on the role of linguistic experience on infants' understanding of language as a conventional system, building on previous work showing that by 9 months monolingual infants expect two speakers to use the same word to refer to an object. They found that variability in language exposure has an important influence on this expectation. Thirteen-month-olds growing up in a bilingual environment did not expect two speakers who speak the same language to use the same word to refer to an object, suggesting that bilingual infants may have expectations of conventionality in language that are different from those of monolingual infants.

The final group of papers (Relation of parent and socialization factors to early individual variability in social cognition) brings up an important source of variability in early experience that can affect social perception and cognition: parent socialization. Upshaw et al. measured 12- and 15-month-old infants' pupil dilation as an index of arousal in response to viewing displays of infants expressing happy and sad emotion. They found that participants' arousal was meaningfully associated with their parents' social dispositions and behaviors. Infants' arousal in response to sad emotion was correlated with parent self-reported empathic perspective taking, and infants' arousal to happy emotion was correlated with parent self-reported altruism. Gross et al. conducted a complementary study which directly assessed parental reports of their infant socialization practices and looked at relationships to infant social attention and prosocial behavior. Overall, parent socialization of prosocial behavior was related to some aspects of prosocial behavior, and moderated the link between social understanding and prosocial behavior. Interestingly, parent socialization mattered more than infant temperament, which did not meaningfully predict differences in prosocial behavior. Finally, Wade et al. show that the way parents socially interact with their infants can

serve as a protective factor against potential risks to infant social cognitive development. In a longitudinal study, they found that cumulative biomedical risk affects variability in social cognition at 18 months, but that this relationship is moderated by parental responsiveness; that is, higher biomedical risk predicted reduced social cognitive abilities only for infants with less responsive parenting.

Taken together, the papers in this Research Topic underscore the importance of considering variability and individual differences in infants' early social perception and cognition, seeking out appropriate methods to accurately measure such variability, and looking for meaningful sources of these differences that can shed light on the mechanisms underlying children's early abilities as well as developmental change over time.

AUTHOR CONTRIBUTIONS

All authors (AM, TZ, JS) contributed equally to the editing process and organization of the Research Topic. AM drafted the editorial, and TZ and JS provided critical feedback and edits.

frontiers
in Psychology

Novel paradigms to measure variability of behavior in early childhood: posture, gaze, and pupil dilation

Robert Hepach[1]*, Amrisha Vaish[2] and Michael Tomasello[1]

[1] *Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany,* [2] *Department of Psychology, University of Virginia, Charlottesville, VA, USA*

A central challenge of investigating the underlying mechanisms of and the individual differences in young children's behavior is the measurement of the internal physiological mechanism and the involved expressive emotions. Here, we illustrate two paradigms that assess concurrent indicators of both children's social perception as well as their emotional expression. In one set of studies, children view situations while their eye movements are mapped onto a live scene. In these studies, children's internal arousal is measured via changes in their pupil dilation by using eye tracking technology. In another set of studies, we measured children's emotional expression via changes in their upper-body posture by using depth sensor imaging technology. Together, these paradigms can provide new insights into the internal mechanism and outward emotional expression involved in young children's behavior.

Keywords: posture, eye tracking, Kinect, pupillometry, pupil dilation, emotion, children, internal arousal

Introduction

Children's navigation through the social world rests on a set of socio-cognitive abilities that emerge during infancy and enable children to tune in to others' emotions and mental states (Carpenter et al., 1998; Tomasello et al., 2005; Grossmann and Johnson, 2007). These abilities in turn allow children to later establish and maintain social relationships. Nevertheless, there are individual differences with regards to the mechanisms that moderate children's social navigation and there is variability in children's responsiveness to others' feelings and desires (Dunn et al., 1991; Eisenberg et al., 1996; Rothbart et al., 2000; Knafo et al., 2008; Salley et al., 2013).

A central challenge to measuring the underlying processes of behavior is that these processes are often either internal or partially based on expressive emotions that occur briefly and rapidly in succession. However, recent advances in eye tracking and depth sensor imaging technology (1) allow us to 'listen in' on the internal states underlying behavior, and (2) provide a new lens through which emotional expressions become visible. Here, we illustrate two novel paradigms recently developed in our lab: one on children's responses to seeing others in need of help and another on children's postural changes following goal-oriented behavior. In our studies, the technology captures children's physiology and physiognomy from a distance so as to retain a natural setting in which the targeted behavior occurs.

The first paradigm was designed to capture children's gaze and pupil dilation as indices of attention and changes in internal arousal, respectively, in behavioral studies. Both variables are

frequently assessed in studies on infants' physical and social cognition (Aslin and McMurray, 2004; Aslin, 2007; Falck-Ytter et al., 2013; Sirois and Brisson, 2014). However, while eye tracking has previously been employed to study cognition, our approach focuses on explaining children's behavior and addressing questions concerning its underlying motives and proximate mechanisms in active behavioral paradigms with children. Here we illustrate the use of this method to address questions regarding the specific motivations underlying children's prosocial behavior (see Gaze and Pupil Dilation). The second paradigm was designed to measure children's emotions, in particular their positive emotions as expressed in their posture. Previous research had focused on facial expressions and composite measures of gestures and posture to identify positive affect in young children through human coders' judgment. However, it has thus far not been possible to automatically capture children's emotions and to focus on specific body parts, e.g., the chest and hip. We apply this technique in situations where children achieve an outcome for themselves and we measure how their posture accordingly changes compared to a prior baseline (see Posture). This allows us to address questions regarding the emotions that accompany behavior, e.g., positive affect following success.

Here we illustrate the two paradigms, which involve capturing children's eye movements and pupil dilation to measure changes in autonomous nervous system (ANS) activity, and measuring children's posture in order to assess changes in their emotional state. We propose that these measures can not only provide insights into the internal mechanisms and emotional bases of children's behavior but also allow researchers to trace the physiological antecedents of children's actions to better understand the sources of variability and individual differences in social cognition and behavior.

Gaze and Pupil Dilation

For children to become competent social partners, they need to realize when others are in need of help, represent the appropriate solution, and have a sympathetic motivation to care for others' needs (Eisenberg and Miller, 1987; Zahn-Waxler et al., 1992; Dunfield, 2014; Warneken, 2015). The attention children pay to others' actions can be measured through tracking their eye movements and mapping them on a visual scene (Aslin and McMurray, 2004). Eye tracking is based on corneal reflection technology and provides numerous non-invasive indicators of attention, including fixations, saccades, anticipatory looking, and scan patterns (Aslin, 2007, 2012; Gredebäck et al., 2009; Oakes, 2012). This has opened up new ways of studying social cognition in infants and young children (Navab et al., 2012; Falck-Ytter et al., 2013; Tenenbaum et al., 2013). Both the time children spend looking at a scene and the pattern of eye movements can reveal the underlying structure of children's social attention (Falck-Ytter et al., 2006; Aslin, 2007; Frank et al., 2012; Fawcett and Gredebäck, 2013; Elsner et al., 2014).

An additional feature of modern corneal reflection eye trackers is the automatic capture of pupil diameter (Wang,

2011). Similar to other physiological measures such as heart rate or skin conductance, changes in pupil dilation reflect activation of the ANS (e.g., Kahneman et al., 1969; Libby et al., 1973; Bradley et al., 2008). This is particularly interesting for developmental research because whereas eye movements can reflect the distribution of attention, changes in pupil dilation may provide a measure of the degree of psychological involvement in pre-verbal and just-verbal populations (see Goldwater, 1972; Laeng et al., 2012; Sirois and Brisson, 2014, for reviews). Similar to the measure of children's eye movements, the measure of pupil dilation has also found application in infancy to study both physical and social cognition early in ontogeny (Falck-Ytter, 2008; Chatham et al., 2009; Jackson and Sirois, 2009; Gredebäck and Melinder, 2010, 2011; Geangu et al., 2011; Sirois and Jackson, 2012; Hepach and Westermann, 2013; Burkhouse et al., 2014).

The majority of previous work had implemented measures of gaze and pupil dilation in response to pictures or prerecorded video stimuli. To study more natural interactions researchers have adapted these set-ups for live paradigms wherein children sit facing an adult while the eye tracker records their eye movements and pupil size. Gredebäck et al. (2010) studied infants between the ages of 2–8 months in a live setting in which the experimenter or the mother sat facing the child. The authors found that infants' ability to follow others' gaze develops linearly and was more stable when facing a stranger compared to their mother (see Falck-Ytter et al., 2015, for a similar set-up). However, none of this previous work attempted to relate children's eye movements and pupil dilation to their behavior as a way to capture the mechanisms underlying children's behavior. Our aim was thus to extend the use of these measures in novel directions to study behavior more generally.

In one example we developed a paradigm to address questions regarding the motivation underlying young children's prosocial behavior. Specifically, we investigated how changes in internal arousal relate to children's own prosocial behavior. For this purpose we devised a behavioral helping paradigm within which we could capture children's gaze and pupil dilation to investigate the mechanisms underlying young children's helping behavior (Hepach et al., 2012). During the first 2 years of life children show a remarkable array of prosocial tendencies, including sharing with and comforting those in need of help (Zahn-Waxler et al., 1992; Warneken and Tomasello, 2009; Svetlova et al., 2010; Dunfield, 2014; Eisenberg and Spinrad, 2014; Paulus, 2014; Warneken, 2015). However, much less is known about children's motives to help others. Therefore we inquired whether changes in young children's internal arousal reflect their motivation.

The general set-up and procedure of our studies is comparable to other developmental studies on children's prosocial behavior, yet we include a crucial difference. At pre-defined time points, children (24-month-olds) sit in front of an apparatus that resembles the facade of a house through which they can view the scene they themselves moments ago participated in. Children watch the scene on a computer screen, which shows a live video feed of the events on the 'other' side of the apparatus. Through a series of familiarizations (see also Troseth and DeLoache,

1998), children learn that what they see on the screen is actually happening and that they can return to the task they were engaged in before they sat down (see **Figure 1** for an illustration). In our studies, children view an adult carry out a task such as stacking cans until at one point the final object accidentally drops to the floor out of his/her reach. Children are then given the opportunity to subsequently help (see Hepach et al., 2012, for details).

While children sit in front of the apparatus, an eye tracker records both their eye movements and changes in pupil dilation (Tobii model X120, SMI models RED and RED-m). The live feed is presented on the computer screen by capturing a video from a USB webcam. The presentation software of the eye tracking system (Tobii Studio with Tobii systems and Experiment Center with SMI) allows both displaying a live video and simultaneously recording eye data at a frequency of at least 60 Hz and uses a standard calibration procedure to map children's gaze onto the computer screen (Gredebäck et al., 2009). It is furthermore possible to apply the same *post-hoc* gaze correction techniques suggested for eye tracking experiments (Frank et al., 2012). This allows for children's gaze to be mapped onto the live scene that they are observing and in turn provides a glimpse into the underlying process guiding their visual attention (see **Figure 2**). To further match children's pupillary responses to the live scene, several additional steps are necessary.

Assessing changes in children's internal arousal in active behavioral paradigms is particularly challenging given that pupil size variations are highly volatile in response to children's body movement during the study. We therefore further developed a technique (first described in detail in Hepach et al., 2012), which focuses on a specific component of pupil dilation rather than a mean over a specified time window. Changes in pupil diameter are a function of both sympathetic and parasympathetic nervous system activity. This results in the typical pupillary oscillations both dilating and constricting the pupil. Even in the dark, the pupils constantly oscillate (Wilhelm, 1991), making the signal of pupil diameter changes over time highly volatile. These oscillations are very different from smooth sine wave patterns given that the magnitude of the positive peaks (dilation) and negative peaks (constriction) as well as the time interval between peaks varies (Loewenfeld, 1993). Psycho-sensory stimulation will increase pupil dilation such that the peaks are higher than before the presentation of the stimulus. Analyses of changes in pupil diameter focus on robust indicators of psychologically induced effects such as the number of oscillations over several minutes (Warga et al., 2009), peak dilation (e.g., Laeng et al., 2012) or the amplitude of the pupillary light reflex (PLR; Steinhauer et al., 2000).

The PLR is the characteristic shape of the change in pupil size upon the presentation of light. With increasing stimulus luminance the pupils constrict. Loewenstein (1920) studied the influence of various emotional states in clinical patients and observed the PLR to be inhibited during induced stress, e.g., when subjects experienced tension or witnessed a startling event. The elicitation of a PLR by shining a light into subjects' eyes is part of the standard procedure in ophthalmology to assess parasympathetic and sympathetic nerves innervating both eyes (Wilhelm, 1991; see Bakes et al., 1990; Heller et al., 1990, for examples). Furthermore, several psychological stimuli inhibit the PLR, e.g., following negative emotional events (Bitsios et al., 2004) and with increased attention during task demands (Steinhauer et al., 2000). An increase in internal arousal results both in overall increased pupil dilation (Bradley et al., 2008) and in an inhibited PLR (Henderson et al., 2014; though see Nyström et al., 2015, for a different interpretation of the PLR in comparison to tonic pupil diameter). The advantage of measuring the PLR as an indicator of internal arousal is its quick assessment within 2–3 s. The crux is that the experimental manipulation has to occur before and not while the PLR is elicited.

During behavioral studies the presentation of visual stimuli on a computer screen causes the pupils to constrict to the luminance properties of an image. We have developed a technique in which we elicit two PLRs in brief succession, i.e., a colorful image flashes twice on the computer screen. The recorded data are exported to a text file and processed using software such as R or Matlab. The exported data need to be pre-processed to remove extreme values (see Hepach et al., 2012, 2013, for filter and interpolation examples). Subsequently, an algorithm identifies the two pupil minima in response to the colorful image and averages both values. The raw value of pupil diameter, i.e., the average minimum, is reflective of individual differences in children's arousal state. To further capture a change
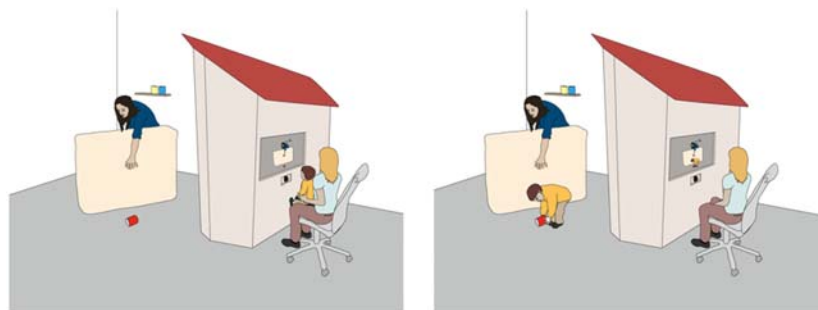
FIGURE 1 | The general set-up of the study. While children sit in front of the computer monitor, the eye tracking unit records both their eye movements and pupil diameter before they return to the adult's side.
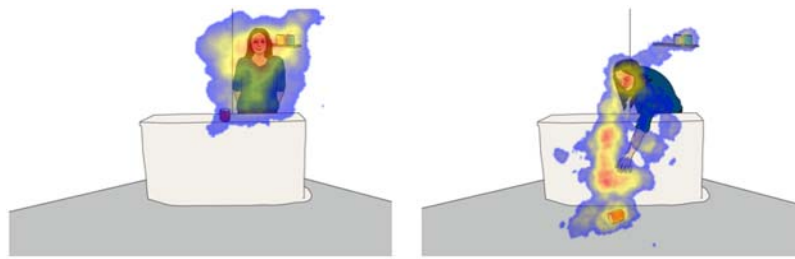
FIGURE 2 | Illustration of the scene children view while sitting in front of the apparatus. In the actual studies, children view a camera image of the adult carrying out her task. Since her behavior is never identical across participants, the illustrations here represent prototypical poses during specific phases of the study. The **(left)** represents the time during which the adult carried out her task, e.g., stacking cans, standing behind the table. The **(right)** represents the time the adult spent reaching for the dropped object. Each phase can last from a few seconds to several minutes. During that time, the eye tracker records children's gaze that can be mapped onto the illustrations to identify focal points of attention. The resulting focus maps reveal areas of low (blue) and high (red) attention.

in children's internal arousal in response to an experimental manipulation, we present the measurement image both before (baseline measure) and after (process measure) the experimental manipulation (e.g., seeing an adult needing help). The change is measured as the percentage increase from baseline to process (see **Figure 3**).

Through recording children's eye movements as well as changes in pupil dilation in the domain of prosocial behavior, we have found that children's own internal arousal increases when others are not helped but decreases to an equal degree when help is provided either by children themselves or by others (Hepach et al., 2012). Crucially, the degree of children's internal arousal reflects individual differences in the latency with which they engage in helpful behavior, i.e., the greater children's pupil dilation is after witnessing the situation, the faster they are to subsequently help others (see Hepach et al., 2013).

While gaze and changes in pupil dilation reveal the processes that precede and underlie a behavior, they do not provide information regarding valence, i.e., whether children's responses are positive or negative in valence. For this purpose it is necessary to use an alternative measure of emotional expressiveness.

Posture

Children's attention to and involvement with others' needs can explain the motivations leading up to a behavior. An equally important aspect of motivation is the question of how a behavior is maintained and reinforced. In the following, we illustrate a novel paradigm to measure the sorts of positive emotions that follow from successful behavior. From as early as 2 years of age, children show noticeable postural changes following their successful attainment of a goal. Such emotional expressions provide a window to assess a subject's feelings, i.e., the internal state, especially if the emotion is studied in context (Lewis, 1997). Postural changes are accompanied by gestures such as pointing to the achievement or self-applauding (Heckhausen, 1987, 1988). By the age of 3, children display an erect posture after succeeding on difficult tasks and conversely their posture decreases if they fail on easy ones (Lewis et al., 1992). Adults display similar changes in posture when they feel proud (Shiota et al., 2003; Tracy and Robins, 2004; Horberg et al., 2013), following athletic success (Weisfeld and Beresford, 1982), as a cue of social dominance (Schwartz et al., 1982), social status (Shariff and Tracy, 2009), and expertise (Martens and Tracy, 2013).
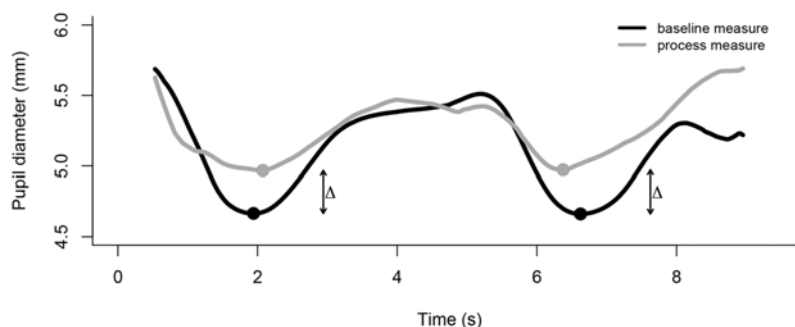
FIGURE 3 | The time course of the change in pupil size averaged across participants for two measurement time points, baseline and process. For each measurement point two pupillary light reflexes (PLRs) are elicited through presenting a bright stimulus on a dark background at 0 and 5 s, respectively. Upon presentation of the bright stimulus, the pupils constrict and re-dilate after reaching their respective minimum. An algorithm identifies the two minima and calculates the change from baseline to process pupil diameter.

Postural changes are reliably identifiable from a person's gait (Montepare et al., 1987) as well as body movement (Dael et al., 2012) and the ability to detect pride from pictures emerges between 3 and 7 years of age (Tracy et al., 2005). The most common way to measure posture is to apply coding criteria to video recordings (e.g., Montepare et al., 1987; Heckhausen, 1988; Lewis et al., 1992; Shiota et al., 2003), photographs (e.g., Tracy and Robins, 2004; Shariff and Tracy, 2009; Martens and Tracy, 2013), drawings (Schwartz et al., 1982), and computer-animated mannequins or point light displays (Atkinson et al., 2004; Coulson, 2004). With these studies, then, there is a documented relation between success and the positive emotion expressed in an expanded upper-body posture. However, little research has assessed young children's postural changes in naturalistic situations without relying on the coding of additional cues such as clapping (e.g., Heckhausen, 1987). Here, we introduce a recently developed paradigm using depth sensor imaging technology to automatically capture individual differences and changes in children's posture.

Automated Posture Assessment in Behavioral Paradigms

To track participants' movement and the location of body points, we use a Mircosoft Kinect adapter in behavioral studies. The Kinect is a specialized camera that captures both an RGB-image, like any regular webcam, as well as the information of how far away each captured pixel is from the device itself. This depth sensor imaging is achieved through an emitted infrared light and a separate lens capturing the reflection of the projected light. If a person is within the device's tracking range (∼1.5–4 m from the Kinect), the system estimates the x, y, and z-coordinates of up to 20 body points from the feet to the head (Shum et al., 2013; Stommel et al., 2015).

The Kinect allows for a relatively accurate and objective tracking of participants' body points and an assessment of body posture expansion (see **Figure 4**). The technology has been employed in several contexts studying infant-caregiver interactions (Nagai et al., 2012), children's behavior while playing cooperative games (Liu and LaFreniere, 2014), to assess and train motor abilities in clinical rehabilitation programs (Chang et al., 2013; de Greef et al., 2013; Luna-Oliva et al., 2013; Anzalone et al., 2014; Chung et al., 2014), and in interaction research (Won et al., 2014). In principal, the system can track multiple participants (Walczak et al., 2013) and it can be used to measure peripheral physiological measures such as respiratory rate (Burba et al., 2012). However, no experimental study has used the technology to capture the change in children's body posture as an indicator of emotional expression. This was our aim in recent work, which we describe next. Before doing so, we present the results from a study with adults in an attempt to validate the use of the technology in this way.

Validation Study with Adults

One of our central assumptions when using the Kinect technology in emotion research is that one can measure changes in upper-body posture that are related to positive and negative internal states. To test this assumption, we investigated whether the chest's center is more elevated when adult participants experience a positive emotion compared to a negative emotion.

Participants

Twelve naïve adult subjects (6 female, 14 years 5 months to 37 years 10 months, median age 26 years 4 months)
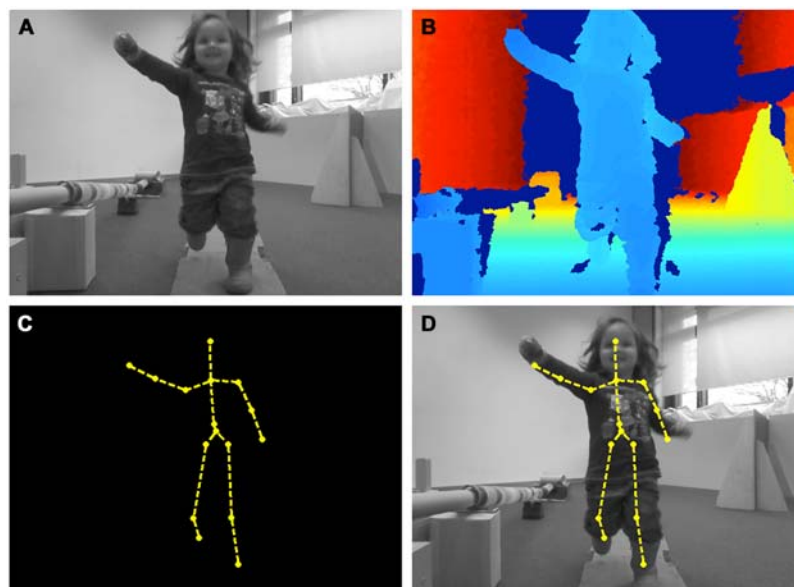
FIGURE 4 | The output of the Kinect motion sensor imaging technique. (A) RGB-channel, **(B)** depth contour image, **(C)** estimated skeletal joints, **(D)** mapping of the skeletal joints onto the RGB image.

were recruited from the Max Planck Institute for Evolutionary Anthropology and gave informed consent prior to participating in the study.

Materials and Design

Each subject was asked to imagine experiencing four specific emotions, one at a time: joy, pride (positive emotions), disappointment, and guilt (negative emotions). We used a Kinect camera to record participants' body posture, i.e., the height of their chest's center as well as of their hip's center. Adults were presented with four test trials with one emotion each. The order was counterbalanced across participants.

Procedure

Before the emotion trials, participants were asked to walk toward the Kinect camera (height = 0.85 m from the ground, distance from participants' starting position = 3.7 m, angle of the camera = 11°) such that a baseline assessment of the position of the chest's body joint could be made. The experimenter (blind to the study's hypotheses) read out the following instructions to participants: "This is a validation for a method to measure emotional expression. For this purpose we would like participants to walk toward the Kinect twice. At the very beginning we will conduct a baseline measurement for which we ask you to walk in a relaxed natural manner. Afterward I will read out the instructions for each emotion, four in total." At the beginning of each subsequent emotion trial the instructions were as follows (example joy): "The emotion to be displayed is joy. Can you recall an event that made you feel joy? Try to recall that feeling. Imagine the situation and surroundings. Take your time until you feel the emotion. Once you are ready give me a sign. Then you can walk toward the Kinect." Next, participants walked in the direction of the Kinect while the positions of both the chest's and hip's body point were recorded. For each trial (including baseline) we asked participants to walk toward the Kinect twice to average the data from both movements.

Data Analysis

The data were recorded with and analyzed in Matlab (see details below, in Section "Studies with Children"). Data from one trial of one participant (disappointment emotion) could not be used for analyses because of a system failure during recording. For each participant we calculated the change in the chest's height from baseline to the test trial for each of the four emotions and for each of 20 distance bins from the Kinect (1.2–3.2 m from the camera, 10 cm bin width). This controlled for differences in participants' walking speed. Furthermore, we averaged the values of joy and pride as well as disappointment and guilt to arrive at one composite positive and one composite negative emotion change score. For statistical analyses we further binned the data into four time windows of equal length and computed Wilcoxon exact paired tests with Bonferroni correction for multiple testing (adjusted significance level $p_{adj} = 0.0125$). We carried out the identical analyses with children's hip point height to investigate whether the effects of emotions were specific to the upper-body posture.

Results

Adults' change in chest height from baseline was more elevated during the positive compared to the negative emotion events immediately after the emotion manipulation, $p = 0.012$ (all other ps > 0.1). This was not the case when performing the identical analyses on the hip's center (all ps > 0.06; see **Figure 5**). These results suggested that adults' upper-body varies with the valence of the induced emotion. Measuring changes in upper body posture using the Kinect system can tap into the types of internal states involved in the experience and display during emotional episodes. This potentially makes the technology an interesting research tool to assess emotional expressions in young children.

Studies with Children

In our behavioral studies with 2-year-old children, participants can move around freely in a naturalistic setting without the need to attach point-light markers to their clothes. The Kinect 'draws' virtual points on participants' bodies. At specific time points during the study, the child moves toward the Kinect camera so that a full body image can be captured. In the following we provide data from one example (not reported with the original study) to illustrate that children's experience of an event that elicits a positive emotion reflects in changes of their body posture. At the beginning of the study we carried out a baseline measure during which children walked toward the Kinect without any experimental manipulation. At a later point in the study children manipulated a box to retrieve a toy that allowed them to continue with an attractive activity. Following this event, children again walked toward the Kinect camera. We hypothesized that experiencing this positive event would increase children's upper-body posture (see **Figure 6** for an illustration).

The tracking of multiple body points allows one to isolate, for example, changes in shoulder and chest height from changes in hip height. That is, even though up to 20 body joints can be tracked with the Kinect, we focused on children's upper-body posture following work on signs of pride in adults (Montepare et al., 1987; Tracy and Robins, 2004). More specifically, we calculated the height of the chest's center as an indicator of postural expansion. Through assuming an upright posture, the shoulders are pushed back which in turn elevates the chest. In principle, a lowering of chest height could reflect a slumped posture, as documented in states of negative affect (Lewis et al., 1992).

The data were recorded running a script written in Matlab. At regular time intervals, the program records (1) information regarding the position of each body point in three-dimensional space, (2) a color image, (3) a depth image, as well as (4) the location of each point on the two-dimensional color image (see **Figure 4**). Separate analyses (written partly in Matlab and R) calculate the difference in chest height between the baseline phase and the measurement taken during the test trial after the experimental manipulation. This results in a baseline-corrected change score that indicates the change in upper-body posture (see **Figure 7**).
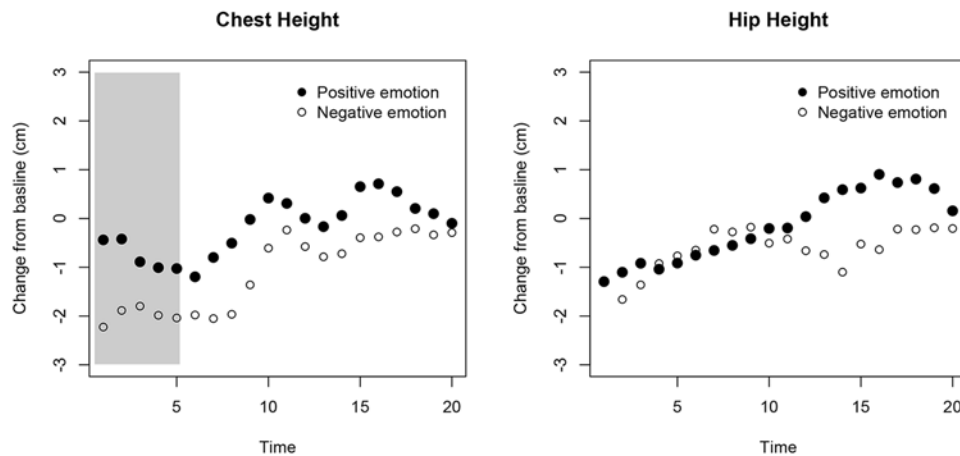
FIGURE 5 | Results from the adult validation study. The *x*-axis represents the time after the emotion was elicited and as adults walked toward the Kinect. The *y*-axis shows the relative change in height for participants' chest **(left)** and hip **(right)**. At each time point the median for the two positive and the two negative emotions is plotted. The gray area marks the time window where the difference between the two types of emotions was statistically significant (corrected for multiple testing).
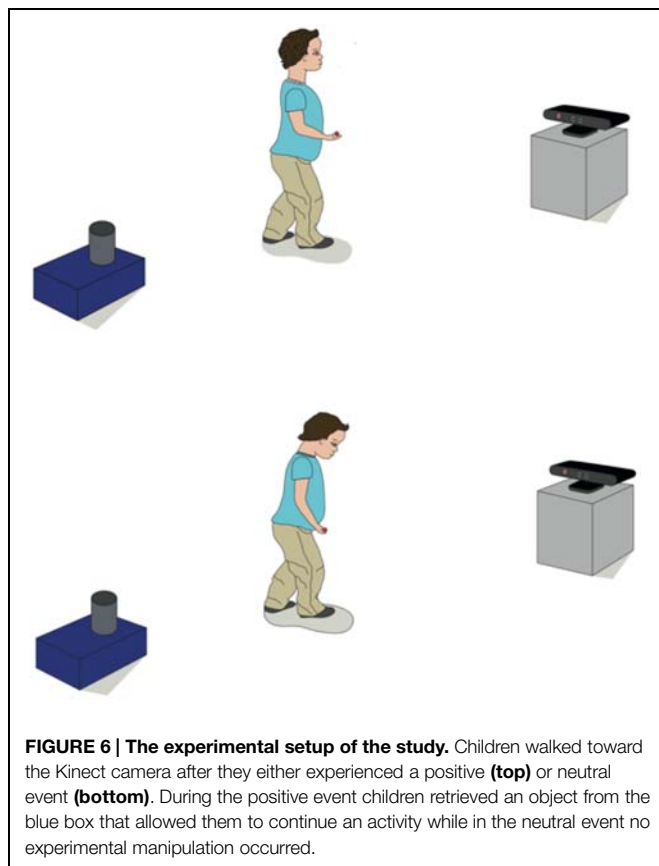
FIGURE 6 | The experimental setup of the study. Children walked toward the Kinect camera after they either experienced a positive **(top)** or neutral event **(bottom)**. During the positive event children retrieved an object from the blue box that allowed them to continue an activity while in the neutral event no experimental manipulation occurred.

FIGURE 7 | An illustration of the data collected by the Kinect camera. Each data point represents the change in the height of the chest's center for each participant. The values are sorted from the lowest change to the largest change in height. For three participants illustrations are provided from the baseline sequence **(top row)** and from the later taken process sequence **(bottom row)**.

Coding of Children's Affect

The assumption underlying the interpretation of the Kinect data is that changes in children's chest but not hip height reflect changes in positive affect. The more positive children feel, the more their upper-body posture should expand. To investigate this relation we asked two adults (blind to the study's hypotheses and type of trial) to rate the recordings of children's behavior along several dimensions.

The material consisted of the recordings of 48 children (25 girls, age range 29 months; 4 days to 31 months; 5 days; median age 30 months; 16 days) with two trials per child (baseline and test). The Kinect system could not record data for seven children on either trial or on both. For each trial coders were provided with the picture frames for the sequence when children started walking toward the Kinect, i.e., the exact same frames that were used for the automated posture analysis. The picture frames did not depict the skeletal information provided by the Kinect system.

For each trial the two coders were given the following instructions along with the SAM (self-assessment-manikin) rating scale (Bradley and Lang, 1994): "The SAM-rating consists of a valence coding and an arousal coding. The scale ranges from 9 to 1. For each trial, please answer the following questions: How pleasant is the emotion that the child is experiencing (very pleasant ∼9, very unpleasant ∼1)? How arousing is the emotion that the child is experiencing (very arousing ∼9, not at all arousing ∼1)?" In addition we asked coders what emotion they saw the child displaying and what features they paid attention to when identifying the emotion. The aim of the latter question was to investigate which emotions coders spontaneously associate with the behavior of the child. The ratings of the two coders were positively correlated, both with regards to rating valence [$\rho_{spearman}(n = 90) = 0.74$, $p < 0.001$, ICC = 0.63] and arousal [$\rho_{spearman}(n = 90) = 0.42$, $p < 0.001$, ICC = 0.36]. We therefore averaged both codings to arrive at composite measures of both valence and arousal.

The results showed that the rated pleasantness of the children's affect was greater in the test ($M = 6.45$, SD = 1.66) compared to the baseline ($M = 5.83$, SD = 1.73) trial, $t(41) = 2.24$, $p = 0.031$. On the other hand, there was no difference in the ratings of children's arousal between the baseline and test trial, $t(41) = 1.17$, $p = 0.25$ (see **Figure 8**). This suggests that the experimental manipulation of attaining a goal for oneself makes children appear to experience more pleasure compared to a baseline level. With regards to the coders' ratings of children's affect during the test trial and the change in children's posture, there was no overall relation between the two variables, $\rho(n = 41) = 0.087$, $p = 0.59$. However, very few trials were coded as 'negative,' i.e., with a value of less than 5 (17%). When focusing the analyses on the positive affect realm, i.e., ratings from 5 to 9, the degree of children's experienced affect was positively related to the change in their chest height from the baseline to the test trial. Children with ratings of high positive affect also tended to show a greater increase in upper-body posture, $\rho(n = 34) = 0.37$, $p = 0.03$ (see **Figure 8**). On the other hand there was no such relation with respect to children's lower-body posture, i.e., the change of hip height, $\rho(n = 34) = 0.08$, $p = 0.67$ (see **Figure 8**). In addition, no statistically significant relations emerged between children's rated degree of arousal and the change in their chest or hip height, ps > 0.09. Furthermore, the most frequently rated emotion after the experimental manipulation was 'happy' (see **Table 1** for details) and the most frequent features that coders paid attention to were children's smile, posture, and gait (see **Table 2** for details).

Overall, the Kinect depth sensor imaging technology not only provides information on individual differences in overall body size but also registers subtle changes in posture. Given the link between positive affect and increased body posture, the Kinect is an extremely useful new tool to measure emotional expressions and motivational states in children.

Summary and Future Directions

The paradigms described in this brief overview aim to capture the underlying mechanism of behavior and the types of expressive emotions that follow from it. Children's eye movements in response to live events in behavioral paradigms reveal how they allocate their attention. Likewise, changes in their pupil

FIGURE 8 | (Left) The average ratings regarding the valence and arousal of children's expressed affect for the baseline and test trial, separately. The vertical bars represent standard error bars. **(Center)** The relation between the rated valence of children's affect and their change in chest height for the test trial. The dashed vertical line represents the value corresponding to the neutral affect coding. Data points for the positive affect realm are highlighted and a regression line is added to illustrate the direction of the association. **(Right)** The relation between the rated change in children's affect valence and the change in hip height for the test trial.

TABLE 1 | The frequency of the type of emotion children expressed during the test trial as identified by the two coders.

	Angry	Sad	Embarrassed /uncertain	Disappointed	Disinterested	Surprise	Neutral	Active	Proud	Happy
Frequency	1	2	23	1	6	3	4	6	14	52

For the 48 children each coder could list two possible emotions per trial. The initial codings were grouped into one of the 10 categories. The maximum possible frequency for an emotion was 96, i.e., both coders identified the emotion for each of the 48 trials.

TABLE 2 | The features of children's behavior during the test trial that coders referred to when making the decision of what emotion the child was expressing.

	Smile	Other facial feature	Looking direction	Posture	Gait
Frequency	43	3	6	41	49

Each coder could provide two features per trial. The initial codings were grouped into one of the five categories. The maximum possible frequency for a feature was 96, i.e., both coders referred to this feature for each of the 48 trials.

dilation indicate the strength of their motivation. Individual differences in children's internal arousal before they carry out an action are related to how quickly they do so. In addition to these measures of internal arousal, changes in upper-body posture reflect children's positive emotional state after carrying out an action. A more straight and upright posture is indicative of a positive emotion while a hanging posture may reflect a negative emotional state (see also Lewis et al., 1992). These methods allow researchers to address novel questions regarding the underlying mechanisms of behavior (using pup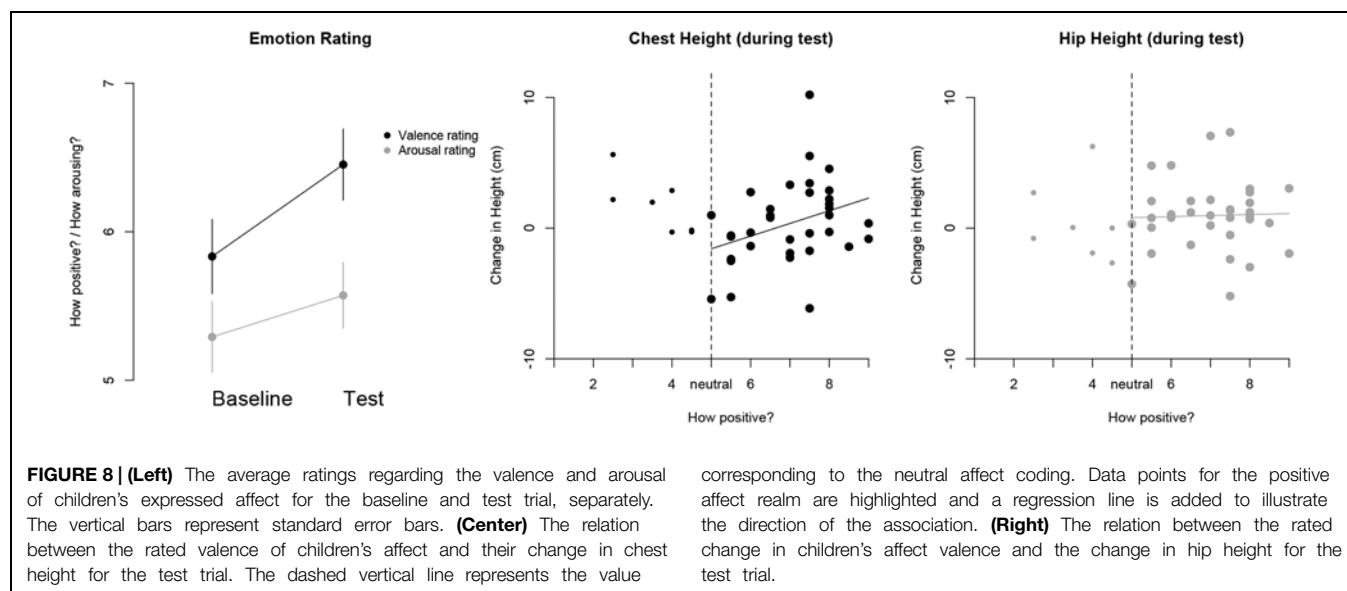il dilation) as well as children's emotional expressions that accompany behavior (using depth sensor imaging). One direction for future studies using eye tracking systems is to collect gaze and pupil data without the need for children to look at a computer screen. The fact that children have to move out of the live situation and temporarily sit in front of a separate apparatus interrupts the task they are involved in. In particular, younger children have difficulty sitting on their parent's lap after an engaging activity. This can result in inattentiveness and fewer chances to gather data points. Moreover, the image on the screen is only a 2-dimensional representation of a 3-dimensional space. In principle, both Tobii and SMI eye tracking systems can map participants' gaze onto a 'real' scene. This is a direction for future research to explore, especially with the emergent use of head-mounted cameras and eye tracking systems (Aslin, 2009; Smith et al., 2014). With regard to using the Kinect camera, an interesting further step is to explore whether the system can also capture other emotions, including those with negative affect such as shame and guilt. In particular, investigating the relation between the various body points, e.g., head vs. shoulders, will provide an interesting avenue for future research given that we have thus far only explored the change in chest height. In this way the Kinect could not only be used to address questions regarding the positive emotions that follow from successful actions but also to study emotional expression in children more broadly.

In the examples given here, measures of pupil dilation assessed children's internal arousal before they carried out an action while measures of posture were taken after children completed their action. In principle, neither technology needs to be restricted to these uses. In fact, there is now work measuring changes in internal arousal both before and after an experimental manipulation to investigate different motivations underlying children's helping behavior (Hepach et al., 2012). Likewise, children's emotional expression in their body posture may already change in anticipation of a positive or negative event.

The study of individual differences in children's behavior relies in part on providing novel dependent measures with which to investigate subtle differences in behavior. In particular, studying the underlying mechanisms allows researchers to address questions that go beyond asking whether or not a behavior occurred in a given context. In the present paper we have provided the example of children's prosocial behavior. Children, much like adults, do not help all the time and understanding the motivations that facilitate or inhibit helping is critical in our understanding of human prosociality. Changes in children's internal arousal, as measured via variation in pupil dilation, do not only reveal how children respond to others in need but also systematically relate to their willingness to engage in helping. Children with greater pupil dilation in response to seeing a person in need are faster to subsequently help (Hepach et al., 2013). In more recent work we have found that children show greater pupil dilation when viewing an adult struggling with an instrumental task compared to a non-social case that portrays an instrumental problem without a person present (Hepach et al., in review). In that study we further specified that it is children's process- but not baseline-measure that systematically relates to individual differences in helping behavior. Similar to the findings by Hepach et al. (2013), children with greater pupil dilation in response to seeing an adult in need were faster to help. Future work will have to investigate whether this relation between internal arousal and the latency to carry out a behavior is specific to children's helping or whether it applies to other contexts as well (such as play).

In addition to understanding the underlying mechanisms of behavior it is equally important to study variability in the expressive emotions that accompany children's behavior. Emotions can be measured via various modalities such as the voice or face. In the present paper we have illustrated a novel paradigm to measure children's posture after achieving a positive outcome. Such changes are relevant to behavior given that children are more likely to carry out an action if they find it enjoyable. Children did not only show variability in their emotional expression (see **Figure 7**) but the change in their upper-body posture systematically related to adult coders' ratings of the valence of children's expressed emotion. Children with greater increase in posture were rated to feel more positive. One possible avenue for future research is to investigate how others

perceive and respond to children's postural changes which may in turn have an impact not only on how children's subsequent posture changes but also on how they experience the actual emotion underlying the postural change (see also Carney et al., 2010).

Early in ontogeny, the technologies described here can provide a window into the underlying mechanisms of behavior. This is particularly relevant given that different processes can result in the same behavior (Karmiloff-Smith et al., 2014). The two examples in the present paper show that variability in children's behavior is meaningful, e.g., children's helping behavior is related to changes in their internal arousal. While the implementation of both pupillometry and depth sensor imaging was here illustrated in two specific contexts, their application need not be limited to children's prosocial and goal-oriented behavior. In fact, any researcher interested in variability of behavior early in ontogeny may find the research tools illustrated here useful for various forms of behavior in different contexts. Internal measures of attention and arousal as well as measures of emotional expressiveness move researchers closer to the source of variability. Together, these measures can be considered additions to the scientific toolbox with which researchers study the origins and development of children's social cognition and behavior. It will be a central challenge for future work to implement these techniques to study age-related changes in children's social cognitive development.

Acknowledgments

References

Anzalone, S. M., Tilmont, E., Boucenna, S., Xavier, J., Jouen, A. L., Bodeau, N., et al. (2014). How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3D++ time) environment during a joint attention induction task with a robot. *Res. Autism Spectr. Disord.* 8, 814–826. doi: 10.1016/j.rasd.2014.03.002

Aslin, R. N. (2007). What's in a look? *Dev. Sci.* 10, 48–53. doi: 10.1111/j.1467-7687.2007.00563.x

Aslin, R. N. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optom. Vis. Sci.* 86, 561–565. doi: 10.1097/OPX.0b013e3181a76e96

Aslin, R. N. (2012). Infant eyes: a window on cognitive development. *Infancy* 17, 126–140. doi: 10.1111/j.1532-7078.2011.00097.x

Aslin, R. N., and McMurray, B. (2004). Automated corneal-reflection eye tracking in infancy: methodological developments and applications to cognition. *Infancy* 6, 155–163. doi: 10.1207/s15327078in0602_1

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., and Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception* 33, 717–746. doi: 10.1068/p5096

Bakes, A., Bradshaw, C. M., and Szabadi, E. (1990). Attenuation of the pupillary light reflex in anxious patients. *Br. J. Clin. Pharmacol.* 30, 377–381. doi: 10.1111/j.1365-2125.1990.tb03787.x

Bitsios, P., Szabadi, E., and Bradshaw, C. M. (2004). The fear-inhibited light reflex: importance of the anticipation of an aversive event. *Int. J. Psychophysiol.* 52, 87–95. doi: 10.1016/j.ijpsycho.2003.12.006

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607. doi: 10.1111/j.1469-8986.2008.00654.x

Burba, N., Bolas, M., Krum, D. M., and Suma, E. A. (2012). "Unobtrusive measurement of subtle nonverbal behaviors with the Microsoft Kinect," in *Proceedings of the 2012 IEEE Virtual Reality Short Papers and Posters* (Washington, DC: IEEE Computer Society), 1–4. doi: 10.1109/vr.2012.6180952

Burkhouse, K. L., Siegle, G. J., and Gibb, B. E. (2014). Pupillary reactivity to emotional stimuli in children of depressed and anxious mothers. *J. Child Psychol. Psychiatry* 55, 1009–1016. doi: 10.1111/jcpp.12225

Carney, D. R., Cuddy, A. J., and Yap, A. J. (2010). Power posing brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychol. Sci.* 21, 1363–1368. doi: 10.1177/0956797610383437

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* 63, 1–174.

Chang, Y. J., Han, W. Y., and Tsai, Y. C. (2013). A Kinect-based upper limb rehabilitation system to assist people with cerebral palsy. *Res. Dev. Disabil.* 34, 3654–3659. doi: 10.1016/j.ridd.2013.08.021

Chatham, C. H., Frank, M. J., and Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5529–5533. doi: 10.1073/pnas.0810002106

Chung, I. C., Huang, C. Y., Yeh, S. C., Chiang, W. C., and Tseng, M. H. (2014). "Developing kinect games integrated with virtual reality on activities of daily living for children with developmental delay," in *Advanced Technologies, Embedded and Multimedia for Human-Centric Computing*, eds Y.-M. Huang, H.-C. Chao, D.-J. Deng, and J. J. J. H. Park (Dordrecht: Springer Netherlands), 1091–1097.

Coulson, M. (2004). Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *J. Nonverbal Behav.* 28, 117–139. doi: 10.1023/B:JONB.0000023655.25550.be

Dael, N., Mortillaro, M., and Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion* 12, 1085–1101. doi: 10.1037/a0025737

de Greef, K., van der Spek, E. D., and Bekker, T. (2013). "Designing kinect games to trainmotor skills for mixed ability players," in *Games for Health: Proceedings of the 3rd European Conference on Gaming and Playful Interaction in Health Care*, eds B. Schouten, S. Fedtke, T. Bekker, M. Schijven, and A. Gekker (Wiesbaden: Springer Fachmedien), 197–205.

Dunfield, K. A. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Front. Psychol.* 5:958. doi: 10.3389/fpsyg.2014.00958

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., and Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Dev.* 62, 1352–1366. doi: 10.2307/1130811

Eisenberg, N., Fabes, R. A., Murphy, B., Karbon, M., Smith, M., and Maszk, P. (1996). The relations of children's dispositional empathy-related responding to their emotionality, regulation, and social functioning. *Dev. Psychol.* 32, 195–209. doi: 10.1037/0012-1649.32.2.195

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91–119. doi: 10.1037/0033-2909.101.1.91

Eisenberg, N., and Spinrad, T. L. (2014). "Multidimensionality of prosocial behavior," in *Prosocial Development: A Multidimensional Approach*, eds L. M. Padilla-Walker and G. Carlo (Oxford: Oxford University Press), 17–39. doi: 10.1093/acprof:oso/9780199964772.003.0002

Elsner, C., Bakker, M., Rohlfing, K., and Gredebäck, G. (2014). Infants' online perception of give-and-take interactions. *J. Exp. Child Psychol.* 126, 280–294. doi: 10.1016/j.jecp.2014.05.007

Falck-Ytter, T. (2008). Face inversion effects in autism: a combined looking time and pupillometric study. *Autism Res.* 1, 297–306. doi: 10.1002/aur.45

Falck-Ytter, T., Bölte, S., and Gredebäck, G. (2013). Eye tracking in early autism research. *J. Neurodev. Disord.* 5, 28. doi: 10.1186/1866-1955-5-28

Falck-Ytter, T., Carlström, C., and Johansson, M. (2015). Eye contact modulates cognitive processing differently in children with autism. *Child Dev.* 86, 37–47. doi: 10.1111/cdev.12273

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.* 9, 878–879. doi: 10.1038/nn1729

Fawcett, C., and Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Dev. Sci.* 16, 841–849. doi: 10.1111/desc.12074

Frank, M. C., Vul, E., and Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy* 17, 355–375. doi: 10.1111/j.1532-7078.2011.00086.x

Geangu, E., Hauf, P., Bhardwaj, R., and Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE* 6:e27132. doi: 10.1371/journal.pone.0027132

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychol. Bull.* 77, 340–355. doi: 10.1037/h0032456

Gredebäck, G., Fikke, L., and Melinder, A. (2010). The development of joint visual attention: a longitudinal study of gaze following during interactions with mothers and strangers. *Dev. Sci.* 13, 839–848. doi: 10.1111/j.1467-7687.2009.00945.x

Gredebäck, G., Johnson, S., and von Hofsten, C. (2009). Eye tracking in infancy research. *Dev. Neuropsychol.* 35, 1–19. doi: 10.1080/87565640903325758

Gredebäck, G., and Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition* 114, 197–206. doi: 10.1016/j.cognition.2009.09.004

Gredebäck, G., and Melinder, A. (2011). Teleological reasoning in 4-month-old infants: pupil dilations and contextual constraints. *PLoS ONE* 6:e26487. doi: 10.1371/journal.pone.0026487

Grossmann, T., and Johnson, M. H. (2007). The development of the social brain in human infancy. *Eur. J. Neurosci.* 25, 909–919. doi: 10.1111/j.1460-9568.2007.05379.x

Heckhausen, H. (1987). "Emotional components of action: their ontogeny as reflected in achievement behavior," in *Curiosity, Imagination and Play: On the Development of Spontaneous Cognitive and Motivational Processes*, eds D. Görlitz and J. F. Wohlwill (Hillsdale, NJ: Lawrence Erlbaum), 326–348.

Heckhausen, J. (1988). Becoming aware of one's competence in the second year: developmental progression within the mother child dyad. *Int. J. Behav. Dev.* 11, 305–326. doi: 10.1177/016502548801100302

Heller, P. H., Perry, F., Jewett, D. L., and Levine, J. D. (1990). Autonomic components of the human pupillary light reflex. *Invest. Ophthalmol. Vis. Sci.* 31, 156–162.

Henderson, R. R., Bradley, M. M., and Lang, P. J. (2014). Modulation of the initial light reflex during affective picture viewing. *Psychophysiology* 51, 815–818. doi: 10.1111/psyp.12236

Hepach, R., Vaish, A., and Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychol. Sci.* 23, 967–972. doi: 10.1177/0956797612440571

Hepach, R., Vaish, A., and Tomasello, M. (2013). A new look at children's prosocial motivation. *Infancy* 18, 67–90. doi: 10.1111/j.1532-7078.2012.00130.x

Hepach, R., and Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *J. Exp. Child Psychol.* 115, 16–29. doi: 10.1016/j.jecp.2012.12.013

Horberg, E. J., Kraus, M. W., and Keltner, D. (2013). Pride displays communicate self-interest and support for meritocracy. *J. Pers. Soc. Psychol.* 105, 24–37. doi: 10.1037/a0032849

Jackson, I., and Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Dev. Sci.* 12, 670–679. doi: 10.1111/j.1467-7687.2008.00805.x

Kahneman, D., Tursky, B., Shapiro, D., and Crider, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *J. Exp. Psychol.* 79, 164–167. doi: 10.1037/h0026952

Karmiloff-Smith, A., Casey, B. J., Massand, E., Tomalski, P., and Thomas, M. S. (2014). Environmental and genetic influences on neurocognitive development

the importance of multiple methodologies and time-dependent intervention. *Clin. Psychol. Sci.* 2, 628–637. doi: 10.1177/2167702614521188

Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., and Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: genetic and environmental contributions. *Emotion* 8, 737–752. doi: 10.1037/a0014179

Laeng, B., Sirois, S., and Gredebäck, G. (2012). Pupillometry a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi: 10.1177/1745691611427305

Lewis, M. (1997). The self in self-conscious emotions. *Ann. N. Y. Acad. Sci.* 818, 119–142. doi: 10.1111/j.1749-6632.1997.tb48251.x

Lewis, M., Alessandri, S. M., and Sullivan, M. W. (1992). Differences in shame and pride as a function of children's gender and task difficulty. *Child Dev.* 63, 630–638. doi: 10.2307/1131351

Libby, W. L., Lacey, B. C., and Lacey, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology* 10, 270–294. doi: 10.1111/j.1469-8986.1973.tb00526.x

Liu, C., and LaFreniere, P. (2014). The effects of age-mixing on peer cooperation and competition. *Hum. Ethol. Bull.* 29, 4–17.

Loewenfeld, I. E. (1993). *The Pupil: Anatomy, Physiology, and Clinical Applications.* Detroit, MI: Wayne State University Press.

Loewenstein, O. (1920). Experimentelle beitrage zur lehre von den katatonischen pupillenveranderungen. *Monatschr. Psychiatr. Neurol.* 47, 194–215. doi: 10.1159/000190690

Luna-Oliva, L., Ortiz-Gutiérrez, R. M., Cano-de la Cuerda, R., Piédrola, R. M., Alguacil-Diego, I. M., Sánchez-Camarero, C., et al. (2013). Kinect Xbox 360 as a therapeutic modality for children with cerebral palsy in a school environment: a preliminary study. *NeuroRehabilitation* 33, 513–521. doi: 10.3233/NRE-131001

Martens, J. P., and Tracy, J. L. (2013). The emotional origins of a social learning bias does the pride expression cue copying? *Soc. Psychol. Pers. Sci.* 4, 492–499. doi: 10.1177/1948550612457958

Montepare, J. M., Goldstein, S. B., and Clausen, A. (1987). The identification of emotions from gait information. *J. Nonverbal Behav.* 11, 33–42. doi: 10.1007/BF00999605

Nagai, Y., Nakatani, A., Qin, S., Fukuyama, H., Myowa-Yamakoshi, M., and Asada, M. (2012). "Co-development of information transfer within and between infant and caregiver," in *Proceedings of the 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, San Diego, CA: IEEE, 1–6. doi: 10.1109/DevLrn.2012.6400879

Navab, A., Gillespie-Lynch, K., Johnson, S. P., Sigman, M., and Hutman, T. (2012). Eye-tracking as a measure of responsiveness to joint attention in infants at risk for autism. *Infancy* 17, 416–431. doi: 10.1111/j.1532-7078.2011.00082.x

Nyström, P., Gredebäck, G., Bölte, S., and Falck-Ytter, T. (2015). Hypersensitive pupillary light reflex in infants at risk for autism. *Mol. Autism* 6, 10. doi: 10.1186/s13229-015-0011-6

Oakes, L. M. (2012). Advances in eye tracking in infancy research. *Infancy* 17, 1–8. doi: 10.1111/j.1532-7078.2011.00101.x

Paulus, M. (2014). The emergence of prosocial behavior: why do infants and toddlers help, comfort, and share? *Child Dev. Perspect.* 8, 77–81. doi: 10.1111/cdep.12066

Rothbart, M. K., Ahadi, S. A., and Evans, D. E. (2000). Temperament and personality: origins and outcomes. *J. Pers. Soc. Psychol.* 78, 122–135. doi: 10.1037/0022-3514.78.1.122

Salley, B., Miller, A., and Bell, M. A. (2013). Associations between temperament and social responsiveness in young children. *Infant Child Dev.* 22, 270–288. doi: 10.1002/icd.1785

Schwartz, B., Tesser, A., and Powell, E. (1982). Dominance cues in nonverbal behavior. *Soc. Psychol. Q.* 45, 114–120. doi: 10.2307/3033934

Shariff, A. F., and Tracy, J. L. (2009). Knowing who's boss: implicit perceptions of status from the nonverbal expression of pride. *Emotion* 9, 631–639. doi: 10.1037/a0017089

Shiota, M. N., Campos, B., and Keltner, D. (2003). The faces of positive emotion. *Ann. N. Y. Acad. Sci.* 1000, 296–299. doi: 10.1196/annals.1280.029

Shum, H. P., Ho, E. S., Jiang, Y., and Takagi, S. (2013). Real-time posture reconstruction for Microsoft Kinect. *IEEE Trans. Cybern.* 43, 1357–1369. doi: 10.1109/TCYB.2013.2275945

Sirois, S., and Brisson, J. (2014). Pupillometry. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 679–692. doi: 10.1002/wcs.1323

Sirois, S., and Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy* 17, 61–78. doi: 10.1111/j.1532-7078.2011.00096.x

Smith, L., Yu, C., Yoshida, H., and Fausey, C. M. (2014). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *J. Cogn. Dev.* doi: 10.1080/15248372.2014.933430

Steinhauer, S. R., Condray, R., and Kasparek, A. (2000). Cognitive modulation of midbrain function: task-induced reduction of the pupillary light reflex. *Int. J. Psychophysiol.* 39, 21–30. doi: 10.1016/S0167-8760(00)00119-7

Stommel, M., Beetz, M., and Xu, W. (2015). Model-Free detection, encoding, retrieval, and visualization of human poses from kinect data. *IEEE Trans. Mechatron.* 20, 865–875. doi: 10.1109/TMECH.2014.2322376

Svetlova, M., Nichols, S. R., and Brownell, C. A. (2010). Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Dev.* 81, 1814–1827. doi: 10.1111/j.1467-8624.2010.01512.x

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., and Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: a longitudinal eye-tracking study. *Infancy* 18, 534–553. doi: 10.1111/j.1532-7078.2012.00135.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/S0140525X05000129

Tracy, J. L., and Robins, R. W. (2004). Show your pride evidence for a discrete emotion expression. *Psychol. Sci.* 15, 194–197. doi: 10.1111/j.0956-7976.2004.01503008.x

Tracy, J. L., Robins, R. W., and Lagattuta, K. H. (2005). Can children recognize pride? *Emotion* 5, 251–257. doi: 10.1037/1528-3542.5.3.251

Troseth, G. L., and DeLoache, J. S. (1998). The medium can obscure the message: young children's understanding of video. *Child Dev.* 69, 950–965. doi: 10.1111/j.1467-8624.1998.tb06153.x

Walczak, N., Fasching, J., Toczyski, W. D., Morellas, V., Sapiro, G., and Papanikolopoulos, N. (2013). "Locating occupants in preschool classrooms using a multiple RGB-D sensor system," in *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS), Tokyo: IEEE, 2166–2172. doi: 10.1109/IROS.2013.6696659

Wang, J. T. Y. (2011). "Pupil dilation and eye tracking," in *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*, eds M. Schulte-Mecklenbeck, A. Kühberger, and R. Ranyard (New York, NY: Psychology Press), 185–204.

Warga, M., Lüdtke, H., Wilhelm, H., and Wilhelm, B. (2009). How do spontaneous pupillary oscillations in light relate to light intensity? *Vision Res.* 49, 295–300. doi: 10.1016/j.visres.2008.09.019

Warneken, F. (2015). Precocious prosociality: why do young children help? *Child Dev. Perspect.* 9, 1–6. doi: 10.1111/cdep.12101

Warneken, F., and Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends Cogn. Sci.* 13, 397–402. doi: 10.1016/j.tics.2009.06.008

Weisfeld, G. E., and Beresford, J. M. (1982). Erectness of posture as an indicator of dominance or success in humans. *Motiv. Emot.* 6, 113–131. doi: 10.1007/BF00992459

Wilhelm, H. (1991). *Pupillenreaktionen – Pupillenstörungen*. Stuttgart: Kohlhammer.

Won, A. S., Bailenson, J. N., Stathatos, S. C., and Dai, W. (2014). Automatically detected nonverbal behavior predicts creativity in collaborating dyads. *J. Nonverbal Behav.* 38, 389–408. doi: 10.1007/s10919-014-0186-0

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., and Chapman, M. (1992). Development of concern for others. *Dev. Psychol.* 28, 126–136. doi: 10.1037/0012-1649.28.1.126

frontiers
in Psychology

Intra-individual variability and continuity of action and perception measures in infants

Anja Gampe[1]*, Anne Keitel[2] and Moritz M. Daum[1]

[1] *Department of Psychology, University of Zürich, Zürich, Switzerland,* [2] *Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK*

The development of action and perception, and their relation in infancy is a central research area in socio-cognitive sciences. In this Perspective Article, we focus on the developmental variability and continuity of action and perception. At group level, these skills have been shown to consistently improve with age. We would like to raise awareness for the issue that, at individual level, development might be subject to more variable changes. We present data from a longitudinal study on the perception and production of contralateral reaching skills of infants aged 7, 8, 9, and 12 months. Our findings suggest that individual development does not increase linearly for action or for perception, but instead changes dynamically. These non-continuous changes substantially affect the relation between action and perception at each measuring point and the respective direction of causality. This suggests that research on the development of action and perception and their interrelations needs to take into account individual variability and continuity more progressively.

Keywords: action, perception, infancy, variability, continuity

Action and Perception in Development

Everyday social interactions involve the production of one's own actions and the perception of actions performed by others (henceforth referred to as action and perception). In the last two decades, a great amount of research has shown that action and perception are mutually related (e.g., Prinz, 1997) and focused on the particular influence of action on perception and vice versa. It has been shown that the perception of others' actions is improved in those with their own action abilities (e.g., Hamilton et al., 2004; Calvo-Merino et al., 2005), and that observing others' actions influences subsequent own action execution (e.g., Craighero et al., 2002; Kilner et al., 2003). This relation between action and perception is especially interesting from a developmental perspective, because during the first months of life infants are about to develop both action and perception skills. It is thus considered possible to disentangle the relative contributions of action and perception for the development of a mutual link. However, there is an ongoing debate about the temporal order of action and perception development, thus whether infants have to be able to perform an action before they can understand it or vice versa (Hauf et al., 2007). Concerning the mutual relation, some studies suggest that a link between action and perception is already present early in life (e.g., Nyström, 2008; van Elk et al., 2008; Kanakogi and Itakura, 2011; Ambrosini et al., 2013). For example, Daum et al. (2011) have found a correspondence between 6-month-old infants' grasping skills (palmar vs. thumb opposition) and their differentiation between expected and unexpected grasping actions (longer looking times toward incongruent grasping actions, i.e., large hand aperture for small objects and vice versa). Studies measuring anticipatory gaze have found that between 4 and 10 months

of age, one-handed grasping was correlated with gaze latency toward the goal of human grasping actions (Kanakogi and Itakura, 2011). Melzer et al. (2012) used a combined perception-action task to investigate the development of contralateral reaching in infants at 6 and 12 months. In the perception task, videos of either contralaterally or ipsilaterally grasped and transported objects were presented and anticipatory gaze behavior was analyzed. In the action task, infants' ipsi- and contralateral reaching behavior toward toys was analyzed to see how often they already reached contralaterally. At 12 months, infants' anticipation of contralateral actions was correlated with their contralateral reaching skills (Melzer et al., 2012). This correlation was not yet evident in 6-month-old children. The above-mentioned studies suggest a link between action and perception in infancy, although the occurrence varies with respect to age and the particular action. Importantly, the state of evidence is not homogenous. When investigating different abilities at different measuring points, different conclusions on the strength and the causality between action and perception are claimed. Some authors suggest that there is an immediate link between action and perception as soon as an action can be produced (Sommerville et al., 2005; Kanakogi and Itakura, 2011; Ambrosini et al., 2013). Others suggest that active experience with an action is necessary before it is linked to perception (cf. Cannon et al., 2012; Melzer et al., 2012). And still other studies report that perception develops to some extent independently of action abilities (Gergely et al., 1995; Hofstadter and Reznick, 1996; Hofer et al., 2005; Biro and Leslie, 2007). Sometimes even the same lab shows a link between action and perception in one study (grasping; Bakker et al., 2014) but not in another (pointing; Gredebäck et al., 2010).

But where do these contradictory results derive from? Potential factors include the designs used, the abilities looked at, the measures calculated, or the age group investigated. In this Perspective Article, we argue that one important but previously neglected factor is the nature of developmental processes: Often, the implicit assumption is that development is more or less continuous. But do abilities really improve steadily and linearly? There is much evidence that, at group level, action and perception skills consistently improve with age (Van der Fits et al., 1999; Hofer et al., 2005; Falck-Ytter et al., 2006; Kanakogi and Itakura, 2011; Ambrosini et al., 2013; Keitel et al., 2013, 2014; Gampe and Daum, 2014). The group level results of Melzer et al. (2012) showed, for example, both an increase in contralateral reaching and an increase in anticipations of contralateral movements between 6 and 12 months. But less is known about the particular shape of developmental trajectories at the individual level. For example, dynamic systems theory suggests that individual development might look quite different from average group development (Thelen and Smith, 2007). According to this approach, abilities self-organize and adapt to their surroundings dynamically (Smith and Thelen, 1993). Behavior emerges as a result of the relationships between abilities. Importantly, abilities are not linearly bound, which means that a small change in one single ability can result in a transformation of the whole system.

The only possibility to investigate individual development is to collect longitudinal data on action and perception skills in infants, and correlate these measures over developmental time. If individual development is linear, good performance at one measuring point should surely entail good performance at another. Such a consistency should also result in high correlations for action and/or perception measures at different measuring points within and between domains. In this Perspective Article, we argue that this is often not the case, and present supporting data from one longitudinal study.

Longitudinal Data on Action and Perception Development

To substantiate our argument, we tested the intra-individual variability and continuity of perception and action in infancy. To this end, we tested 25 infants longitudinally at 7, 8, 9, and 12 months of age (see **Figure 1A** for details), using the action-perception paradigm developed by Melzer et al. (2012), in which perception and production of contralateral grasping movements were measured.

In the perception task (see **Figure 1B** for details), children observed videos of an actor grasping a ball (either ipsilaterally or contralaterally) and transporting it into a bucket (either contralaterally or ipsilaterally). The frequency of anticipatory gaze shifts toward the goal of contralateral movements was used as a performance measure. In the action task (see **Figure 1C** for details), the children's ability to reach contralaterally was tested. The frequency of contralateral responses produced toward a presented toy was used as a measure of action. Action and perception measures were both expressed in per cent, which makes them easily comparable.

At group level, we found similar action and perception abilities to those in the original study (see **Figure 2A** for individual and group means). The anticipation frequency increased from $M_{7\,months} = 16.8 \pm 22.4\%$ ($\pm$ SD) to $M_{12\,months} = 64.2 \pm 28.6\%$ (Melzer et al., 2012: $M_{6\,months} = 19.1 \pm 3.2\%$, $M_{12\,months} = 61.8 \pm 3.8\%$). The frequency of contralateral reaching increased from $M_{7\,months} = 18.2 \pm 14.8\%$ to $M_{12\,months} = 34.3 \pm 18.1\%$ (Melzer et al., 2012: $M_{6\,months} = 18.9 \pm 15.9\%$ to $M_{12\,months} = 30.7 \pm 15.4\%$).

However, we were interested in a systematic evaluation of the continuity of action and perception measures at group level (linear regression) and at individual level (correlations). To this end, we ran linear regression analyses for action and perception with age in days as the between-subject factor. Performance increased linearly for action, $R^2 = 0.09$, $p = 0.007$, and for perception, $R^2 = 0.25$, $p < 0.001$. The regression coefficients for action and perception differed significantly, $t = 2.15$, $p = 0.03$, suggesting that age is a stronger predictor for perception than for action. A steeper increase in perception abilities than in action abilities is thus evident. *Post hoc* Bonferroni-corrected *t*-tests with all infants revealed that, for action, performance differed significantly between 7 and 12 months ($p = 0.003$). For perception, performance differed between the following age groups: 7–9: $p = 0.002$; 7–12: $p < 0.001$; 8–9: $p = 0.009$; 8–12: $p < 0.001$, but not for 7–8 and 9–12 months.

In a second step, we looked at the correlations within a performance measure between measuring points to see whether the abilities also increase linearly at individual level (**Figure 2B**, yellow and red bars). More precisely, we correlated perception at all
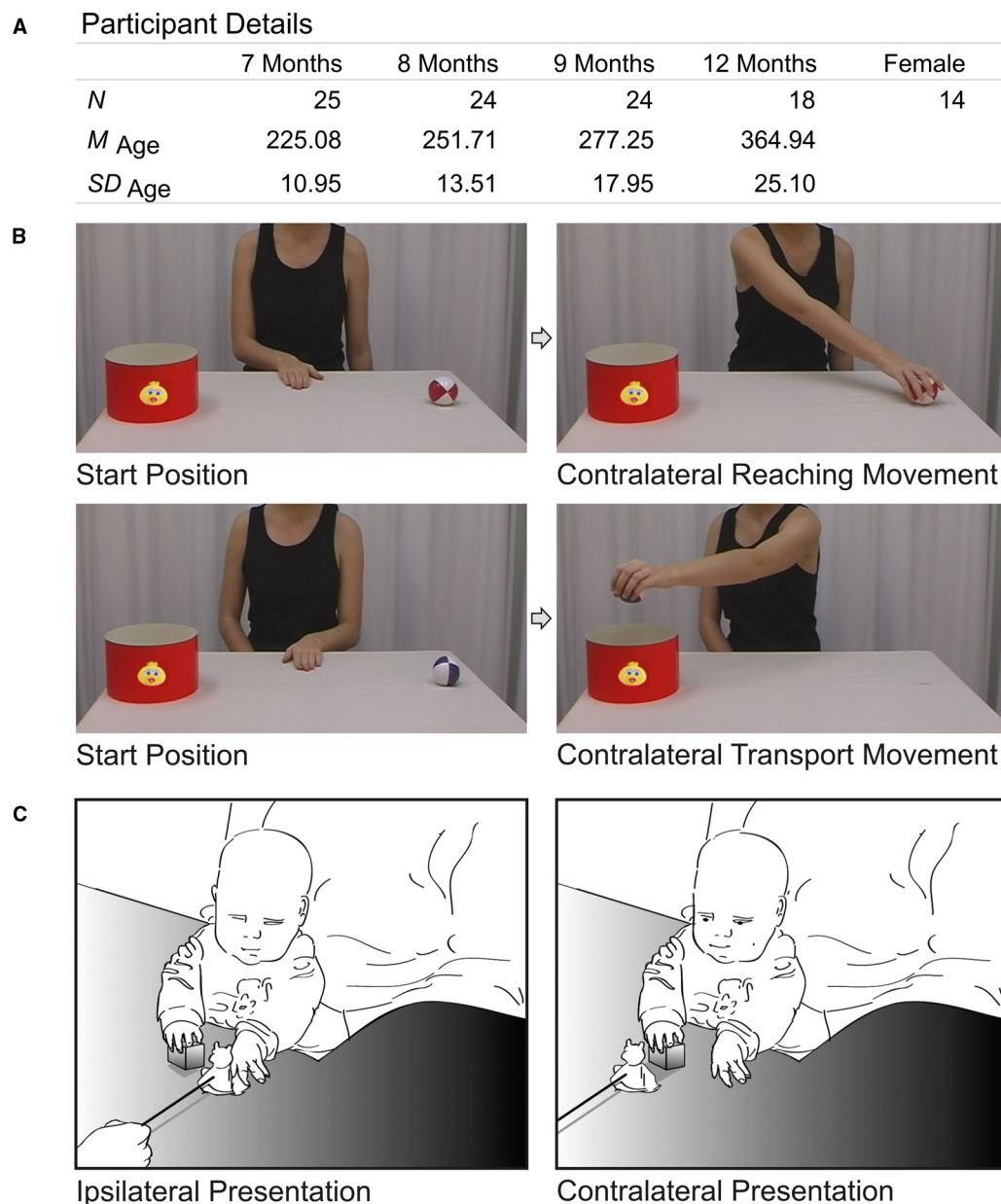
FIGURE 1 | Methods. (A) Participant details, including number of participants per measuring point, mean age in days (and standard deviation), and number of females in the sample. **(B)** Screen shots of the perception task. Upper row illustrates a trial with a contralateral reaching movement (and, therefore, an ipsilateral transport movement, not shown). Lower row illustrates a trial with a contralateral transport movement (and preceding ipsilateral reaching movement, not shown). While observing the actions, infants' eye movements were recorded using an eyetracker (SR Research Eyelink Plus, 500 Hz, monocular). As perception measure, we calculated the anticipation frequency for contralateral movements, that is, the number of trials in which the gaze arrived at goal areas (i.e., ball AOI and bucket AOI)

prior to the model's hand, divided by all actions that were perceived. We used the same criteria for analysis as in the original study (Melzer et al., 2012). **(C)** Illustration of the action task. We presented similar toys on sticks as Melzer et al. (2012), while the child held a small cube in one hand. The toys were either presented to the empty hand of the child (to elicit an ipsilateral reaction) or to the occupied hand (to elicit a contralateral reaction with the empty hand). As a measure of action, we calculated the ratio between the performed contralateral grasping movements (interrater-reliability $\kappa = 0.93$) and presented contralateral trials: $N_{contralateral\_grasped}/N_{contralateral\_presented}$. Licenses for re-used figures from Melzer et al. (2012) have been obtained.

different measuring points, and action at all different measuring points. If individual development was linear, there should be significant correlations between different measuring points, with the highest correlation coefficients at adjacent measuring points.

To get a better estimation of the correlation, we bootstrapped the pairwise correlations and calculated the correlation bias and the standard error of the mean correlation coefficient. The p-values were corrected for multiple testing (FDR correction) and showed
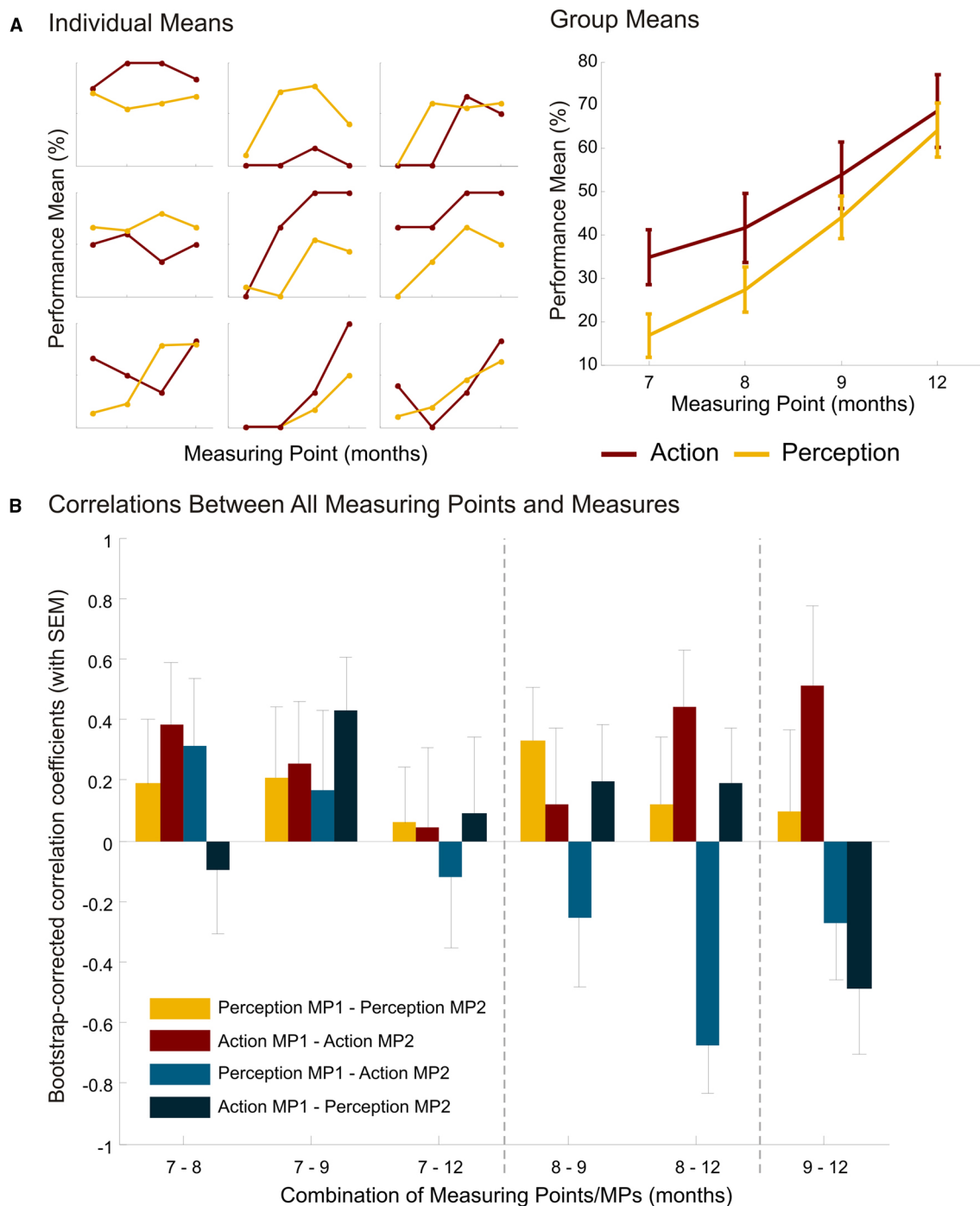
FIGURE 2 | Results. (A) Mean performance for action and perception (in %) of nine individual children and of the group with standard error of the mean at the different measuring points (7, 8, 9, and 12 months of age). Individual data displayed is of the children who provided action and perception measures at all measuring points. Although the number of participants was initially 25, only 18 children participated at age 12 months. Of those 18 who participated at all measuring points, only 9 consistently provided performance measures in both tasks. Note that the y-axis for each individual plot is scaled from 0 to 100%. **(B)** Bootstrap-corrected correlation coefficients (with SEM) for perception, action, perception-action and action-perception between measuring points (MPs) 7–8, 7–9, 7–12, 8–9, 8–12, and 9–12 months.

that none of the perception abilities (all $p > 0.67$) and none of the action abilities (all $p > 0.24$) were correlated with the same ability at another measuring point.

A further analysis targeted questions of the temporal order of action and perception. Are we able to perform actions ourselves only after having understood other people's actions, or do we need

own action abilities for observational understanding of others? We again calculated bivariate, bootstrap-corrected correlations with FDR-corrected p-values between perception ability at one measuring point and action ability at another measuring point and vice versa (**Figure 2B**, light and dark blue bars). Correlations yielded no significance for action predicting perception at any measuring point (all $p > 0.20$). Perception at 8 months negatively predicted action at 12 months ($r = -0.659$, $p < 0.05$); and no other significant predictions from perception to action (all other $p > 0.48$).

Finally we looked at the relations between action and perception at one measuring point, as was done in the original study. We did not find a correlation between action and perception measures at 12 months of age, nor at any other age tested. The highest correlation we found is $r = 0.33$ at 7 months of age, which did not reach significance ($p = 0.16$).

The longitudinal data presented illustrate two points: First, action and perception increase linearly at group level but not at individual level. Second, correlations between action and perception within and between measuring points are unstable and transient. Not one level of abilities relates to its ability at a later stage, although abilities at group level increase steadily. The relations between the domains are of different strengths at different points in time and between points in time. Together, these findings suggest that individual development does not take place linearly, but might depend on various interactions of specific abilities within the child, which affects performance at any given time.

Action and Perception Development within a Dynamic System

This idea is congruent with the view that abilities self-organize and adapt to their surroundings dynamically as proposed, for example, by the dynamic systems approach (Smith and Thelen, 1993). When looking at the longitudinal results presented above, it appears that, in contrast to group level, at individual level perception and action do not develop in a continuous manner, but rather in developmental trajectories that differ greatly between individuals (for a discussion of a variety of individual developmental trajectories, see, e.g., Adolph et al., 2008). The present findings add to this knowledge that, resulting from these individual differences, the relation between perception and action is not one of continuous stability but also subject to fluctuations over age. Transferred to system dynamics, this means that action and perception abilities are themselves the result of relationships with other abilities that can change at any moment. How each ability develops over time therefore depends on various interactions with other abilities within each infant. As a consequence, no individual correlation was found within one domain over the measuring points. Some of the infants improved in comparison to the last measure, while others remained constant or declined. At group level, a linear increase can be observed because performance increases in more infants than it decreases. And even within action and perception abilities a small, but critical change in one sub-system can cause the whole system to shift, resulting in a new action or perception behavior. This way, the strength of the relation between action and

perception and predictive power in different measuring points varied enormously.

Consequences and Possible Solutions

The most important message of the above findings and theoretical considerations is that a cautious interpretation should be made when relations and especially temporal order of action and perception are investigated in infants. Unsteady individual development can make a replication of results difficult, which is evident in the heterogeneity of previous findings, as well as in the discrepancy between the current data and Melzer et al. (2012). Although we found the same level of abilities in action and perception at group level, we were unable to replicate the interrelation between them. One reason might be the difference in design (cross-sectional vs. longitudinal), another might be the sample size for infants who provided data in action and perception measures at each measuring point ($N_{7\ months} = 20$; $N_{8\ months} = 24$; $N_{9\ months} = 20$; $N_{12\ months} = 14$ vs. $N = 24$ in the original study). But as we have replicated the results at group level, it also seems plausible that system dynamics might account for the missing relation. Abilities in dynamic systems are unstable and unpredictable in transition phases (Lewis, 2000). As a result, some studies will find no relations while others might see incidental relations. Non-linear individual development could, consequently, also cause non-linear results at group level (van der Maas and Molenaar, 1992; van Geert, 1994). This is rarely found in published data, although this could be due to the fact that researchers usually expect continuous results, and do not attempt to publish erratic data (but see Keitel and Daum, 2015). Answers to simple questions of temporal order or functional relations between action and perception cannot therefore be unidimensional but depend on the age group chosen, the distance between measures and the domains and abilities looked at.

There are some methodological precautions one could take to ensure that an interpretation of findings is reliable, at least to some extent. For example, sample size should be large enough to accurately reflect the population, going beyond the 10–12 children per group sometimes reported (Gredebäck et al., 2010; Kanakogi and Itakura, 2011; Ambrosini et al., 2013). A large number of trials helps to yield the most reliable results, although this might not always be easy to achieve with infants. Collecting a larger number of trials offers the possibility to compute system dynamics, which in turn might offer better pathways in understanding changes in development and the relations between different components (Spivey and Dale, 2006; Reddy et al., 2013). Non-linear analyses have the strength to better capture the complexity of each individual. Non-linearity underscores the observation that behaviors are not proportional to their causes (Carver and Scheier, 1998). The outcome behavior might appear chaotic and noiselike where it is in fact deterministic and predictable (Heath et al., 2000). One easily applicable method for computing non-linear system dynamics is recurrence quantification analysis (RQA). RQA quantifies aspects of the temporal evolution of a collected time series, such as its predictability, variability, or repetitiveness (Webber and Zbilut, 2007). For example, Reddy et al. (2013) analyzed infants' force data when being picked up by their mothers and found that

3-month-olds already showed anticipatory adjustments to the approach of their mother's arms. We applied RQA to the perception measure of the data presented above and computed the recurrence rate of shifts to the goal location. Next, we correlated the recurrence rates at the different measuring points to look at individual stability and continuity. The analyses revealed that system dynamics are stable within the individual between three measuring points, 7–12: $r = 0.628$, $p = 0.009$; 8–9: $r = 0.413$, $p = 0.045$; 9–12: $r = 0.444$, $p = 0.05$. Thus, measures that take into account non-linearity may possibly reveal reliable developmental interrelations in infants (for more examples of non-linear analyses, see Giese et al., 1996; Boker et al., 1998; Taga et al., 1999; Deffeyes et al., 2009). Furthermore, there are other non-linear analyses that could meet the obvious non-linear characteristics of development, like fractality and 1/f (for an introduction to the different non-linear measures and calculations, see Heath et al., 2000; Riley and van Orden, 2005; Holden et al., 2013). These kinds of analyses can complement traditional analyses and might eventually lead to a better understanding of children's development.

What is equally important is to run longitudinal studies when aiming at investigating developmental changes in certain abilities. The heterogeneity in individual development can tell us more about the mechanisms than a cross-sectional growth curve does (Jenni et al., 2013; Lindenberger et al., 2013).

To conclude, we presented theoretical considerations and supporting data that imply inconsistency and discontinuity of individual action and perception skills in infancy. Even though there are some precautions one could take to address this individual discontinuity, we believe that no definite conclusions can be drawn about the development of the link between action and perception in infancy. More precisely, with current methodological standards, there can be no accurate interpretation about the time when a link between action and perception is established, or which ability develops first. The nature of individual discontinuity results in the fact that some samples will show incidental correlations, while others will not. In our opinion, valid conclusions can only be achieved by applying a multi-method approach in order to better capture individual variance in development.

References

Adolph, K. E., Robinson, S. R., Young, J. W., and Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychol. Rev.* 115, 527–543. doi: 10.1037/0033-295X.115.3.527

Ambrosini, E., Reddy, V., de Looper, A., Costantini, M., Lopez, B., and Sinigaglia, C. (2013). Looking ahead: anticipatory gaze and motor ability in infancy. *PLoS ONE* 8:e67916. doi: 10.1371/journal.pone.0067916

Bakker, M., Daum, M. M., Handl, A., and Gredebäck, G. (2014). Neural correlates of action perception at the onset of functional grasping. *Soc. Cogn. Affect. Neurosci.* nsu119. doi: 10.1093/scan/nsu119 [Epub ahead of print].

Biro, S., and Leslie, A. M. (2007). Infants' perception of goal-directed actions: development through cue-based bootstrapping. *Dev. Sci.* 10, 379–398. doi: 10.1111/j.1467-7687.2006.00544.x

Boker, S. M., Schreiber, T., Pompe, B., and Bertenthal, B. I. (1998). "Nonlinear analysis of perceptual-motor coupling in the development of postural control," in *Nonlinear Analysis of Physiological Data*, eds H. Kantz, J. Kurths, and G. Mayer-Kress (Berlin: Springer), 251–269.

Calvo-Merino, B., Glaser, D. E., Passingham, R. E., and Haggard, P. (2005). Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* 15, 1243–1249. doi: 10.1093/cercor/bhi007

Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., and Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Dev. Sci.* 15, 35–42. doi: 10.1111/j.1467-7687.2011.01095.x

Carver, C. S., and Scheier, M. F. (1998). *On the Self-Regulation of Behavior*. Cambridge: Cambridge University Press.

Craighero, L., Bello, A., Fadiga, L., and Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia* 40, 492–502. doi: 10.1016/S0028-3932(01)00134-8

Daum, M. M., Prinz, W., and Aschersleben, G. (2011). Perception and production of object-related grasping movements in 6-month-old infants. *J. Exp. Child Psychol.* 108, 810–818. doi: 10.1016/j.jecp.2010.10.003

Deffeyes, J. E., Harbourne, R. T., Kyvelidou, A., Stuberg, W. A., and Stergiou, N. (2009). Nonlinear analysis of sitting postural sway indicates developmental delay in infants. *Clin. Biomech. (Bristol, Avon)* 24, 564–570. doi: 10.1016/j.clinbiomech.2009.05.004

Falck-Ytter, T., Gredebäck, G., and von Hofsten, C. (2006). Infants predict other people's action goals. *Nat. Neurosci.* 9, 878–879. doi: 10.1038/nn1729

Gampe, A., and Daum, M. M. (2014). Productive verbs facilitate action prediction in toddlers. *Infancy* 19, 301–325. doi: 10.1111/infa.12047

Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193. doi: 10.1016/0010-0277(95)00661-H

Giese, M. A., Dijkstra, T. M. H., Schöner, G., and Gielen, C. C. (1996). Identification of the nonlinear state-space dynamics of the action-perception cycle for visually induced postural sway. *Biol. Cybern.* 74, 427–437. doi: 10.1007/BF00206709

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

Hamilton, A., Wolpert, D., and Frith, U. (2004). Your own action influences how you perceive another person's action. *Curr. Biol.* 14, 493–498. doi: 10.1016/j.cub.2004.03.007

Hauf, P., Aschersleben, G., and Prinz, W. (2007). Baby do—baby see! How action production influences action perception in infants. *Cogn. Dev.* 22, 16–32. doi: 10.1016/j.cogdev.2006.09.002

Heath, R. A., Kelly, A., and Longstaff, M. (2000). Detecting nonlinearity in psychological data: techniques and applications. *Behav. Res. Methods Instrum. Comput.* 32, 280–289. doi: 10.3758/BF03207796

Hofer, T., Hauf, P., and Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behav. Dev.* 28, 466–480. doi: 10.1016/j.infbeh.2005.04.002

Hofstadter, M., and Reznick, J. S. (1996). Response modality affects human infant delayed-response performance. *Child Dev.* 67, 646–658. doi: 10.1111/j.1467-8624.1996.tb01757.x

Holden, J. G., Riley, M. A., Gao, J., and Torre, K. (2013). Fractal analyses: statistical and methodological innovations and best practices. *Front. Physiol.* 4:97. doi: 10.3389/fphys.2013.00097

Jenni, O. G., Chaouch, A., Caflisch, J., and Rousson, V. (2013). Infant motor milestones: poor predictive value for outcome of healthy children. *Acta Paediatr.* 102, e181–e184. doi: 10.1111/apa.12129

Kanakogi, Y., and Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. *Nat. Commun.* 2, 1–6. doi: 10.1038/ncomms1342

Keitel, A., and Daum, M. M. (2015). The use of intonation for turn anticipation in observed conversations without visual signals as source of information. *Lang. Sci.* 6, 108. doi: 10.3389/fpsyg.2015.00108

Keitel, A., Prinz, W., and Daum, M. M. (2014). Perception of individual and joint action in infants and adults. *PLoS ONE* 9:e107450. doi: 10.1371/journal.pone.0107450

Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C., and Daum, M. M. (2013). Perception of conversations: the importance of semantics and intonation in children's development. *J. Exp. Child Psychol.* 116, 264–277. doi: 10.1016/j.jecp.2013.06.005

Kilner, J., Paulignan, Y., and Blakemore, S. (2003). An interference effect of observed biological movement on action. *Curr. Biol.* 13, 522–525. doi: 10.1016/S0960-9822(03)00165-9

Lewis, M. D. (2000). The promise of dynamic systems approaches for an integrated account of human development. *Child Dev.* 71, 36–43. doi:10.1111/1467-8624.00116

Lindenberger, U., Burzynska, A. Z., and Nagel, I. E. (2013). "Heterogeneity in frontal lobe aging," in *Principles of Frontal Lobe Function*, 2nd Edn, eds D. T. Stuss and R. T. Knight (New York: Oxford University Press), 609–627.

Melzer, A., Prinz, W., and Daum, M. M. (2012). Production and perception of contralateral reaching: a close link by 12 months of age. *Infant Behav. Dev.* 35, 570–579. doi: 10.1016/j.infbeh.2012.05.003

Nyström, P. (2008). The infant mirror neuron system studied with high density EEG. *Soc. Neurosci.* 3, 334–347. doi: 10.1080/17470910701563665

Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154. doi: 10.1080/713752551

Reddy, V., Markova, G., and Wallot, S. (2013). Anticipatory adjustments to being picked up in infancy. *PLoS ONE* 8:e65289. doi: 10.1371/journal.pone.0065289

Riley, M. A., and van Orden, G. C. (2005). *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences*. Washington, DC: National Science Foundation.

Smith, L. B., and Thelen, E. (1993). *A Dynamic Systems Approach to Development: Applications* (Bd. xviii). Cambridge, MA: The MIT Press.

Sommerville, J. A., Woodward, A. L., and Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96, B1–B11. doi: 10.1016/j.cognition.2004.07.004

Spivey, M. J., and Dale, R. (2006). Continuous dynamics in real-time cognition. *Curr. Dir. Psychol. Sci.* 15, 207–211. doi: 10.1111/j.1467-8721.2006.00437.x

Taga, G., Takaya, R., and Konishi, Y. (1999). "Analysis of general movements of infants towards understanding of developmental principle for motor control," in *Proceedings of the 1999 IEEE International Conference on Systems, Man, and Cybernetics, 1999*, Vol. 5, (Tokyo: IEEE SMC '99 Conference Proceedings), 678–683. doi: 10.1109/ICSMC.1999.815633

Thelen, E., and Smith, L. B. (2007). "Dynamic systems theories," in *Handbook of Child Psychology*, eds W. Damon and R. M. Lerner (New York, NY: John Wiley & Sons, Inc.).

Van der Fits, I. B. M., Otten, E., Klip, A. W. J., Van Eykern, L. A., and Hadders-Algra, M. (1999). The development of postural adjustments during reaching in 6- to 18-month-old infants Evidence for two transitions. *Exp. Brain Res.* 126, 517–528. doi: 10.1007/s002210050760

van der Maas, H. L., and Molenaar, P. C. (1992). Stagewise cognitive development: an application of catastrophe theory. *Psychol. Rev.* 99, 395–417. doi: 10.1037/0033-295X.99.3.395

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., and Bekkering, H. (2008). You'll never crawl alone: neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage* 43, 808–814. doi: 10.1016/j.neuroimage.2008.07.057

van Geert, P. (1994). *Dynamic Systems of Development: Change Between Complexity and Chaos* (Bd. xii). Hertfordshire: Harvester Wheatsheaf.

Webber, C. L., and Zbilut, J. P. (2007). Recurrence quantifications: feature extractions from recurrence plots. *Int. J. Bifurcat. Chaos* 17, 3467–3475. doi: 10.1142/S0218127407019226

Developmental pathways for social understanding: linking social cognition to social contexts

Kimberly A. Brink [1], Jonathan D. Lane [2] and Henry M. Wellman [1]*

[1] Department of Psychology, University of Michigan, Ann Arbor, MI, USA, [2] Peabody College of Education and Human Development, Vanderbilt University, Nashville, TN, USA

Contemporary research, often with looking-time tasks, reveals that infants possess foundational understandings of their social worlds. However, few studies have examined how these early social cognitions relate to the child's social interactions and behavior in early development. Does an early understanding of the social world relate to how an infant interacts with his or her parents? Do early social interactions along with social-cognitive understandings in infancy predict later preschool social competencies? In the current paper, we propose a theory in which children's later social behaviors and their understanding of the social world depend on the integration of early social understanding and experiences in infancy. We review several of our studies, as well as other research, that directly examine the pathways between these competencies to support a hypothesized network of relations between social-cognitive development and social-interactive behaviors in the development from infancy to childhood. In total, these findings reveal differences in infant social competences that both track the developmental trajectory of infants' understanding of people over the first years of life and provide external validation for the large body of social-cognitive findings emerging from laboratory looking-time paradigms.

Keywords: infancy, social cognition, theory of mind, continuity, longitudinal predictions

Introduction

Human infants live in a social world and they develop expectations and understandings about people's actions and interactions in that world. Hypothetically, their understandings—their social cognitions—simultaneously shape and are shaped by their social lives and interactions. Moreover, early infant social understanding and interactions should hypothetically shape later social cognition and social behavior in preschool, and beyond.

Figure 1 outlines a theoretical framework for thinking about these developmental transactions, pathways, and achievements. Much is known about several of the topics within each box—e.g., infants' understanding of intentional action, preschoolers' understanding of false belief, and the nature of various parent-child interactions in infancy. The connections between these boxes, however, are only just beginning to be examined. Thus, considering this framework as a guide, our current knowledge is patchy and incomplete. In this article, we aim to help fill in this bigger theoretical picture about how early social cognition is informed by social context and *vice versa*. First, we review and report three illustrative studies of our own that tackle the ways in which early social cognition and social behavior fit together. These studies focus on the links labeled as 1, 2, and 3 in **Figure 1**. Using those studies as anchors, we review other emerging research that addresses these

FIGURE 1 | Hypothesized theoretical framework for comprehensively looking at the pathways between early childhood social cognitions and social behavior. Path 1 is examined in Study 1. Path 2 is examined in isolation in Study 2. Paths 2 and 3 are both assessed in Study 3. Social cognitions and behaviors not explicitly dealt with in this paper appear in parentheses (topics listed are for illustrative purposes and are not exhaustive). Social cognitions and social behaviors outside of parentheses are those we directly address in this paper.

same links. Finally, we briefly discuss portions of this network that still need to be studied.

Study 1: Relations between Intention Understanding and Infants' Larger Social Experiences

Background

Two methodological approaches characterize contemporary research on infant development: looking-time methods that reveal infants' basic understandings of the world, and measures of individual differences that characterize variation among infants in their social contexts (e.g., parent–infant interaction or infant temperament). Surprisingly little research combines both of these methods and perspectives and this is particularly true in research encompassing infant social cognition.

Based on numerous recent studies of infant social cognition, it is now well accepted that during their first year of life infants come to understand that intentional, or goal-directed, states underlie the actions and expressions of others (Baird and Astington, 2005; Sommerville et al., 2005; Akhtar and Martinez-Sussmann, 2007; Tomasello et al., 2005). As noted in **Figure 1**, this includes appreciating agents as intentional actors (engaged in deliberate acts like reaching for and getting things) and as intentional experiencers (experiencing states like desires for, emotions about, and perceptions of things). Although these characterizations rest partly on studies where infants engage in active interaction with others (Carpenter et al., 1998; Saylor et al., 2007; Gräfenhain et al., 2009), the large majority of the relevant studies have used looking-time paradigms where infants look at and react to agents' actions and expressions (e.g., Gergely et al., 1995; Woodward, 1998; Sodian and Thoermer, 2004).

At the same time, much research in the attachment and temperament traditions has examined infant and parent–infant social

actions and interactions. Specifically, attachment research has shown that the sensitivity of mothers' responses to their infants (Rosenblum et al., 2008; Bigelow et al., 2010) is consistently associated with a secure attachment style. Also, infant attachment status predicts and is predicted by aspects of mother–infant responsiveness to each other's actions, communications, and mother–infant affect attunement (e.g., Beebe et al., 2010). These findings suggest that the quality of mother–infant interaction could also lead to enhanced social-cognitive understanding of the sort measured in infant laboratory assessments. Sensitive, responsive, contingent mother–child interactions might sensitize infants to the intentions and desires that underpin such behaviors. Beyond attachment research, infant temperament additionally seems a good candidate to consider because it is widely argued to reflect important individual differences in infancy that further influence infants' experiences and interactions within their social worlds.

Although infant looking-time studies have provided the foundation for many theories of social-cognitive development, these studies rarely consider individual differences (as opposed to age-related differences) in infants' understandings. Thus, they also rarely include analyses of infants' social context, temperament, and parent–infant interaction, and so almost never examine the relations between infants' social experiences or social behavior and their laboratory-based social cognitions. This seems like an important task in its own right and, moreover, such integrative research would also shed light on the ecological validity of the paradigms and findings so often used to study infant social cognition in the laboratory.

In a recent study (Dunphy-Lelii et al., 2014), therefore, we reported on two tasks implemented concurrently when infants were about 12 months old: (1) a traditional looking-time paradigm in which infants witness intentional action-emotion displays, and (2) caregiver–infant interaction episodes. This allowed us to examine relations between social cognition, via looking-time intention-understandings, and social behavior (pathway 1 in **Figure 1**) by taking advantage of individual differences in both. Thus, this first study had the goal of providing insight into relations between social behavior and social cognition (as well as traditional looking-time paradigms and measurements of social interaction) by investigating links between infant habituation to intentional-action displays and their social interactions and social temperaments.

Overview of Study 1

For a looking-time paradigm we used one first reported by Phillips et al. (2002). In habituation trials infants saw a person look at one of two objects with an expression of interest and joy. Then they saw two test events in which the person either reached for the "liked" object (consistent event) or the not-liked object (inconsistent event). By 12–14 months of age most infants (80%) looked longer at the inconsistent test event, a finding consistent with other research demonstrating infants' developing understanding of intentional agents. We chose this task because (a) it is representative of many used to assess intention understanding (e.g., Gergely et al., 1995; Woodward, 1998; Brandone and Wellman, 2009), (b) it involves infant appreciation of both actions (reaching) and emotions (liking), and (c) this same task captures reasonable

individual differences in infants' looking-time as shown in a prior study (Wellman et al., 2004).

To assess infant and mother–infant social interaction differences we included measures of the sort often used to examine the quality of mother–infant interaction in attachment and temperament research. In particular, by observing mother–infant interactions, we assessed the quality of the interaction and the infant's temperamental orientation to and away from other people.

Almost one hundred 10- to 12-month-old infants were seen in a multi-phase research session lasting 40–50 min. After warm up, the protocol began with a looking-time procedure using the paradigm from Phillips et al. (2002) just described. Caregiver and infant then adjourned to a large playroom housing several toys and furniture where they participated in a 16-min session consisting of free play and novel object interaction (95% of the caregivers were mothers, so we will refer to our data as mother–infant interaction).

Social Cognition via Looking-time Measures

Looking-time tasks typically have two phases: a habituation or familiarization phase and a test phase. In social-cognitive research, during infant-controlled habituation, infants view an agent's (or agents') intentional behavior and emotional expressions over multiple trials until they become habituated—for example the agent looks at object-A rather than object-B with interest and pleasure. Then infants see test events that probe their understanding of the original actions—for example, the agent then reaches for object-A (a consistent test event, consistent with the agent liking and wanting that object) or object-B (an inconsistent test event). In this study we focused on individual differences in infants' decrement (or, conversely, maintenance) of attention during habituation to our emotion–action displays. Attention during habituation acts as a measure of infants' ability to parse displays meaningfully and their interest in those displays; that is, differences on this measure represent differences in processing or interest as infants come to a stable impression of the displayed action as being intentional. Infants' attention to intentional-action displays during habituation, as measured by decrement of attention, has been especially revealing in prior research (Wellman et al., 2004; Aschersleben et al., 2008). Our measure (see also, Wellman et al., 2004) essentially involved subtracting infant looking on the last trials of habituation from their looking on the first trials (and then dividing by infant total habituation looking, as recommended by Bornstein and Sigman, 1986). So a higher score means infants' attention decreased sizably and quickly whereas a lower score means infants sustained their attention longer and at higher levels.

We do not focus here on infant differential looking to consistent vs. inconsistent test events because in past social-cognitive research (Wellman et al., 2004; Aschersleben et al., 2008), it has often proven less revealing as an individual-differences measure. More focally, test event looking yielded little of significance in our research as well.

Social Interaction

Measures of mother–infant free play behavior could be collected by coding highly quantified observational tallies of specific acts (numbers of infant reaches, or points, or mother looks at infant per minute). But, more typically research has coded global, aggregated categories (Hofer et al., 2008; Slaughter et al., 2008; Kaitz et al., 2010). Arguably, global aggregates capture important individual differences at a more informative level of analysis (Sroufe and Waters, 1977) and we concentrated on those. Using such aggregates, we focused on four key constructs inspired by findings from attachment, temperament, and social-interaction research: quality of mother–infant interaction, socially observant temperament, joint attention, and imitation.

Because the quality of mother–infant interaction leads to enhanced attachment, it might additionally lead to enhanced social-cognitive understanding of the sort measured in infant laboratory assessments as well; infants who interact contingently with others who are sensitive to their own states and actions would be well situated to notice intentional actions and interactions and thus to achieve a more advanced understanding of them. In our data, *quality of mother–infant interaction* was based on aggregating four sets of global ratings from the free play observations as outlined on the left in **Figure 2**.

We also examined infants' *socially observant temperament* (i.e., an infant's attentiveness to social phenomena). Some aspects of temperament (e.g., activity level) describe infants' motoric proclivities, but others (e.g., social reactivity or avoidance) are more related to infants' social tendencies. Because these latter aspects of temperament can influence a child's experiences and interactions within their social world, they could also impact early social-cognitive understanding. Indeed Wellman et al. (2011) showed that a "socially observant" temperament in early preschool longitudinally predicts enhanced theory of mind 2 years later (see also Lane et al., 2013). This temperament profile, however, was established in work with preschool children. Based on temperament items designed for infants and toddlers (Putnam et al., 2008), we devised items to assess something equivalent in infants. In general, socially observant infant temperament was coded based on whether infants noticed their parent's facial expressions, liked to sit and watch their parents do things, and made talking or vocal sounds when parents talked to them.

Two other relevant social experiences arguably could play supporting roles in infants' emerging understanding of others as intentional agents. *Joint attention* is a hallmark of early cognitive development in which infants begin to coordinate objects into their previously dyadic social interactions—they begin to monitor another's attentional stance toward themselves and the object (see, e.g., Bakeman and Adamson, 1984). Infant *imitation* also has a long history of being studied as both a precursor to, and result of, increased social-cognitive competence. Deliberate infant imitation is evident as early as the second half of the first year (see Bauer and Kleinknecht, 2002), and it is related to both infant language development and mother–infant responsiveness (e.g., Masur and Olson, 2008).

Findings

Many of the numerous aspects of infants' and mothers' behavior in our free-play/teaching scenarios that we coded are listed in **Figure 2**. In that figure, on the left behaviors are organized into *a priori* aggregates that we hypothesized might predict infants' looking-times based on the four constructs we outlined earlier

FIGURE 2 | Two separate sets of aggregates examined in Study 1. The set on the left is organized into four *a priori* aggregates and color-coded. The set on the right is organized into four aggregates that were validated empirically within the study. Specific items from the *a priori* aggregates on the left contributed to the empirically validated aggregates, as indicated by the color of those items, on the right.

(i.e., the quality of mother–infant interaction, socially observant infant temperament, joint attention, and imitation). In initial analyses, we correlated these aggregates with differences in decrement of attention in the looking-time task. In Dunphy-Lelii et al. (2014) we reported similar relations without controlling for age, as all the participants were infants. But these infants ranged in age from 10 to 12 months and the age span between youngest and oldest infants was 10 weeks. This is sizable for infants aged, on average, 49 weeks, so in what we now report, all analyses are controlled for age in days at infant testing.

The four left-column aggregates describing the interactions between mother and child all correlated with our habituation measure when controlling for age: quality of mother–infant interaction, $r(69) = 0.31$, $p < 0.01$; socially observant infant temperament, $r(85) = 0.25$, $p < 0.05$; imitation, $r(72) = 0.27$, $p < 0.05$; and joint attention, $r(70) = 0.24$, $p < 0.05$. Not shown in **Figure 2**, we also coded for infant attentiveness to objects (rather than persons) as an infant behavior that could provide discriminant validity because we predicted it would not correlate with social-cognitive understandings. As expected, that score did not correlate with our habituation measure either initially, $r(86) = -0.03$, $p = 0.77$ or when controlling for age, $r(85) = -0.04$, $p = 0.75$.

Equally important, however, was a more comprehensive empirical analysis of the interactions and behaviors that predict looking-time measures of infant intention understanding. Here we began with the individual items listed on the left in **Figure 2**, but considered their internal organization further. Beginning with factor analysis and then adjusting and deleting items to achieve factors with good internal consistencies, we arrived at the four factors on the right in **Figure 2**. Items such as *infant and maternal affect* and *joint engagement* summed into a highly consistent factor (Factor 1), which we labeled as "action–emotion synchrony." *Infant responsiveness* and *parental non-intrusiveness* loaded highly onto Factor 2, which we called, "mother–infant responsiveness." The items *maternal sensitivity* and *infant gaze following* loaded highly onto Factor 3, which we labeled "mother–infant sensitivity." And finally, several additional items captured a social-temperament factor that we called "infant social attentiveness" (Factor 4).

We entered these four derived factors—those on the right in **Figure 2**—into a regression predicting our primary habituation measure. To test the possibility that something like general object-centered attentiveness or general cognitive "maturity" would account for our findings, we entered the object-attentiveness measure noted earlier as well as age into this regression predicting our primary habituation measure; this step was not significant. We then entered our four key social factors, and those factors accounted for an additional 17% of variance in our habituation

measure, $R^2_{change} = 0.17$, F_{change} (4,63) = 3.31, $p < 0.05$. In this model, the infant social attentiveness aggregate (β = 0.26, $t = 2.01$, $p < 0.05$) and mother–infant responsiveness aggregate (β = 0.25, $t = 2.01$, $p < 0.05$) were independent predictors of social attention during the habituation portion of the looking-time study.

Discussion

These findings indicate that infant social cognition, as demonstrated by performance in infant looking-time procedures, does relate to infants' social interactive behavior. More specifically, individual differences in how infants parse and habituate to intentional action displays within laboratory looking-time research clearly relates to mother–infant interaction patterns, in particular the quality of mother–infant interaction. Additionally, infant looking-time performance also relates to infants' social temperament, in particular, their disposition for attending to another person's social behaviors (e.g., facial expressions and speech). These relations between looking-time performance and aspects of infant social experience provide validity to the assertion that common laboratory assessments of social-cognitive understanding do, in fact, tap important formative social understandings in infants.

Fortunately, our study does not stand alone; a small set of other research has tackled similar issues (generally pathway 1 in **Figure 1**). Brune and Woodward (2007) reported that 10-month-olds' looking-time responses to a display linking a person's gaze to an intended object was related to their engagement in joint attention during a mother–infant play session. Hofer et al. (2008) reported that 6-month-olds whose mothers were rated to have a modestly controlling interaction style encoded actions in terms of the agent's goals, while those whose mothers were rated as sensitive or unresponsive did not. In an integration of looking-time and attachment research, Johnson et al. (2007) showed that securely attached infants (but not insecurely attached ones) looked significantly longer at an animated display in which a large "mother" object appeared to intentionally abandon a smaller "baby" object.

More recently, Licata et al. (2014) reported a study that partly parallels ours, but with 7-month-olds. These authors had 37 infants participate in Woodward's intention-understanding looking-time task (e.g., Woodward, 1998) and then videotaped mother–infant interaction in a 10-min free play episode. They assessed mother–infant interaction by using Biringen's (2000, 2008) Emotional Availability Scale, which involves global ratings of six dimensions such as maternal-sensitivity, maternal non-intrusiveness, and child-responsiveness. These dimensions were highly intercorrelated, so in a regression analysis controlling for age and infant activity level they entered only a general emotion-availability (EA) aggregate. EA was a significant predictor of children's looking-time performances. Licata et al. (2014) concluded that EA captured general mother–infant interaction quality; thus combining their findings with ours, the quality of mother–infant interaction significantly relates to infant intention understanding as measured in looking-time tasks for 7-, 10-, 11-, and 12-month-olds. Interestingly, both we (Dunphy-Lelii et al., 2014) and Licata et al. (2014) also collected measures of maternal-mind-mindedness (Meins et al., 2003) from the free-play interactions, and in neither study did mind-mindedness predict infants' looking-time intention understanding.

Because this first study (as well as the ones by Brune and Woodward, 2007; Licata et al., 2014) was concurrent, we could reveal relations between the laboratory and semi-naturalistic social experiences, but could not determine the direction of these relations: Infant social-interactive experiences and proclivities could contribute to enhanced intention understanding and *vice versa*. Indeed, as we hypothesize in pathway 1 of **Figure 1**, most likely the relationship is bi-directional and transactional.

Study 2: Looking-Time Differences Predict 4-Year-Old Theory of Mind

In Study 2, we examined whether looking-time measures in infancy predict later social-cognitive understanding. Beginning with a study by Wellman et al. (2004), several studies have now examined this link—pathway 2 in **Figure 1**. But the current study, an updated version of Wellman et al. (2008), provides the most extensive and controlled evidence that we are aware of for a pathway from infants' laboratory assessed social cognition to later preschool theory of mind. For measures of infant social cognition, we examined looking-time behaviors on the habituation task used in Study 1. For preschool theory of mind, we measured false-belief understandings.

Overview of Study 2

Forty-five of the 10- to 12-month-old infants that participated in the looking-time habituation task in Study 1 returned to the laboratory at approximately 4 years of age to participate in a series of cognitive assessments. Focally these children were assessed on their theory of mind.

We also assessed children's IQ, language competence (vocabulary), and executive functioning at 4-years because theory of mind has been linked to maturity of general information processing abilities. Moreover, for 40 years, it has been clear that infant attention to perceptual-object displays (such as familiar vs. novel objects and images) in looking-time studies predicts later IQ (Bornstein and Sigman, 1986; McCall and Carriger, 1993). These perceptual-attention findings are consistently interpreted as demonstrating developmental continuity in general information processing, such as memory-encoding or executive function. Conceivably, correlations between infant intention-understanding and preschool theory of mind might represent just another example of continuity in such general cognitive processing, as opposed to continuity more specific to a domain of social cognition. Thus, we included measures to account for this possibility.

Social-Cognitive Measures

Study 2 utilized the same infant looking-time measure described in Study 1—decrement of attention during habituation to intentional-action displays. Children's preschool theory of mind was assessed with several measures. Importantly, children completed two explicit false-belief tasks because these provide the most often used, standard assessment for preschool theory of mind (Wellman et al., 2001): a standard, contents false-belief task (from Wellman and Liu's (2004), theory-of-mind scale) and a

standard, change-of-locations task (a Sally–Anne task of the type first used by Baron-Cohen et al., 1985). The scores from these two tasks were summed to form a false-belief composite.

IQ and Executive Function Measures

For a brief IQ assessment we used two subscales of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1989) Vocabulary (a measure of verbal aptitude) and Block Design (a non-verbal measure tapping various capacities, including spatial understanding and logic).

Executive functions encompass several constructs (e.g., Zelazo et al., 1997), but, of these constructs, inhibitory control yields the strongest relations to theory of mind (Carlson and Moses, 2001). Inhibition also deserves attention because it has been posited to explain the continuity of IQ from infancy to childhood (McCall, 1994; McCall and Mash, 1995). We used two common assessments of preschool inhibitory control—Whisper and Bear/Dragon (Kochanska et al., 1996; Carlson and Moses, 2001). For example, in the whisper task, children see 10 cartoon characters, some well-known (e.g., Winnie the Pooh, Elmo) and some not (e.g., Marvin the Martian, Petunia the Pig) and are asked to whisper the names of the characters that they know. Children are inclined to blurt loudly the names of the characters they know, and thus whispers reflect more sophisticated inhibitory control and earn children higher scores on this task.

Findings

Focally, children's scores as infants in the looking-time task were correlated with their preschool theory of mind. Children's looking-time performance during infancy negatively correlated with the false belief composite when controlling for age, $r(42) = -0.38$, $p = 0.01$. That is, infants who sustained interest in (and therefore had smaller decrements in attention to) human intentional action displays during habituation at 1 year had enhanced theory of mind at age 4-years as indexed by better performance on the false belief tasks. Conceivably this association might be fully explained by verbal competence, performance competence (or a combination of the two representing overall IQ), and/or executive function. However, when we simultaneously controlled for age at infant testing, WPPSI Vocabulary and Block Design, and both executive function tasks, the relation between infant decrement of attention and preschool false-belief understanding remained significant, $r(35) = -0.38$, $p < 0.05$.

Discussion

This second study demonstrated that, not only are individual differences in infant looking-time performance to intentional action displays predicted by parent–infant social interactive experiences, those looking-time differences also predict later theory of mind at 4 years of age. Moreover, just as for Study 1, our data are not the only relevant findings.

Specifically, Wellman et al. (2004) initially found a correlation between infant social attention and 4-year-olds' performance on a false-belief composite for eighteen 14-month-olds. Replicating Wellman et al. (2004), Aschersleben et al. (2008) found a similar correlation between infant attention to goal-directed action and

4-year-old false-belief understanding for 20 German 6-months-olds. Yamaguchi et al. (2009) also reported a similar result: 4-month-old attention to goal-directed action (in a simpler procedure using animated circles and triangles) predicted later theory of mind (false-belief understanding) in a group of 17 U.S. children. Moreover, in a parallel study of fifteen 4-month-old infants, attention to *physical* stimuli (i.e., discrimination of tones rather than intentional-action stimuli) did not predict later theory of mind at 4 years (Yamaguchi et al., 2009). Notably, the 45 infants in our Study 2 more than doubled the samples used in these other studies and yielded a substantial correlation between later false-belief understanding and 10- to 12-month olds' social attention.

Across several of these studies, the relation between infant social attention and later theory of mind is consistent and robust in the sense of remaining essentially undiminished when measures of more general cognitive processes are controlled. Wellman et al. (2004) used a single measure of verbal IQ (Peabody Picture Vocabulary Test); Aschersleben et al. (2008) used a single composite measure of language competence (the SETK; Grimm, 2001). And, as just noted, Yamaguchi et al. (2009) showed that one measure of attention to physical displays failed to predict later theory of mind. Of course, any single control is limited. For this reason, we examined multiple measures of general cognitive processing including verbal and non-verbal measures of general IQ and measures of executive functioning. More generally, verbal competence, general IQ, and executive function are complementary aspects of general information processing that make substantial and independent contributions to preschool theory-of-mind performance. Our Study-2 findings demonstrate continuity from infant social attention to preschool theory of mind with all three factors measured and controlled.

This social-cognitive continuity is consistently evident for measures of attention during habituation, but within the same studies (Wellman et al., 2004, 2008; Aschersleben et al., 2008), it does not appear for measures of test event looking (but see, Yamaguchi et al., 2009). We return to this issue in the General Discussion. Regardless, social cognition evidences distinctive infant-to-preschool continuities indicating that theory of mind constitutes its own domain of cognitive development. Infant social-cognitive understanding is not only early achieved as revealed in sensitive looking-time paradigms, it is formative for further developmental advances in theory of mind as hypothesized in pathway 2 of **Figure 1**.

Study 3: Using Infant Individual Differences to Predict Preschool Social Cognition

While important, infant looking-time measures do not capture all of infant social cognitions or all of the predictive variance between infant social competence and experience and later social cognition. Indeed, Study 1 shows that the infant looking-time data and social-interactive measures interrelate. One possibility, therefore, is that looking-time measures are essentially a proxy for early infant–mother social experiences that themselves influence later social cognition including theory of mind. Or, more in line with the pathways outlined in **Figure 1**, it is possible that both social-cognitive competence (indexed in looking-time studies)

and social-interactive behavior and temperament might independently contribute to the further development of preschool social cognition. If that is the case, then we would want to know the extent to which both types of measures utilized in concert predict preschool social-cognitive outcomes. We tackle these issues in Study 3.

In this final study we take advantage of the fact that the children for whom we have longitudinal data in Study 2 were also in Study 1. Thus we can combine information about social differences in the parent–infant interaction measures focal to Study 1, plus looking-time differences focal to Studies 1 and 2, as well as theory-of-mind data from the same children at 4 years of age. This allowed us to address several crucial questions: What features of infant social-cognitive and social-interaction experiences combine to best predict later theory of mind? What factors are, separately, the most influential? And what is the total predictive power of these factors?

Overview of Study 3

Forty-three children participated in a series of social interaction and social-cognitive assessments: mother–infant interactions and looking-time habituation tasks at 10–12 months of age, as well as theory of mind tasks at 4 years. Social-interaction measures, looking-time measures, and preschool theory of mind measures were those described in Studies 1 and 2.

Findings

We used regression modeling to determine the independent and joint contributions of infant social-interactive measures (i.e., the four measures: quality of mother–infant interaction, socially observant infant temperament, infant joint attention, and imitation) and habituation scores to children's theory of mind at age 4. Given the number of measures, not all participants had complete data. Specifically, seven participants were missing data at random; therefore, five iterations of imputation were performed in order to predict the values of the missing data. The participants that were missing data did not differ systematically from the other children in terms of demographics or the other variables of interest (quality of mother–infant interaction, socially observant temperament, joint attention, or imitation). In all five iterations the exact same patterns of significance and non-significance emerged, therefore we report the results based on the pooled values of the five imputations.

Again, to test the possibility that something like general object-centered attentiveness or general "maturity" would account for our findings, we first entered the infant object-attentiveness measure as well as age into a regression predicting false-belief understanding at age 4; this was not significant, $R^2 = 0.07$, $F(2,38) = 0.27$, $p = 0.47$. Entering our four infant social-interactive measures combined with habituation scores in a second separate analysis did significantly predict children's theory of mind at age 4, $R^2 = 0.27$, $F(5,37) = 2.79$, $p < 0.05$. Of the five measures, the infant habituation measure ($t = 2.83$, $p < 0.01$) and socially observant temperament ($t = 2.34$, $p < 0.05$) independently significantly predicted performance on the false-belief tasks.

Similarly, using stepwise regression, the model including the four infant social-interactive aggregates and the habituation measure was reduced to two independent and significant predictors of preschool false belief understanding: socially observant infant temperament and the infant habituation measure. Socially observant infant temperament and infant performance on the looking-time task predicted 26% of the variability in preschool false belief understanding, $R^2 = 0.26$, $F(2,40) = 2.87$, $p < 0.05$.

Discussion

Study 3 demonstrates that both infant performance differences in looking-time paradigms and parent–infant interaction differences, and especially infant social temperament differences, independently predict later theory-of-mind performance. Neither is a mere proxy for the other, and together infant individual differences of both sorts more powerfully predict preschool social cognition. Indeed, when our several predictive factors were entered jointly, the overall regression model accounted for a sizable amount of variance in preschool theory of mind, and two variables alone—one looking-time measure and one social-interaction measure—accounted for 26% of the variability in 4-year-olds' false belief performance.

Pathways from infant social-interaction experiences to later social cognition (pathway 3 in **Figure 1**) are not well studied, but, still, our Study-3 findings are complemented by several others. Nelson et al. (2008) examined 30 min of mother–infant interaction for children aged 18–21 months. From that they extracted several measures of joint attention or joint engagement and related these to children's false-belief competence as 4-year-olds. Two measures were particularly predictive: *Coordinated joint engagement* (where infant as well as mother managed the dyad's joint attention to events) and *symbol-infused joint engagement* (where child participation in joint events included verbal reference to both the mother and to the events). Higher amounts of coordinated and symbol-infused joint engagement in these toddlers significantly predicted better false-belief performance when the children were 4 years of age, and did so even after language competence was controlled. Likewise, in an early study of a small sample of 13 infants, joint attention at 20 months (measured as gaze switching between an adult and a salient toy in a laboratory task) was predictive of later preschool-age theory of mind, and remained significant after IQ and language competence were partialed out (Charman et al., 2000).

Other measures of infant social-cognitive understanding could also be related to later preschool theory-of-mind performance. Recently, investigators of infant social cognition have expanded beyond examination of infants' understanding of intentional action and emotion to focus on infants' implicit understanding of agents' knowledge and beliefs as well (e.g., Onishi and Baillargeon, 2005; Southgate et al., 2007; Buttelmann et al., 2009). Implicit understanding of false belief is assessed via violation of expectation looking-time studies as well as anticipatory looking measured via eye-tracking methods. Thoermer et al. (2012) have shown that belief-based anticipatory looking measures at 18 months of age longitudinally predict false-belief reasoning on standard verbal preschool tasks at 48 months.

Note that our study goes beyond such prior research in including younger infants and in including both social-interactive *and* looking-time measures. To reiterate, together infant individual

differences of both sorts more powerfully predicted preschool social cognition.

General Discussion

Returning to **Figure 1**, contemporary research on early childhood understanding of the social world should ideally encompass numerous constructs studied in the lab and in social interactions and, crucially, the transactional and longitudinal pathways that form and contextualize such understandings. While the field has, for the most part, explored bits and pieces of this developmental system separately, the appropriate combination of constructs, approaches, and ages can provide a clearer picture of the developmental trajectory of children's social-cognitive understanding.

In this paper, we reviewed evidence from three studies conducted by our research group that examined pathways 1, 2, and 3 in our theoretical framework. We demonstrated that infants' social-cognitive understanding, as measured by laboratory performance on social-cognitive tasks, was significantly associated with individual differences in social-interactive experiences and infant social temperaments (pathway 1). We also demonstrated that both infant social-cognitive understanding and infant social-interactive experiences and temperaments contributed independently to later social-cognitive competencies, specifically preschool theory of mind (pathways 2 and 3). And they do so even after age, IQ, language competence, and executive function are controlled.

It seems of special note that among our predictors, socially observant infant temperament was an important predictor of infant social cognition in looking-time tests, as well as a particularly important predictor of later, preschool social cognition. This finding takes its place beside an emerging set of findings concerning relations between preschool temperament and theory of mind achievements—preschoolers who are shy, socially observant, and non-reactive outperform their peers on false-belief tasks (Wellman et al., 2011; Lane et al., 2013). What we add to these findings are data from infants and that infants' socially-observant temperament relates to their social cognition both concurrently (at 1 year) but also longitudinally, during the preschool years. Importantly, Mink et al. (2014) have also recently reported that infants with a shy temperament at 18-months exhibit a more advanced theory of mind by the early preschool years.

Note that this research showing early links between socially observant temperament and theory of mind spans several countries and cultures: the U.S. (the current studies; Wellman et al., 2011), China (Lane et al., 2013), and Germany (Mink et al., 2014). This is an impressive beginning, but of course, further research with other samples and especially with infants would be welcome. More generally, it will be important to determine the extent to which these relations between infant understanding of intentional agents, their social-interactive experiences, and early social-cognitive developments exist for other realms of social-cognition, not just theory of mind, but, for example, in infants' and preschoolers' moral intuitions.

We hasten to emphasize that this social attentiveness—reflected in both infant socially observant temperament as well as

maintenance of interest in intentional action displays in our looking-time paradigms—should not be thought of as a purely dispositional factor inherent to the infant. In Study 1 social attentiveness, as measured by infant "temperament" ratings, was related to measures of mother–infant interaction and interaction quality. In Licata et al. (2014) attention to intentional action in their looking-time task was related to maternal emotional availability. In short, enhanced attention to social actors and social interactions in infancy powerfully impacts early childhood social cognition, but that enhanced social attention can be due to the efforts of the infant, the infant's social partners, or mostly likely to both in concert. Determining the nature and social shaping of social attentiveness in early life is probably the single topic now most worth increased research efforts.

There are other pathways apparent in **Figure 1** that we have not touched on here. For example, it is now well established that preschool theory of mind impacts preschoolers' social behavior, such as the skilled interactions with and hence popularity with their peers (e.g., Watson et al., 1999; Diesendruck and Ben-Eliyahu, 2006), their ability to tell lies (e.g., Lee, 2013), their attempts to persuade others (e.g., Bartsch and London, 2000) and their engagement in games like hide-and seek (Peskin and Ardino, 2003). And preschool social behaviors, such as engaging in pretend play (e.g., Jenkins and Astington, 2000), engaging in family conversation about mind and emotion (e.g., Ruffman et al., 2002), and living in a family full of siblings and extended family members (e.g., Peterson, 2000), enhance preschool theory of mind. However, in keeping with a focus that includes infants, the remaining pathway of most import is the diagonal from infant social cognition to preschool social behavior. We know of no research that has tackled this pathway; it now seems of special import for future research.

More specific aspects of our findings also deserve mention. First, *both* more and less sustained attention to intentional action displays relate to infants' interactions and later competences. Both of these relations make sense: Steeper decrement of attention in habituation can reflect coming more quickly to a consolidated and intentional understanding of action. It can index infants' ability to parse intentional action displays meaningfully and thus habituate to them. Hypothetically, infants who more quickly habituate to complex displays of the sort that we present could be more practiced at understanding intentional action regularities and thus more quickly become habituated to them. That sort of relation may very well underpin the positive relation between decrement of attention and parent–infant interaction measures we found in Study 1.

At the same time, more sustained attention–and hence coming more slowly and less "steeply" to habituation criteria—could reflect greater interest in and deeper processing of intentional action. And such sustained attention to human intentional actions could promote later more advanced understanding. That sort of relation may very well underpin the negative relation we found between decrement of attention and later theory of mind in Study 2. Similarly, the relation between our infant temperament ratings and later theory of mind that we found in Study 3 seems to reflect the power of sustained, enhanced, infant social attentiveness.

Notably, this mix of significant relations, including both more quickly parsing intentional action and also greater attention to intentional action, occurs not only in our three studies but also throughout the literature. In Wellman et al. (2004) the relation between infant decrement of attention and later false-belief was positive, suggesting that by 14 months more quickly parsing intentional action predicts theory of mind at 4 years. Yet, to repeat, in Study 2 here (and thus in Wellman et al., 2008) the relation was negative. So for 10- and 12-month-olds, more sustained attention to intentional action displays predicted theory of mind at 4 years. Perhaps, then, for younger infants the key factor is sustained attention in habituation, but for older infants it is quicker, fluent intention-action processing. That cannot be the whole story, however, because in Aschersleben et al.'s (2008) study using a very different intention-action habituation display with 6-month-olds, it was infants who more quickly habituated—thus showing a larger attention decrement in habituation—that proved better at false-belief understanding as 4-year-olds. This abundance of differing relations requires closer scrutiny in future research.

Second, in our research and others (e.g., Aschersleben et al., 2008), it was infants' attention to intentional-action displays during habituation trials and not test trials that was especially revealing. Thus, social-cognitive processing revealed during habituation has proved particularly important. Yet, Brune and Woodward (2007) and Licata et al. (2014) found that, given their looking-time task, infant differences in response to novelty were also significantly related to parent–infant social behavior. Notably, our data, in contrast to theirs, demonstrate important continuities between our infant measures and preschool social cognition. This is an important addition, but the full pattern of relations between social cognition and social interaction, and the full pattern of continuities from infant intentional understandings plus social interactions to later theory of mind is an important topic for further research.

To summarize, future research should continue to explore the theoretical framework that we present here. Our research, along with the work of others, provides a strong foundation supporting conclusions that infants' social-cognitive understanding, their social-interactive experiences, and their social temperaments significantly inter-correlate and contribute to later social-cognitive competencies. However, several remaining portions of the framework deserve further exploration. First, we proposed a number of pathways that have not yet been evaluated in the literature. How does infant social cognition relate to later preschool social interactions and behaviors? Additionally, future research should continue to assess the nature of the pathways between laboratory-based assessments of social cognition and later social understanding that we have only just begun to explore. More evidence is required to understand how individual differences in sustained attention and decrement of attention relate to later social developments. Additionally, the influence of sensitivity to novelty and sensitivity to familiar social stimuli should also be assessed.

To conclude, in keeping with the focus of this special issue, we demonstrate the importance of individual differences in research on early social cognition. Our specific focus was on individual differences anchored in infancy, particularly differences in everyday mother–infant interaction as well as laboratory-based looking-time tasks. Our results demonstrate that both approaches generate informative measures of individual differences, and that moreover used together such measures can be especially compelling. Future research on children's social development should not focus solely on laboratory-based habituation measures nor individual differences in social experiences, but the combination of the two. More generally, our studies, along with recent confirmatory findings from Licata et al. (2014), Mink et al. (2014), Brune and Woodward (2007) and the like, underscore that social-cognitive development is a constructive process based in social interaction and observation.

Acknowledgments

References

Akhtar, N., and Martinez-Sussmann, C. (2007). "Intentional communication," in *Socioemotional Development in the Toddler Years: Transitions and Transformations*, eds C. A. Brownell and C. B. Kopp (New York, NY: Guilford Press), 201–220.

Aschersleben, G., Hofer, T., and Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Dev. Sci.* 11, 862–868. doi: 10.1111/j.1467-7687.2008.00736.x

Baird, J. A., and Astington, J. W. (2005). "The development of the intention concept: from the observable world to the unobservable mind," in *The New Unconscious*, eds R. R. Hassin, J. S. Uleman, and J. A. Bargh (New York, NY: Oxford University Press), 256–276.

Bakeman, R., and Adamson, L. B. (1984). Coordinating attention to people and objects in mother–infant and peer-infant interaction. *Child Dev.* 55, 1278–1289. doi: 10.2307/1129997

Baron-Cohen, S., Leslie, A. M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition* 21, 37–46. doi: 10.1016/0010-0277(85)90022-8

Bartsch, K., and London, K. (2000). Children's use of mental state information in selecting persuasive arguments. *Dev. Psychol.* 36, 352–365. doi: 10.1037/0012-1649.36.3.352

Bauer, P. J., and Kleinknecht, E. E. (2002). To 'ape' or to emulate? Young children's use of both strategies in a single study. *Dev. Sci.* 5, 18–20. doi: 10.1111/1467-7687.00197

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., et al. (2010). The origins of 12-month attachment: a microanalysis of 4-month mother–infant interaction. *Attach. Hum. Dev.* 12, 3–141. doi: 10.1080/14616730903338985

Bigelow, A. E., MacLean, K., Proctor, J., Myatt, T., Gillis, R., and Power, M. (2010). Maternal sensitivity throughout infancy: continuity and relation to attachment security. *Infant Behav. Dev.* 33, 50–60. doi: 10.1016/j.infbeh.2009.10.009

Biringen, Z. (2000). Emotional availability: conceptualization and research findings. *Am. J. Orthopsychiatry* 70, 104–114. doi: 10.1037/h0087711

Biringen, Z. (2008). *The Emotional Availability (EA) Scales*, 4th Edn. Boulder: Colorado State University.

Bornstein, M. H., and Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Dev.* 57, 251–274. doi: 10.2307/1130581

Brandone, A. C., and Wellman, H. M. (2009). You can't always get what you want: infants understand failed goal-directed actions. *Psychol. Sci.* 20, 85–91. doi: 10.1111/j.1467-9280.2008.02246.x

Brune, C. W., and Woodward, A. L. (2007). Social cognition and social responsiveness in 10-month-old infants. *J. Cogn. Dev.* 8, 133–158. doi: 10.1080/15248370701202331

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342. doi: 10.1016/j.cognition.2009.05.006

Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333

Carpenter, M., Akhtar, N., and Tomasello, M. (1998). Fourteen- to 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behav. Dev.* 21, 315–330. doi: 10.1016/S0163-6383(98)90009-1

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., and Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cogn. Dev.* 15, 481–498. doi: 10.1016/S0885-2014(01)00037-5

Diesendruck, G., and Ben-Eliyahu, A. (2006). The relationships among social cognition, peer acceptance, and social behavior in Israeli kindergarteners. *Int. J. Behav. Dev.* 30, 137–147. doi: 10.1177/0165025406063628

Dunphy-Lelii, S., LaBounty, J., Lane, J. D., and Wellman, H. M. (2014). The social context of infant intention understanding. *J. Cogn. Dev.* 15, 60–77. doi: 10.1080/15248372.2012.710863

Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193. doi: 10.1016/0010-0277(95)00661-h

Gräfenhain, M., Behne, T., Carpenter, M., and Tomasello, M. (2009). One-year-olds' understanding of nonverbal gestures directed to a third person. *Cogn. Dev.* 24, 23–33. doi: 10.1016/j.cogdev.2008.10.001

Grimm, H. (2001). *Language Development Test for 3- to 5-Year-Old Children [SETK 3–5]*. Gottingen: Hogrefe.

Hofer, T., Hohenberger, A., Hauf, P., and Aschersleben, G. (2008). The link between maternal interaction style and infant action understanding. *Infant Behav. Dev.* 31, 115–126. doi: 10.1016/j.infbeh.2007.07.003

Jenkins, J. M., and Astington, J. W. (2000). Theory of mind and social behavior: causal models tested in a longitudinal study. *Merrill Palmer Q.* 46, 203–220.

Johnson, S. C., Dweck, C. S., and Chen, F. S. (2007). Evidence for infants' internal working models of attachment. *Psychol. Sci.* 8, 501–502. doi: 10.1111/j.1467-9280.2007.01929.x

Kaitz, M., Maytal, H. R., Devor, N., Bergman, L., and Mankuta, D. (2010). Maternal anxiety, mother–infant interactions, and infants' response to challenge. *Infant Behav. Dev.* 33, 136–148. doi: 10.1016/j.infbeh.2009.12.003

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., and Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Dev.* 67, 490–507. doi: 10.2307/1131828

Lane, J. D., Wellman, H. M., Olson, S. L., Miller, A. L., Wang, L., and Tardif, T. (2013). Relations between temperament and theory of mind development in the United States and China: biological and behavioral correlates of preschoolers' false-belief understanding. *Dev. Psychol.* 49, 825–836. doi: 10.1037/a0028825

Lee, K. (2013). Little liars: development of verbal deception in children. *Child Dev. Perspect.* 7, 91–96. doi: 10.1111/cdep.12023

Licata, M., Paulus, M., Thoermer, C., Kristen, S., Woodward, A. L., and Sodian, B. (2014). Mother–infant interaction quality and infants' ability to encode actions as goal-directed. *Soc. Dev.* 23, 340–356. doi: 10.1111/sode.12057

Masur, E. F., and Olson, J. (2008). Mothers' and infants' responses to their partners' spontaneous action and vocal/verbal imitation. *Infant Behav. Dev.* 31, 704–715. doi: 10.1016/j.infbeh.2008.04.005

McCall, R. B. (1994). What process mediates predictions of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence* 18, 107–125. doi: 10.1016/0160-2896(94)90022-1

McCall, R. B., and Carriger, M. S. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Dev.* 64, 57–79. doi: 10.2307/1131437

McCall, R. B., and Mash, C. W. (1995). "Infant cognition and its relation to mature intelligence," in *Annals of Child Development: A Research Annual*, Vol. 10, ed. R. Vasta (London: Jessica Kingsley Publishers), 27–56.

Meins, E., Fernyhough, C., Wainwright, R., Clark-Carter, D., Gupta, M. D., Fradley, E., et al. (2003). Pathways to understanding mind: construct validity and predictive validity of maternal mind-mindedness. *Child Dev.* 74, 1194–1211. doi: 10.1111/1467-8624.00601

Mink, D., Henning, A., and Aschersleben, G. (2014). Infant shy temperament predicts preschoolers Theory of Mind. *Infant Behav. Dev.* 37, 66–75. doi: 10.1016/j.infbeh.2013.12.001

Nelson, P. B., Adamson, L. B., and Bakeman, R. (2008). Toddlers' joint engagement experience facilitates preschoolers' acquisition of theory of mind. *Develop. Sci.* 11, 847–852.

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255–258. doi: 10.1126/science.1107621

Peskin, J., and Ardino, V. (2003). Representing the mental world in children's social behavior: playing hide-and-seek and keeping a secret. *Soc. Dev.* 12, 496–512. doi: 10.1111/1467-9507.00245

Peterson, C. C. (2000). Kindred spirits: influences of siblings' perspectives on theory of mind. *Cogn. Dev.* 15, 435–455. doi: 10.1016/S0885-2014(01)00040-5

Phillips, A. T., Wellman, H. M., and Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition* 85, 53–78. doi: 10.1016/S0010-0277(02)00073-2

Putnam, S. P., Rothbart, M. K., and Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant Child Dev.* 17, 387–405. doi: 10.1002/icd.582

Rosenblum, K. L., McDonough, S. C., Sameroff, A. J., and Muzik, M. (2008). Reflection in thought and action: maternal parenting reflectivity predicts mind-minded comments and interactive behavior. *Infant Ment. Health J.* 29, 362–376. doi: 10.1002/imhj.20184

Ruffman, T., Slade, L., and Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Dev.* 73, 734–751. doi: 10.1111/1467-8624.00435

Saylor, M. M., Baldwin, D. A., Baird, J. A., and LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *J. Cogn. Dev.* 8, 113–128. doi: 10.1080/15248370709336996

Slaughter, V., Peterson, C. C., and Carpenter, M. (2008). Maternal talk about mental states and the emergence of joint visual attention. *Infancy* 13, 640–659. doi: 10.1080/15250000802458807

Sodian, B., and Thoermer, C. (2004). Infants' understanding of looking, pointing and reaching as cues to goal-directed action. *J. Cogn. Dev.* 53, 289–316. doi: 10.1207/s15327647jcd0503_1

Sommerville, J. A., Woodward, A. L., and Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96, B1–B11. doi: 10.1016/j.cognition.2004.07.004

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592. doi: 10.1111/j.1467-9280.2007.01944.x

Sroufe, L. A., and Waters, E. (1977). Attachment as an organizational construct. *Child Dev.* 48, 1184–1199. doi: 10.2307/1128475

Thoermer, C., Sodian, B., Vuori, M., Perst, H., and Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *Br. J. Dev. Psychol.* 30, 172–187. doi: 10.1111/j.2044-835X.2011.02067.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi: 10.1017/S0140525X05000129

Watson, A. C., Nixon, C. L., Wilson, A., and Capage, L. (1999). Social interaction skills and theory of mind in young children. *Develop. Psychol.* 35, 386–391. doi: 10.1037/0012-1649.35.2.386

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wellman, H. M., Lane, J. D., LaBounty, J., and Olson, S. L. (2011). Observant, nonaggressive temperament predicts theory of mind development. *Dev. Sci.* 14, 319–326. doi: 10.1111/j.1467-7687.2010.00977.x

Wellman, H. M., and Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Dev.* 75, 523–541. doi: 10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., Lopez-Duran, S., LaBounty, J., and Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Dev. Psychol.* 44, 618–623. doi: 10.1037/0012-1649.44.2.618

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., and LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Dev. Sci.* 7, 283–288. doi: 10.1111/j.1467-7687.2004.00347.x

Wechsler, D. (1989). *Primary Scale of Intelligence-Revised (WPPSI-R)*. San Antonio, TX: The Psychological Corporation.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34. doi: 10.1016/S0010-0277(98)00058-4

Yamaguchi, M., Kuhlmeier, V. A., Wynn, K., and VanMarle, K. (2009). Continuity in social cognition from infancy to childhood. *Dev. Sci.* 12, 746–752. doi: 10.1111/j.1467-7687.2008.00813.x

Zelazo, P. D., Carter, A., Reznick, J. S., and Frye, D. (1997). Early development of executive function: a problem-solving framework. *Rev. Gen. Psychol.* 1, 198–226. doi: 10.1037/1089-2680.1.2.198

A longitudinal study of the emerging self from 9 months to the age of 4 years

Susanne Kristen-Antonow*, Beate Sodian, Hannah Perst and Maria Licata

Department of Psychology, Ludwig-Maximilians-University, Munich, Germany

The aim of this study was to investigate if children's early responsiveness toward social partners is developmentally related to their growing concept of self, as reflected in their mirror self-recognition (MSR) and delayed self-recognition (DSR). Thus, a longitudinal study assessed infants' responsiveness (e.g., smiling, gaze) toward social partners during the still-face (SF) task and a social imitation game and related it to their emerging MSR and DSR. Thereby, children were tested at regular time points from 9 months to 4 years of age. Results revealed significant predictive relations between children's responsiveness toward a social partner in the SF task at 9 months and their MSR at 24 months. Further, interindividual differences in children's awareness of and responsiveness toward being imitated in a social imitation game at 12 months proved to be the strongest predictor of children's DSR at 4 years, while some additional variance was explained by MSR at 24 months and verbal intelligence. Overall, findings suggest a developmental link between children's early awareness of and responsiveness toward the social world and their later ability to form a concept of self.

Keywords: longitudinal studies, self concept, social cognition, conceptual development, infancy research

Introduction

Self Development: The Importance of Longitudinal Data

The ability to represent oneself as an intentional agent is foundational for the development of social cognition. Meltzoff (1990) has argued that, from birth, infants are able to recognize others as "like me." Based on this fundamental human ability to establish correspondence between oneself and another agent, infants' increasing ability to represent themselves as intentional agents leads to an understanding of others' intentional action. Evidence for this view comes from a study of goal-encoding in very young infants which showed that infants as young as 3 months can encode others' reaching and grasping actions as goal-directed after having experienced themselves as goal-directed agents with the help of Velcro mittens (Sommerville et al., 2005). While a representation of the self as an intentional agent remains implicit in social interactions throughout the first and second years of life, first evidence for an explicit self-representation emerges close to the second birthday, when children recognize themselves in the mirror and begin to use self-referential language. Theories of the developing self (e.g., Damon and Hart, 1982; Meltzoff, 1990; Rochat, 2003) have emphasized the importance of experience in reciprocal social interaction during the first and second years of life in leading up to the developmental milestone of mirror self-recognition (MSR). Furthermore, MSR has been theoretically linked to later Theory of Mind development (Gallup and Suarez, 1986; Parker et al., 1994). However, to date, there is little evidence for these views, for lack of longitudinal data. If social responsiveness in specific types of early social interactions

which provoke self-awareness is developmentally linked to later explicit self-representation, then individual differences in these tasks should be correlated independently of more general cognitive abilities. There is longitudinal evidence for long-term conceptual continuity in understanding *others*' intentional states from infancy to preschool age (cf. Aschersleben et al., 2008; Wellman et al., 2008; Thoermer et al., 2012), but no comparable studies have been conducted with respect to the self. In the present longitudinal study of self-development, we investigate the developmental relation of two markers of self-representation, indexing different levels of self-awareness: The MSR task at 24 months and the delayed self-recognition (DSR) task at 4 years of age. Note that these classic, yet controversial, tests have been widely used as markers of the self in empirical studies across cultures (e.g., Keller et al., 2005; Broesch et al., 2010), in typical and atypical development (e.g., Povinelli et al., 1996; Lind and Bowler, 2009), as well as across species (e.g., Povinelli et al., 1993). Specifically, we explore the predictive relations between precursor abilities emerging in social interaction in the first and second years of life and these two markers of self-representation.

Levels of Self-development

The developmental process of self-understanding has been described as "forward engineering" (Rochat, 2003, p. 117, p. 10), meaning that different constituents of the self develop chronologically during infancy and early childhood (Damon and Hart, 1982). This view implies that interindividual differences in competencies at the different theorized levels are related and that the self develops as a differentiated, yet conceptually coherent concept. Rochat described five levels of self-awareness of which two levels are of central interest, since they pertain to MSR and DSR, respectively. The first level of interest is the "identification"-level at which toddlers are able to express an identified self and comprehend that the mirror reflects their self-experienced "me," not some other individual. In other words, at this level the toddler is able to detect the correspondence between a mental representation of the self and an observed marked mirror image and at the same time is able to differentiate between the two. The second level of interest is the level at which the self can be truly represented independently of featural information and temporal contingency. At this level pre-schoolers begin to recognize themselves in videos and photographs taken in the past as opposed to live videos or contingent mirror images. It is called the "permanence"-level. It has been controversially discussed to what extent toddlers' MSR and their ability to recognize themselves in videos index self-awareness or a permanent awareness of self, respectively. The following paragraph will discuss different theoretical viewpoints.

Interpretations of MSR and DSR

Classic interpretations see MSR as an evidence of children's knowledge about what they look like (Amsterdam, 1972; Lewis and Brooks-Gunn, 1979; Bischof-Köhler, 1988, 1991), since the child is required to use a mirror to detect a mark covertly placed on her or his nose or cheek. In order to do so, the child has to detect the discrepancy between the mental representation of his or her own body (e.g., cheek without rouge) and the observed

marked mirror image (e.g., cheek with rouge). Therefore, fitting with Rochat's (2003) label, children's mastery of this task at around 18–24 months of age (Amsterdam, 1972; Lewis and Brooks-Gunn, 1979; Asendorpf and Baudonniére, 1993; Nielsen and Dissanayake, 2004) has been regarded as evidence for being able to identify oneself.

This view has been challenged, however. Overall, there are lean interpretations, rich interpretations and proposals somewhere between lean and rich. While some theorists (Lewis and Brooks-Gunn, 1979; Courage and Howe, 2002; Nielsen et al., 2003) state that mastery of this task proves that children know what they look like, leaner interpretations have stated that children pass this test because of kinaesthetic-visual matching skills (Mitchell, 1993; Heyes, 1994). In contrast, richer theories have claimed that beyond identifying themselves, children's mark-directed behavior is evidential of their underlying introspective abilities and reflective capacities (Gallup, 1998; Gallup et al., 2002). For instance, Bischof-Köhler (1991) argued that the ability for mental imagination is necessary for MSR. Mental imagination involves self-objectification and describes the ability to represent objects, including the self, independently of the immediate perceptual reality (see also Moore et al., 2007). In the case of MSR, the child must be able to couple the "I" (the subject of one's own experience) with the objectified and reflected-on "Me." Thereby, the "I" can recognize the mirror image as "Me" (Bischof-Köhler, 1991, p.12). Lewis (2003) considers MSR as an indicator of self-metarepresentation abilities: by recognizing oneself in the mirror, the mental state of "Me" (as opposed to an implicit knowledge of the self) is established. The mental state of "Me" in turn gives rise to mental state attributions to others and awareness of the relation between self and other (Lewis, 2003). In contrast, in a more cautious proposal, Perner (1991) argued that self-recognizers have formed multiple representations of one situation or event, so called "secondary representations" and thus, understand the relation between the real situation and the situation in the mirror. However, MSR does not require a representation of the representational relation between oneself and the mirror image of oneself. Similarly, Suddendorf and Butler (2013) argue that MSR requires the ability to collate representations, rather than a metarepresentational understanding of the relation between these representations. Some proposals have been more domain-specific and state that children's developing cognitive skills in regard to analyzing their own face result in mastery of the MSR task (e.g., Neisser, 1995). It is argued that it is only when children understand that their face is important to other people in order to identify them, they start to use mirrors to see their reflection.

Empirical evidence seems to rule out extremely lean interpretations of MSR, as well as proposals focusing exclusively on children's developing recognition of their own face. For instance, Nielsen et al. (2006) showed that children can recognize features of their whole body such as their legs, instead of just their face, in the mirror. Further, when altering children's appearance by putting them in pants, 24-month-olds updated their representation of what they looked like during an exposure phase that had just lasted for 30 s. Thus, instead of facing problems because their proprioceptive matching capacities were handicapped by wearing pants, they detected the mark more

rapidly. They only faced problems when they had not been previously exposed to their altered looks and thus, could not build an expectation of what they were supposed to look like. This result supports that children recognize the mark because of a match between the image in the mirror and their expected image of themselves (see also Moore, 2007).

While MSR measures a temporally restricted self-representation, a representation of oneself as a temporally extended individual seems to develop only around the age of 4 years (e.g., Povinelli et al., 1996). This is what Rochat (2003) calls the "permanent self" and what according to Povinelli (1995) is the "proper self," characterized by children's comprehension of the fact that different temporal representations (past, present, future) of the self belong to one underlying unifying entity. The standard task to measure this more elaborate concept of self is the DSR task (Povinelli et al., 1996). In the DSR task children have to relate their current self to their temporally-delayed self as shown on a videotape. Mastery of this task emerges between 3 and 5 years of age (depending on the time-delay between recording and showing the video to the child; cf. Povinelli, 2001). While replicating the developmental asynchrony between MSR and DSR, a study by Suddendorf (1999) has challenged the view that young children have specific problems with self-awareness. Rather, they seem to have general problems in relating information shown in a video to the current situation, whether self-related or not (as in the case of a surprising object in the room). Similarly, in case of the MSR-task, live video versions lead to a significant drop in children's performance below chance level (Vyt, 2001; Suddendorf et al., 2007). While it remains debatable if these difficulties merely reflect children's problems with the video as a medium itself (Suddendorf, 1999, 2003; Troseth, 2003), the validity of the DSR task as a measure of self-representation has been questioned. Within-subject longitudinal data on different measures of self-awareness can contribute to a better understanding of this marker.

The Social Origins of Self

The general idea that social experiences in the first year of life support self-development (e.g., Damon and Hart, 1982; Rochat, 2009) is corroborated by empirical evidence from cross-cultural studies showing that variations in social experiences impact children's performance at the MSR-task. In cultures with a distal parenting style with a lot of face-to-face contact and object manipulation (e.g., Greece) children recognize themselves earlier than in cultures with a stronger emphasis on body contact and body stimulation (e.g., Cameroon). Cultures (e.g., Costa Rica) utilizing proximal as well as distal parenting practices fall between the two other cultural groups in regard to the onset of children's MSR (Keller et al., 2004). Further, a parenting style where mothers react more contingently (e.g., German mothers when compared to Nso mothers in Cameroon) has been shown to lead to a higher rate of MSR (Keller et al., 2005).

Thus, contingency detection in reciprocal social interaction appears to be one mechanism supporting the development of self-awareness. Rochat (2009) emphasizes that infants' engagement in reciprocal social interaction allows them to use the adult as a screen providing an opportunity for self-objectification leading to an emerging sense of shared experience with others. Within these reciprocal exchanges infants can develop a sense of contingency and agency as they experience causal efficacy between their own and the other person's behavior, as well as bodily reactions. A frequently investigated phenomenon assessing infants' sensitivity to what is reflected back by their interaction partner is the so-called still-face (SF) effect (Tronick et al., 1978). In reacting with irritation when their interaction partner interrupts the interaction (frozen face), and by showing different gaze and smile patterns while trying to bring one's social partner back into interaction, as well as by displaying reengagement behaviors such as vocalizing or bodily movements, children show social responsiveness (see Mesman et al., 2009, for a meta-analysis of 39 studies employing the SF task). Individual differences in social responsiveness to contingency disruption may thus be predictive of the age of mastery of MSR and DSR.

Another mechanism promoting self-awareness may be synchronic imitation. Based on infants' early imitation skills as a "foundation and earliest manifestation" (Meltzoff, 1990, p.141) of the self, infants have ample opportunity to detect the structural equivalence between the acts they perform themselves and the acts they see others perform in everyday interaction. A study by Asendorpf and Baudonniére (1993) found that the extent to which 19-month-old infants engaged in synchronic imitation as a measure of other-awareness was affected by their MSR status. Consistently, strong relations were found between 18-month-olds' MSR skills and their concurrent imitation skills (Zmyj et al., 2013).This has been interpreted as evidence for a developmental synchrony between self- and other-awareness during imitation. However, a study by Nielsen and Dissanayake (2004) found that while both abilities emerge around the same age, synchronic imitation skills (Asendorpf et al., 1996) and MSR proved to be unrelated. The inconsistent findings may be due to the fact that imitation tasks pose many different cognitive and motivational demands beyond a basic self and other-awareness. A clearer measure of self-awareness can be attained in tasks which tap infants' awareness of their own actions being mirrored by another person. Meltzoff (1990) designed a social mirroring task, in which 9-to 18-month olds' had to discriminate between an experimenter mimicking their actions on an identical object and the object "mimicking" them without being manipulated by an experimenter, and found clear evidence for an awareness of being imitated in the majority of the infants above the age of 14 months. Agnetta and Rochat (2004) did not find significant predictive correlations between 14-month-olds' awareness of being mirrored and their MSR at 18 months. However, the relation was only explored at one measurement point, and only with respect to MSR. If the awareness of being imitated is closely linked to an understanding of others as intentional agents in infants above the age of 12 months as Agnetta and Rochat suggest, then individual differences in this ability may very well be predictive of later DSR which has been theoretically linked to metarepresentation and mental state understanding.

Hypotheses

Summing up, based on theories distinguishing between different levels of self-development in the sense of one underlying

unitary concept (e.g., Rochat, 2003) we assume MSR-skills to be related to DSR-skills. Further, based on Rochat's (2009) social construction theory and Meltzoff's (1990) "like me" hypothesis, we assume interindividual differences in (a) social responsivity to contingency disruption, and (b) self-awareness in social mirroring to be predictively related to interindividual differences in a time-restricted concept of self as indexed by MSR-skills, as well as in a temporally-extended concept of self as indexed by DSR-skills. We expect these developmental pathways to be specific and thus, independent of general cognitive abilities.

To test these hypotheses, we studied the predictive relations between interindividual differences (1) in social responsiveness in dyadic interaction (SF reaction) at 9 months of age, (2) in social mirroring assessed at 12 months of age, (3) in MSR at 24 months of age and (4) in DSR at 50 months of age, in a longitudinal design. Interindividual differences based on gender and verbal IQ were systematically taken into account.

Materials and Measures

Participants
Overall, 89 full-term children (41 female, 48 male) participated in this comprehensive longitudinal study, while due to attendance the *n* at the later measurement points could vary. The mean age at the first measurement point relevant for this study was 9.00 months (SD = 9 days; *n* = 88), 12.01 months (SD = 7 days, *n* = 89) at the second measurement point and 24.03 months (SD = 8 days; *n* = 81) at the third measurement point. At the fourth measurement point, children were, on average, 4.21 years of age (SD = 0.93 months; *n* = 70). Children were tested in a child-friendly research laboratory at the University of Munich and all came from lower to upper middle-class families in an urban area in the South of Germany. On average, children had one sibling, while the number of siblings ranged from 0 to 3. The study followed the ethical standards for experiments involving humans and was approved of by the University of Munich's ethics committee.

Still-face Task
The paradigm was adapted from Striano and Stahl (2005) and involved two interruptive situations (adopting a neutral face and ignoring the child). The main purpose of this task was to measure children's social responsiveness when confronted with an interruption of communication. Infants were seated on a highchair facing the experimenter at a distance of about 45 cm. Once infants were seated, two identical plastic objects (10 cm of height) were unobtrusively placed to the infant's left and right side at a distance of about 70 cm. The procedure always started off and finished with a normal interaction (NI), while in between the NI and the two different SF phases were alternated in a randomized order. The five phases lasted for 30 s each. In the NI phase children were involved in a natural dyadic interaction. To render the communication situation as natural as possible, experimenters were free to react intuitively to children's social interaction bids by talking, singing or laughing. In the SF face-to-face condition, the experimenter adopted a neutral facial expression and looked at the infant's face without any affect. During the SF ignore phase,

the experimenter ensured that the infant held eye contact, then adopted a neutral facial expression and turned to the side of one of the two objects (the side of the objects was counterbalanced across children). Thus, the infant was ignored during the whole phase. No smiling or gazing back at the infant or touching the infant occurred. Thus, the two SF variants did not differ in the facial expression (neutral) and were both characterized by the absence of communicative bids.

Based on the coding scheme by Striano and Stahl (2005), infants' behaviors were coded using the INTERACT® software. Percent duration of time that infants engaged in a particular behavior was used as dependent measure. The dependent measures included the amount of smiling (raised cheeks, upward turned lips), gazes at experimenter and reengagement behaviors. Reengagement behaviors involved movements (arm or leg movements or pick-me-up gestures accompanied by looks directed at the experimenter) and communicating (e.g., babbling, squeaking, laughing or whining) while gazing at the experimenter. In order to analyze whether a SF effect was manifested, smiling, gaze and reengagement behaviors were averaged across the three NI episodes and then compared to the average duration of smiling, gaze and reengaging behavior across the two SF episodes.

Note that according to the literature competent children should, on average, show important differences in their gaze, smile and behavior when comparing the SF phase with the NI phase. More specifically, the SF effect involves a decrease in smiling and gazing behavior, as well as an increase of reengagement behaviors displayed toward the interaction partner during SF (interrupted) compared to NI episodes.

Thus, for subsequent correlational analyses, difference scores were computed for all three behaviors: the duration of each behavior during NI was subtracted by the duration during SF phases. Based on these difference scores, the following competence levels were defined: for smiling and gazing, a subject was classified as competent (and assigned a competence score of 1) if the difference scores in smiling and gazing behavior between the averaged NI and SF episodes was greater than zero, that is, the child spent less time smiling or gazing respectively, during SF phases than during NI. A child who received a negative value or a score of zero was classified as incompetent (and was assigned a competence score of 0). For reengagement behavior the rationale was different. Note that responsivity to social interaction cues is characterized by an increase of reengagement behaviors after the interruption of an ongoing interaction and a decrease of such behaviors once NI was re-established. Therefore, a competence score of 1 was assigned if the amount of reengagement behaviors during the SF phases was greater than during NI. Note that all scores around 0 and below were assigned a 0. In order to be assigned a score of 1 instead of 0, the differences between the NI and the SF effect had to be significantly different from zero.

Consequently, infants were classified as incompetent (and assigned a competence score of 0) if they did not show an increase of reengagement behaviors during SF episodes. These dichotomous variables were included in correlational analyses within the SF task as well as in correlations with measures from other tasks. A second independent coder, who was blind to the experimental hypotheses, coded a random 30% of all infants and

measures for reliability. Cohen's Kappa for all measures was 0.74. In regard to excluded children, the task was not administered to $n = 8$ children, while $n = 8$ children had to be excluded due to technical errors of the camera system and $n = 6$ children due to crankiness and being fuzzy. An additional five children could not be included in the smiling analyses because they did not display any smiling behavior.

Social Mirroring Task

This task was adapted from Meltzoff (1990). The main purpose of this task was to test infants' beginning self-awareness by assessing if they show preference for an experimenter who mirrors their own actions.

Two experimenters sat across the child at a table, while the child was seated within the caretaker's lap (120 cm away from both experimenters). To make sure children would not develop a preference for one of the experimenters, both experimenters were interacting with the child for about 5–10 min before the experiment started. At the beginning of the experiment, identical toys were handed to both experimenters, as well as to the child at the start of each trial. Each trial lasted for about 45 s. The toys were a car (10 cm long × 9 cm tall), a cup (6.5 cm diameter × 8.5 cm tall), a shovel (20 cm long × 6 cm wide) and a round form (12 cm diameter × 4 cm tall). While the starting object was counterbalanced across infants, in the following, the object order remained the same: the car was always preceded by the cup, the shovel was always preceded by the car, and the form would always follow the shovel independent of the respective starting object. Note that this was done in order to avoid order effects in this comprehensive longitudinal study. We predefined children's actions. The experimenter to the left always imitated the infant's actions. Correspondingly, the experimenter to the right performed control actions. The predefined actions (control actions in brackets) were as follows: Shake (Slide), Slide (Shake), Pound (Poke), Poke (Pound), Mouth (Touch Body), Touch Body (Mouth), Passive (Passive). It was made sure that experimenters showed the same activity level, which according to Meltzoff (1990) rules out that the infant prefers one of the experimenters because of the way he or she manipulated the toy. Further, since children might prefer the experimenter who acts temporally contingent upon their own actions, both the imitating and the non-imitating experimenter started and stopped acting at the same time as the infant. The task was filmed and there were three target behaviors which were then coded using the INTERACT® software. The average duration of smiling and looking at either the imitating or the non-imitating experimenter, as well as the number of instances children showed testing behavior averaged across the four objects were used as dependent variables. Thereby, in order to be coded as smiling infants had to display raised cheeks and lips which were turned upward. Testing behavior was defined as sudden and unexpected actions (sudden stop and restart) on the toy while eyeing the experimenter.

To build competence scores, first difference scores were calculated by calculating the duration or number of times (in the case of testing behavior) of a particular behavior which was directed at the imitating experimenter minus the duration or number of times the same behavior was directed at the non-imitating experimenter. First, preferences score were created: Any subject receiving a difference score larger than zero preferred the imitating experimenter and therefore obtained the preference value 1, while subjects who did not differentiate between both experimenters obtained value 0. Subjects preferring experimenter 2 displayed more behavior toward experimenter 2 resulting in a negative difference score. Those infants obtained the preference value −1. Additionally, competence scores were established by merging the preference value −1 and 0. Thus, if infants preferred the non-imitating experimenter 2 or showed no preference for either experimenter they were classified as incompetent and were assigned a score of 0. In contrast, if infants preferred the imitating experimenter they were classified as competent (score of 1). The sum score of the competence scores in regard to all three behaviors was used in analyses. Cohen's Kappa for the competence scores of all three behaviors ranged from 0.73 to 1.0.

In regard to excluded children, the task was not administered to $n = 6$ infants due to time restrictions, while $n = 6$ infants had to be excluded due to not showing sufficient interest in the objects and $n = 8$ children due to technical problems with the camera system.

Mirror Self-recognition Task

The task was adapted from Asendorpf and Baudonniére (1993). The main purpose of this task was to assess children's growing concept of self by testing if children recognize themselves in the mirror.

Prior to testing, child and experimenter engaged in a warm-up phase involving a mirror (dimensions: 110.5 × 104 cm), during which the experimenter made sure the child fixated the mirror for a minimum of three times and at least once for 2 s (baseline-phase). Then, the parent approached the child to apply a mark on either the child's nose or cheek. This was done in a way to ensure that the parent was not visible in the mirror for the child or that the child and parent were not standing in front of the mirror. The parent used a cloth with lipstick traces (invisible to the child) and wiped the child's nose thereby leaving a mark on the child's nose or cheek. For none of the children this served as a clue leading them to touch their face after the mark-application. This application-phase was followed by the test-phase during which the experimenter made sure to focus the child's attention back on the mirror (e.g., by moving a puppet between the child and the mirror) so that the child would look into the mirror at least three times and at least once for 2 s.

The task was filmed and the videos were analyzed using a coding scheme adopted from the Mirror Behavior Checklist by Amsterdam (1972) in both the baseline-phase and test-phase. A child was coded as 1 (recognizer) if he or she touched the mark [e.g., while verbally referring to either the mark, both the mark and the self or to self (child's name or I)]. If the mark-touching behavior was not present, the child received a score of 0 (non-recognizer). Interrater-reliability for the recognizer/non-recognizer coding was assessed by calculating Cohen's Kappa and was 0.92.

In regard to excluded children, the task was not administered in $n = 3$ children due to time restrictions, while $n = 5$ children had to be excluded because could not be brought to focus on the mirror; $n = 2$ did not concentrate during the task and were distracted and

$n = 2$ additional infants had to be excluded because of technical problems with the camera system.

Delayed Self-recognition Task

This task was adapted from Povinelli et al. (1996). The main purpose was to assess children's understanding of the self as possessing explicit temporal features. At the beginning of the experiment the child and experimenter 1 sat across from each other at a table surrounded by black walls to secure a high contrast-video. Experimenter 2 sat at the child's right to operate a hand camera which stood ca. Two meters away from the child. Further, a covered video monitor within the child's visual field, but not yet visible to the child, was part of the setting. The monitor had a width of 39.5 cm and a height of 35 cm. Two cameras were used to film the whole setting, as well as close-ups of the child during the experiment.

At the beginning of the marking-phase, experimenter 2 showed the camera to the child. The child was told that the child and experimenter 1 would play a game and that the camera would record everything so that they could look at the video later. Then, experimenter 1 and the child began playing a search game lasting for five trials, where the child had to search for a cracker which could be hidden under three different opaque containers. Trials 1 and 2 were used to habituate the child to experimenter 2 touching his or her forehead (this was done while praising the child for his or her success at the search task). During trial 3 a sticker (a yellow one was used for dark-haired children and a blue one was used for blond-haired children) was placed at the child's forehead. The post-it stickers measured 76 × 76 mm. Trials 4 and 5 served as control trials to ensure the child had not detected the sticker. In these trials the child was only praised, but not touched.

The test-phase followed 2 min after trial 5 and experimenter 2 informed the child that they would now look at the video. It was made sure that the overall setting remained the same in the test-phase as compared to the marking-phase and that the child's face was not reflected in the monitor. The child was shown the video beginning in trial 3 and 15 s after the start the first prompt was given: "Who is this?," while pointing at the child in the video. If children gave no answer the prompt was repeated. After an additional 15 s, experimenter 2 gave the second prompt and asked: "What is this?," while pointing at the sticker in the video. If the child did not answer or answered incorrectly the experimenter said: "This is a sticker!" followed by "Where is the sticker now?" After 15 additional seconds the third prompt followed and the experimenter said: "Can you find the sticker? Where it is now?" Thereby, the word "now" was emphasized.

It was coded whether children took off the sticker or touched it, as well as when this behavior occurred. Based on this, children could receive scores ranging from 0 to 4. They received a score of 0 if they did not touch the sticker at all. They received a score of 1 if they touched it only after the third cue, a score of 2 if they touched it after the second cue, a score of 3 if they touched it after the first cue and a score of 4 if they touched it even before any cues were given. Interrater-reliability for this coding was 0.79 (Cohen's Kappa).

In regard to excluded children, the task was not administered in $n = 3$ children due to time restrictions, while $n = 9$ children

TABLE 1 | Overview of measurement points when particular tasks were administered.

	Still-face	Social mirroring	Mirror self-recognition	Delayed self-recognition	Verbal IQ
9 months	X				
12 months		X			
24 months			X		
4 years				X	X

TABLE 2 | Descriptive statistics of the study variables.

	N	%/M	SD	Categories/Range
SF effect mile 9 months	61	0 = 12/1 = 88	0.32	0/1
SF effect gaze 9 months	66	0 = 30/1 = 70	0.43	0/1
SF effect reengagement 9 months	66	0 = 21/1 = 79	0.41	0/1
Social mirroring 12 months	69	1.39	1.20	0–3
Mirror self-recognition 24 months	69	0 = 32/1 = 68	0.47	0/1
Delayed self-recognition 4 years	55	1.80	1.31	0–4
Verbal IQ 4 years	63	107.44	13.72	67–137

mo, months; SF, still-face.

detected the sticker during application and in $n = 2$ children the sticker fell off during the play-back phase. Finally, $n = 1$ child did not pay attention to the video at the critical time points.

WPPSI Verbal IQ Subtest

WPPSI verbal IQ subtest (Petermann, 2009). In order to control for children's general cognitive skills, at the fourth measurement point, their verbal IQ was assessed and used as a control variable.

For this purpose two subtests of the German version of the WPPSI (Petermann, 2009) verbal IQ scale were administered. The procedure followed the test manual. While the subtest *Information* measures children's basic knowledge in regard to a variety of topics, the subtest *Similarities* involves that children display verbal reasoning skills and engage in concept formation. First, raw scores were calculated which were transformed into normalized scores for the respective age group. Since we used two out of three subtests we arrived at the estimated Verbal IQ scores by building a sum of the normalized values, subsequently dividing it by two and finally, multiplying it by three. These steps followed the standard procedure for calculating IQ estimates as proposed in the test manual (Petermann, 2009). In $n = 7$ children the test could not be administered due to concentration issues. See **Table 1** for an overview of the measures.

Results

An overview of descriptive statistics of the study variables is presented in **Table 2**. In preliminary analyses, in regard to group differences based on gender, girls, on average, were more advanced in their social mirroring skills at 12 months, $t(67) = 2.25$, $p = 0.03$ [girls ($n = 31$): $M = 1.74$, SD $= 1.09$; boys ($n = 38$): $M = 1.11$, SD $= 1.23$], and in their MSR at 24 months of age, $\chi^2(1) = 9.11$, $p = 0.00$; $n = 69$, [girls ($n = 34$): recognizers: 85%; non-recognizers: 15%; boys ($n = 35$): recognizers: 51%; non-recognizers: 49%]. Further, girls showed more of a SF smile effect,

TABLE 3 | Zero-order point-biserial correlations between the study variables.

	1	2	3	4	5	6	7
SF effect smile	1						
SF effect gaze	0.12[#](60)	1					
SF effect reengagement	−0.18[#](61)	−0.18[#](66)	1				
Social mirroring	0.29*(47)	−0.11(51)	0.08(51)	1			
Mirror self-recognition	0.15(47)	0.29*[#](50)	0.01(50)	0.09(51)	1		
Delayed self-recognition	0.01(39)	−0.03(43)	0.01(43)	0.36*[≠](39)	0.26[+](49)	1	
Verbal IQ	0.13(45)	−0.05(50)	0.33*(50)	−0.06[≠](51)	−0.03(54)	0.22[≠](48)	1

*$p < 0.05$; [+]$p < 0.10$ (two-sided significance level); [#]Phi-coefficients; [≠]Pearson correlation; SF, still face.

$\chi^2(1) = 4.22$, $p = 0.04$; $n = 61$, [girls ($n = 31$): SF smile effect: 97% yes, 3% no; boys ($n = 30$): 80% yes, 20% no].

In contrast, boys and girls did not differ significantly in their SF gaze effect [$\chi^2(1) = 0.00$, $p = 1.00$; $n = 66$], nor in their SF reengagement effect, [$\chi^2(1) = 1.45$, $p = 0.23$; $n = 66$]. Further, girls and boys did not differ significantly in regard to verbal IQ, $t(61) = 0.64$, $p = 0.52$, as well as in regard to DSR skills at 50 months of age, $t(53) = 1.50$, $p = 0.14$.

To assess relations among the study variables, correlational analyses were conducted with a two-tailed significance level. As can be seen in **Table 3** we found that interindividual differences in social mirroring skills at 12 months of age were predicted by interindividual differences in the SF smile effect at 9 months, while in turn, interindividual differences in social mirroring at 12 months predicted interindividual differences in DSR at 4 years of age. Further, interindividual differences in MSR at 24 months and in DSR at 4 years were related.

To assess the influence of possible mediators (gender and verbal IQ), partial correlations were conducted.

The significant relation between children's social mirroring skills at 12 months and DSR at 4 years of age $r_{\text{partial}} (36) = 0.32$, $p = 0.048$, remained significant when controlling for gender, while the relation between MSR at 24 months and DSR at 4 years of age remained marginally significant when controlling for gender; $r_{\text{partial}} (46) = 0.23$, $p = 0.05$, as did the relation between the SF smile effect and social mirroring at 12 months, $r_{\text{partial}} (44) = 0.22$, $p = 0.07$.

In order to assess the influence of early social responsiveness on children's MSR, while also considering possible mediators, in addition to correlational analyses, regression analyses were conducted.

First, a binary-logistic regression analysis (inclusion method) was performed with the complete set of early social responsiveness measures as theoretically important predictors. Further, based on Baron and Kenny (1986) mediation model, we included all possible mediator variables based on if they showed significant or marginally significant correlations with the outcome into the regression analysis. Thus, verbal IQ could be excluded as a mediator variable since it proved to be unrelated to MSR at 24 months (see **Table 3**), while gender had to be included. The overall model correctly predicted 76.5% of recognizers and non-recognizers. Only the SF gaze effect proved to be an independent predictor of MSR at 24 months (see **Table 4**).

In order to assess the respective importance of early social responsiveness for children's DSR, while also taking into

TABLE 4 | Binary-logistic regression predicting MSR at 24 months.

	B	Wald	Odd's ratio	p
SF effect smile	0.842	0.363	2.321	0.547
SF effect gaze	2.266*	4.439	9.640	0.035
SF effect reengagement	0.614	0.246	1.847	0.620
Social mirroring	0.193	0.185	1.213	0.667
Gender	1.090	1.249	2.976	0.264
Constant value	−2.502	1.309	0.082	0.253

*$p < 0.05$; SF, still-face; SF, still-face.

TABLE 5 | Regression analysis to predict delayed self-recognition at 4 years of age.

	ß	T	R^2	F-value
SF effect smile	−0.24	−1.43		
SF effect gaze	−0.06	−0.35		
SF effect reengagement	−0.19	−1.16		
Social mirroring	0.43*	2.69		
Mirror self-recognition	0.29[^]	1.84		
Verbal IQ	0.34*	2.12		
			0.303	2.32[^]

[^]$p < 0.10$; *$p < 0.05$; SF, still-face.

consideration possible mediators, a linear regression analysis (inclusion method) was conducted with the complete set of early social responsiveness measures as predictors. Based on Baron and Kenny (1986), verbal IQ was included as a possible mediator, while gender could be excluded. As is shown in **Table 5**, the overall regression model, $F(6, 32) = 2.32$, $p = 0.056$, was marginally significant and explained 30% of variance in children's DSR at 4 years of age. Looking at single predictors, interindividual differences in social mirroring skills at 12 months were found to significantly predict DSR at 4 years of age, beyond verbal IQ at 4 years as another significant predictor and MSR at 24 months as a marginally significant predictor.

Discussion

This longitudinal study had two major aims: It explored the conceptual coherence of the self as a construct, as well as the specificity of the social origins of the self. It included precursor abilities and measures of self-awareness from infancy through toddlerhood to preschool age, as well as control measures. Most importantly, the findings suggest quite specific developmental

pathways. Using (cf., Rochat's, 2003) terminology, one such pathway appears to lead from infants' gaze reaction to contingency disruption at 9 months to self-"identification" at 24 months, while the other leads from self-awareness in a social imitation game at 12 months to self-"permanency" at 4 years of age. These developmental pathways indicate a fairly long-term continuity of social cognition in regard to self-understanding that is consistent with the long-term continuity of social cognition in regard to other-understanding as indicated by developmental pathways from infants' understanding of goals and intentions to their later theory of mind (cf., Aschersleben et al., 2008; Wellman et al., 2008; Thoermer et al., 2012). Further, interindividual differences in regard to the identified self and in regard to the permanent self proved to be moderately related, independently of gender or verbal IQ. Thus, the self seems to develop as a multi-dimensional, and at the same time, (moderately) coherent concept.

The findings provided some support for theories emphasizing the role of social interaction in the development of a concept of self (e.g., Moore, 2007; Rochat, 2009). Specifically, interindividual differences in children's responses to contingency disruption at 9 months as indicated by their gaze predicted MSR, while sensitivity toward social mirroring predicted DSR. This is consistent with Rochat's (2009) idea that children do not only use reflecting surfaces of all kinds, but their social world as a mirror. Thus, infants expect others to project their inner self back at them, just like a real mirror would do. Note that this expectation might have been rapidly formed during infants' early dyadic social interactions with their caregivers and might not be entirely endogenous, since exogenous, parental factors have been shown to impact children's behavior during SF (e.g., Rosenblum et al., 2002).

Further the results are consistent with cross-cultural findings (e.f., Keller et al., 2005) showing that toddlers show earlier MSR-skills if they are socialized in cultures with a distal parenting style promoting attentiveness toward the human face. Note also that the correlational link between gaze behavior during the SF task and MSR found in this study supports theories proposing that the development of domain-specific cognitive structures supports the development of MSR-skills (e.g., Neisser, 1995). More specifically, it was children's sensitivity toward interruption as indicated by gaze which predicted mark removal, while the arguably more emotional-evaluative reactions during the SF task, such as smiling and trying to re-engage the partner, were not related to mark removal. It is possible that smiling and re-engagement would be related to more emotional-evaluative reactions to the mirror image (such as puzzlement or coy reactions). This is a task for future research, since those behaviors are expected to develop well before mark removal. For instance, Reddy (2003) proposes an affective-engagement account claiming that the experience of self as an object to others as reflected in affective responses or coy reactions (i.e., smile with gaze and/or head aversion), occurs at a very early age, around 2 months. Links between the SF situation and these measures in the mirror situation need to be explored in future research involving younger age groups.

Again in favor of specific rather than more general developmental links and in favor of the self being a multi-dimensional construct, our data showed that while predicting

DSR, social mirroring in infancy did not predict MSR. Consistent with our findings, a study in 9-to-18-month olds by Agnetta and Rochat (2004), using a modification of Meltzoff's social mirroring task, also did not find significant predictive relations between social mirroring and MSR. Thus, while these skills may emerge around the same age (Nielsen and Dissanayake, 2004) and while interindividual differences may be concurrently related, especially later in development (Asendorpf and Baudonniére, 1993), social mirroring skills toward the end of the first year of life do not seem to be predictive of MSR. However, they seem to predict DSR. There are several possible explanations for these differential relations.

Nielsen and Dissanayake (2004) explained the missing link between MSR and synchronic imitation skills found in their own research referring to Perner's (1991) developmental theory. They argue that while secondary representations can be applied in one field of development (e.g., featural information), children might not be able to apply them in another field of development (e.g., temporal information). This interpretation implies domain-specific, rather than domain-general pathways of self-development and is thus empirically supported by the results of this study showing very specific relations. During the SF task, as well as during the mirror rouge test children have to use featural information to detect contingency. Consistently, the MSR and the SF gaze effect in this study were also related. In contrast, as is argued by Meltzoff (1990) in social mirroring contingency detection based on the infant's understanding that an interaction partner behaves "like me" rather than on featural information as in MSR is assessed. In DSR, children have to use featural, as well as temporal cues to represent a temporally-extended self. Consistent with this interpretation, both MSR and social mirroring predicted DSR. Further, another explanation would be that social imitation gets increasingly complex and with increasing age it is supposed to signify a higher level of mental state understanding and higher meta-cognitive skills. More specifically, while the early detection of "being imitated" in Meltzoff's task may be based on mere surface detection of temporal contingency with increasing age it might reflect infants' intention understanding (Tomasello, 1995; Striano and Rochat, 1999). Thus, the infant, who understands that the interaction partner systematically matches his intended acts, will thereupon attribute the *intention* to behave "like me" to the imitating partner. Similarly, Agnetta and Rochat (2004) argued that while at 9 months the discrimination of an imitating adult is based primarily on the detection of contingency and a sense of self-agency, toward the end of the first of year of life it is based on intention understanding. Thus, social mirroring at 12 months is likely to already reflect higher cognitive processes. Consistent with this interpretation, social mirroring skills predicted DSR as the cognitively more demanding measure of self-awareness.

While experimental research needs to identify the underlying cognitive mechanisms of the developmental links identified in this study, it seems plausible that children's underlying metacognitive awareness of being intentionally imitated as is the case in the social mirroring task at 12 months and their metacognitive awareness of having experienced what can be seen on the video at 4 years of age, may be the source for the developmental link between the tasks. Longitudinal research, using metacognition

as an outcome measure, could shed light on that interpretation. Finally, while the present longitudinal study only started at the age of 9 months, theories of self-development have viewed the first few months of life as particularly important. For instance, children's awareness of their own body as an object and its relations to other objects has been conceptualized as a unique part of self-awareness (e.g., Brownell et al., 2007; Moore et al., 2007). Rochat (2009) proposed that it is from the second month of age that infants engage in reciprocal exchanges and thereby begin to objectify themselves as objects of shared attention. Already at the age of 6 weeks, according to Rochat (2003), the "situated self" is established, a sense of how one's own body is situated in relation to other entities in the environment. Therefore, to further complement the picture, future longitudinal research on developmental precursors of MSR and DSR should include measures of social responsiveness in the first months of life.

The present longitudinal study supports the notion that MSR seems to develop earlier than the permanent self which involves the understanding that the self is invariant over time (Povinelli et al., 1996; Suddendorf, 1999). Importantly, since identification of the mirror image is tied to the temporal simultaneity of the body and its mirror reflection, future research needs to shed light on the exact nature of the capacities underlying the ability to recognize one's own mirror image. Interestingly, while related to different precursor abilities, MSR and DSR were found to be moderately developmentally related to each other, indicating conceptual continuity. This result is in support of Rochat (2003) theory of self-development. Rochat compares the self to an onion with different layers. Thus, the self as a whole always comprises earlier and later-developing stages. More specifically, one possible developmental mechanism linking MSR and DSR proposed by Gallup (1998; cited in Nielsen and Dissanayake, 2004) is that during MSR children need to engage in introspection in order to focus on themselves and to become the object of their own attention. Thus, self-recognizers possess higher introspection skills than non-recognizers. Subsequently, if applying introspection across a variety of contexts during development, combined with meta-representational abilities, introspection should not only promote children's time-restricted self, but also their temporally extended sense of self.

A limitation of this study is that it did not include a socially or culturally diverse sample. Cross-cultural research comparing cultures with distal and proximal parenting styles, as well as research in clinical samples need to corroborate the developmental links identified in this study. Are deficits in social responsiveness and imitation skills (e.g., in autistic children) related to decreased MSR- and DSR-skills?

In sum, the present longitudinal findings show that infants make use of their social world to form an understanding of who they are. This seems to result in very specific, rather than general developmental links between early social responsiveness and children's later understanding of self. To further explore these specific developmental pathways more longitudinal work, focusing on interindividual differences, is clearly needed. Note that the same developmental links identified in self-development might also be found when studying children's early understanding of others (cf., Moore and Corkum, 1994). Similarly, on the neural level, there seems to be a shared, while not completely overlapping, representation network for self and other (Decety and Sommerville, 2003). Thus, future studies should look simultaneously at the development of self and other. In sum, the present findings suggest that one of the multi-faceted interrelations between self-understanding and other-understanding originates from infants' understanding of the other's intention to "act like me" which seems to lay the ground for an advanced concept of the temporally extended self.

Acknowledgment

References

Agnetta, B., and Rochat, P. (2004). Imitative games by 9-, 14-, and 18-month-old infants. *Infancy* 6, 1–36. doi: 10.1207/s15327078in0601_1

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Dev. Psychobiol.* 5, 297–305. doi: 10.1002/dev.420050403

Aschersleben, G., Hofer, T., and Jovanovic, B. (2008). The link between infant attention to goal directed action and later theory of mind abilities. *Dev. Sci.* 11, 862–868. doi: 10.1111/j.1467-7687.2008.00736.x

Asendorpf, J. B., and Baudonniére, P.-M. (1993). Self-awareness and other-awareness: mirror recognition and synchronic imitation among unfamiliar peers. *Dev. Psychol.* 29, 88–95. doi: 10.1037/0012-1649.29.1.88

Asendorpf, J. B., Warkentin, V., and Baudonniére, P. M. (1996). Self-awareness and other-awareness. II: mirror self-recognition, social contingency awareness, and synchronic imitation. *Dev. Psychol.* 32, 313–321. doi: 10.1037/0012-1649.32.2.313

Baron, R. M., and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173

Bischof-Köhler, D. (1988). Über den zusammenhang von empathie und der fähigkeit, sich im spiegel zu erkennen (On the relation between empathy and mirror self-recognition). *Schweiz. Z. Psychol.* 47, 147–159.

Bischof-Köhler, D. (1991). "The development of empathy in infants," in *Infant Development: Perspectives from German-Speaking Countries*, eds M. E. Lamb and H. Keller (Hillsdale: Lawrence Erlbaum), 245–273.

Broesch, T. L., Callaghan, T., Henrich, J., Murphy, C., and Rochat, P. (2010). Cultural variations in children's mirror self-recognition. *J. Cross Cult. Psychol.* 42, 1018–1029. doi: 10.1177/0022022110381114

Brownell, C. A., Zerwas, S., and Ramani, G. B. (2007). "So big": the development of body self-awareness in toddlers. *Child Dev.* 78, 1426–1440. doi: 10.1111/j.1467-8624.2007.01075.x

Courage, M. L., and Howe, M. L. (2002). From infant to child: the dynamics of cognitive change in the second year of life. *Psychol. Bull.* 128, 250–277. doi: 10.1037/0033-2909.128.2.250

Damon, W., and Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child Dev.* 53, 841–864. doi: 10.2307/1129122

Decety, J., and Sommerville, J. A. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends Cogn. Sci.* 7, 527–533. doi: 10.1016/j.tics.2003.10.004

Gallup, G. G. (1998). Self-awareness and the evolution of social intelligence. *Behav. Process.* 42, 239–247. doi: 10.1016/S0376-6357(97)00079-X

Gallup, G. G. Jr., Anderson, J. R., and Shillito, D. J. (2002). "The mirror test," in *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*, eds M. Bekoff, C. Allen, and G. M. Burghardt (Cambridge, MA; MIT Press), 325–333.

Gallup, G. G. Jr., and Suarez, S. D. (1986). "Self-awareness and the emergence of mind in humans and other primates," in *Psychological Perspectives on the Self*, Vol. 3, eds J. Suls and A. G. Greenwald (Hillsdale, NJ; Erlbaum), 3–26.

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biol. Rev.* 69, 207–231. doi: 10.1111/j.1469-185X.1994.tb01506.x

Keller, H., Kärtner, J., Borke, J., Yovsi, R., and Kleis, A. (2005). Parenting styles and the development of the categorical self: a longitudinal study on mirror self-recognition in Cameroonian Nso and German families. *Int. J. Behav. Dev.* 29, 496–504. doi: 10.1080/01650250500147485

Keller, H., Yovsi, R., Borke, J., Kärtner, J., Jensen, H., and Papaligoura, Z. (2004). Developmental consequences of early parenting experiences: self-recognition and self-regulation in three cultural communities. *Child Dev.* 75, 1745–1760. doi: 10.1111/j.1467-8624.2004.00814.x

Lewis, M. (2003). The emergence of consciousness and its role in human development. *Ann. N. Y. Acad. Sci. U.S.A.* 1001, 104–133. doi: 10.1196/annals.1279.007

Lewis, M., and Brooks-Gunn, J. (1979). *Social Cognition and the Acquisition of Self.* New York, NY: Plenum Press.

Lind, S. E., and Bowler, D. M. (2009). Recognition memory, self-other source memory, and theory-of-mind in children with autism spectrum disorder. *J. Autism. Dev. Disord.* 39, 1231–1239. doi: 10.1007/s10803-009-0735-2

Meltzoff, A. N. (1990). "Foundations for developing a concept of self: the role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy," in *The Self in Transition: Infancy to Childhood*, eds D. Cicchetti and M. Beeghly (Chicago: University of Chicago press), 139–164.

Mesman, J., van Ijzendoorn, M. H., and Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: a review and meta-analysis. *Dev. Rev.* 29, 120–162. doi: 10.1016/j.dr.2009.02.001

Mitchell, R. W. (1993). Mental models of mirror-self-recognition: two theories. *New Ideas Psychol.* 11, 295–325. doi: 10.1016/0732-118X(93)90002-U

Moore, C. (2007). "Understanding self and others in the second year. Socioemotional development in the toddler years: transitions and transformations," in *Transition in Early Socioemotional Development: The Toddler Years*, eds C. A. Brownell and C. B. Kopp (New York, NY: Guilford Press), 43–65.

Moore, C., and Corkum, V. (1994). Social understanding at the end of the first year of life. *Dev. Rev.* 14, 349–372. doi: 10.1006/drev.1994.1014

Moore, C., Mealiea, J., Garon, N., and Povinelli, D. J. (2007). The development of body self-awareness. *Infancy* 11, 157–174. doi: 10.1111/j.1532-7078.2007.tb00220.x

Neisser, U. (1995). "Criteria for an ecological self," in *The Self in Infancy. Advances in Psychology Book Series*, ed. P. Rochat (Amsterdam: Elsevier), 17–34.

Nielsen, M., and Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: a longitudinal investigation through the second year. *Infant Behav. Dev.* 27, 342–365. doi: 10.1016/j.infbeh.2003.12.006

Nielsen, M., Dissanayake, C., and Kashima, Y. (2003). A longitudinal investigation of self–other discrimination and the emergence of mirror self-recognition. *Infant Behav. Dev.* 26, 213–226.

Nielsen, M., Suddendorf, T., and Slaughter, V. (2006). Mirror self-recognition beyond the face. *Child Dev.* 77, 176–185. doi: 10.1111/j.1467-8624.2006.00863.x

Parker, S. T. E., Mitchell, R. W., and Boccia, M. L. (1994). Self-awareness in animals and humans: developmental perspectives. *In Based on Papers Presented at a Conference on Self-Awareness in Monkeys, Apes, and Humans*, Sonoma State U, Rohnert, CA, 1991. Cambridge University Press.

Perner, J. (1991). *Understanding the Representational Mind.* Cambridge, MA: The MIT Press.

Petermann, F. (2009). *WPPSI-III. Wechsler Preschool and Primary Scale of Intelligence–Third Edition. Deutsche Version.* Frankfurt: Pearson Assessment.

Povinelli, D. J. (2001). "The self: elevated in consciousness and extended in time," in *The Self in Time: Developmental Perspectives*, eds C. Moore and K. Lemmon (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 75–95.

Povinelli, D. J. (1995). "The unduplicated self," in *The Self in Early Infancy*, ed. P. Rochat (Amsterdam: Elsevier), 161–192.

Povinelli, D. J., Landau, K. R., and Perilloux, H. K. (1996). Self-recognition in young children using delayed versus live feedback: evidence of a developmental asynchrony. *Child Dev.* 67, 1540–1554. doi: 10.1111/j.1467-8624.1996.tb01813.x

Povinelli, D. J., Rulf, A. B., Landau, K. R., and Bierschwale, D. T. (1993). Self-recognition in chimpanzees (*Pan troglodytes*): distribution, ontogeny, and patterns of emergence. *J. Comp. Psychol.* 107, 347–372. doi: 10.1037/0735-7036.107.4.347

Reddy, V. (2003). On being the object of attention: implications for self–other consciousness. *Trends Cogn. Sci.* 7, 397–402. doi: 10.1016/S1364-6613(03)00191-8

Rochat, P. (2003). Five levels of self-awareness as they unfold early in life. *Conscious. Cogn.* 12, 717–731. doi: 10.1016/S1053-8100(03)00081-3

Rochat, P. (2009). *Others in Mind: Social Origins of Self-Consciousness.* Cambridge: University Press.

Rosenblum, K. L., McDonough, S., Muzik, M., Miller, A., and Sameroff, A. (2002). Maternal representations of the infant: associations with infant response to the still face. *Child Dev.* 73, 999–1015. doi: 10.1111/1467-8624.00453

Sommerville, J. A., Woodward, A. L., and Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96, B1–B11. doi: 10.1016/j.cognition.2004.07.004

Striano, T., and Rochat, P. (1999). Developmental link between dyadic and triadic social competence in infancy. *Br. J. Dev. Psychol.* 17, 551–562. doi: 10.1348/026151099165474

Striano, T., and Stahl, D. (2005). Sensitivity to triadic attention in early infancy. *Dev. Sci.* 8, 333–343. doi: 10.1111/j.1467-7687.2005.00421.x

Suddendorf, T. (2003). Early Representational Insight: Twenty?Four?Month?Olds Can Use a Photo to Find an Object in the World. *Child Dev.* 74, 896–904.

Suddendorf, T. (1999). Children's understanding of the relation between delayed video representation and current reality: a test for self-awareness? *J. Exp. Child Psychol.* 72, 157–176. doi: 10.1006/jecp.1998.2485

Suddendorf, T., and Butler, D. L. (2013). The nature of visual self-recognition. *Trends Cogn. Sci.* 17, 121–127. doi: 10.1016/j.tics.2013.01.004

Suddendorf, T., Simcock, G., and Nielsen, M. (2007). Visual self-recognition in mirrors and live videos: evidence for a developmental asynchrony. *Cogn. Dev.* 22, 185–196. doi: 10.1016/j.cogdev.2006.09.003

Thoermer, C., Sodian, B., Vuori, M., Perst, H., and Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *Br. J. Dev. Psychol.* 30, 172–187. doi: 10.1111/j.2044-835X.2011.02067.x

Tomasello, M. (1995). "Joint attention as social cognition," in *Joint Attention: Its Origins and Role in Development*, eds C. Moore and P. J. Dunhum (Hillsdale: Lawrence Erlbaum), 103–130.

Tronick, E., Als, H., Adamson, L., Wise, S., and Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *J. Am. Acad. Child. Psychiatry* 17, 1–13. doi: 10.1016/S0002-7138(09)62273-1

Troseth, G. L. (2003). Getting a clear picture: young children's understanding of a televised image. *Dev. Sci.* 6, 247–253. doi: 10.1111/1467-7687.00280

Vyt, A. (2001). Processes of visual self-recognition in infants: experimental induction of 'mirror' experience via video self-image presentation. *Infant Child Dev.* 10, 173–187. doi: 10.1002/icd.277

Wellman, H. M., Lopez-Duran, S., La Bounty, J., and Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Dev. Psychol.* 44, 618–623. doi: 10.1037/0012-1649.44.2.618

Zmyj, N., Prinz, W., and Daum, M. M. (2013). The relation between mirror self-image reactions and imitation in 14-and 18-month-old infants. *Infant Behav. Dev.* 36, 809–816.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The neural basis of non-verbal communication—enhanced processing of perceived give-me gestures in 9-month-old girls

Marta Bakker[1]*, Katharina Kaduk[2], Claudia Elsner[1], Joshua Juvrud[1] and Gustaf Gredebäck[1]

[1] Uppsala Child and Baby Lab, Department of Psychology, Uppsala University, Uppsala, Sweden
[2] Department of Psychology, Fylde College, Lancaster University, Lancaster, UK

This study investigated the neural basis of non-verbal communication. Event-related potentials were recorded while 29 nine-month-old infants were presented with a give-me gesture (experimental condition) and the same hand shape but rotated 90°, resulting in a non-communicative hand configuration (control condition). We found different responses in amplitude between the two conditions, captured in the P400 ERP component. Moreover, the size of this effect was modulated by participants' sex, with girls generally demonstrating a larger relative difference between the two conditions than boys.

Keywords: give-me gesture, ERP, P400, sex differences, non-verbal communication, social perception, infancy

INTRODUCTION

Gestures may be used as social tools for expressing one's own feelings and thoughts, cooperating with others, and drawing others' attention to objects and events (Tomasello et al., 2007; Carpendale and Carpendale, 2010). In early childhood gestures may be expressed in grimaces and smiles (Caselli, 1990) and are later exhibited with fingers, hands, and arms (Crais et al., 2004). By the end of the first year of life, gestures such as giving (Mundy et al., 1986; Caselli, 1990) or pointing (Bates et al., 1975; Tomasello, 2008) become meaningful for expressing goals and communicating with others.

Research exploring the development of the pointing gesture is quite prevalent (e.g., Butterworth, 2003; Camaioni et al., 2004; von Hofsten et al., 2005; Liszkowski et al., 2006; Tomasello et al., 2007; Daum et al., 2013). In contrast, the give-me gesture (a face-up palm directed toward the observer; Mundy et al., 1986) has received little attention. We believe that the give-me gesture warrants more interest from the scientific community considering its communicative importance in serving multiple functions, such as referring to a specific object, expressing a request and communicating an action goal (Shwe and Markman, 1997).

From a behavioral perspective, we know that children begin to give and request objects to and from others at around 9- to 12-month of age (Bates et al., 1975; Masur, 1983; Carpenter et al., 1998; Crais et al., 2004). Recent eye tracking studies show that infants are sensitive to the communicative properties of the give-me gesture by 12-month of age (Elsner et al., 2014). In this study, infants observed a give-and-take interaction between two individuals. At the beginning of each trial the receiving hand formed either a give-me gesture or an inverted hand shape

(hand shaped as a give-me gesture but presented upside-down). Subsequently, the passing hand (hand from another individual) that was located on the opposite side of the screen transferred the ball to a receiving hand. The authors assessed differences in latency of goal-directed gaze shifts from the hand transporting the ball to the receiving hand. The results revealed that infants shifted their gaze significantly earlier toward the goal, the receiving hand, if it was shaped as a give-me gesture in comparison to an inverted hand shape. Additional control conditions ruled out that the effect was based on affordance, e.g., a simple match between the ball and the receiving hand, or attentional differences (Elsner et al., 2014). Jointly, the results indicate that infants are sensitive to the communicative intent of a hand shaped in a give-me gesture. Another eye tracking study demonstrated that 14-month-old infants have a clear expectation of adequate responses to the give-me gesture. That is, when observing an interaction between two people, infants anticipate that an object will be passed to another person when the give-me gesture request is presented, suggesting again that infants at this age can recognize the communicative intent of the gesture (Thorgrimsson et al., 2014). Interestingly, perception of give-me gestures may be different for typically developing children than children with autism spectrum diagnosis (ASD). In a recent study 5- to 6-year old children with ASD were found to look differently at social interactions incorporating give-me gestures differentially than typically developing children (Falck-Ytter et al., 2013). This may suggest that children with this clinical diagnosis might be less able to read the meaning of the give-me gesture or that they are less interested in the people's reactions that are confronted with give-me gestures (Falck-Ytter et al., 2013).

Motivated by eye tracking studies that highlight the importance of the give-me gesture in goal understanding and encoding social interaction during development (Falck-Ytter et al., 2013; Elsner et al., 2014; Thorgrimsson et al., 2014), as well as a desire to learn more about the neural mechanisms that are involved in processing of give-me gestures, the current study investigated the neural activation that is evoked when observing give-me gestures. To our knowledge, only two studies have investigated the neural correlates of gesture perception early in development. The first study investigated the neurodevelopment of pointing perception (Gredebäck et al., 2010), whereas the second the perception of grasping gestures (Bakker et al., 2014). In those studies, the authors reported the ERP component P400 to be sensitive to the congruency of pointing or grasping, revealing higher mean amplitudes for the congruent (gestures directed toward an object) compared to the incongruent condition (gestures directed away from the object). Here, we aim to explore if the same ERP component generalizes over communicative settings, from hand configurations directed toward objects (pointing; Gredebäck et al., 2010, and grasping; Bakker et al., 2014) to more socially oriented gestures, in this case the give-me gesture directed toward the infant. If the same underlying neural processes are involved in processing of a large array of gestures, than we would expect larger amplitudes of the P400 for the give-me gesture than a hand configuration that is perceptually very similar but has no communicative intent (from here labeled as non-communicative hand configuration).

In addition, we aim to investigate the relation between infants' neural response to the give-me gesture and infants' own ability to respond to the same gesture on a behavioral level. Prior work has demonstrated that infants process both pointing (Gredebäck et al., 2010) and grasping gestures (Bakker et al., 2014) by 9 months of age. At the same age, infants also start to engage in producing give-me gestures (Bates et al., 1975; Masur, 1983; Carpenter et al., 1998; Crais et al., 2004). Based on the revealed correspondence between infants' neural potentials and behavior in prior EEG studies (i.e., Bakker et al., 2014), the current study targets both 9-month-old infants' neural correlates of the give-me gesture and their behavioral responses to give-me requests (Responding to Behavioral Request procedure from the Early Social Communication Scales; Mundy et al., 2003). We expect that behavioral responses to the give-me gesture will correspond with P400 amplitudes. That is, relative amplitudes (give-me gesture vs. non-communicative hand configuration) should be higher in infants that are proficient in responding behaviorally to the give-me gesture.

Further analyses in this study explored individual differences in gesture perception with respect to infants' sex. Based on prior studies revealing that girls are ahead of boys in the onset of gesture and language production (Butterworth and Morissette, 1996; Özçalışkan and Goldin-Meadow, 2010), it is possible that girls are more proficient in discriminating between the give-me gesture and the non-communicative hand configuration than boys. If we find such an effect we would expect an interaction effect between sex and condition. That is, both boys and girls should be able to differentiate between the two conditions, but we would expect the effect to be bigger in girls than boys.

In summary, the current study has three aims: to investigate the give-me gesture perception on a neural level, to investigate infants' behavioral response to the give-me gesture and to investigate the presence of sex differences in social perception mechanisms.

MATERIALS AND METHODS

PARTICIPANTS

The final sample consisted of twenty-nine 9-month-olds (15 girls, mean age 8 months and 28 days, SD = 6 days). An additional 30 infants (16 girls) participated but were excluded due to fussiness (less than 10 artifact-free trials, $n = 25$) or technical problems ($n = 5$). Parents completed informed consent prior to participation and received a gift voucher of approximately 10° for participating. The study was conducted in accordance with the standards specified in the 1964 Declaration of Helsinki and approved by the local ethics committee.

EEG STIMULI

The give-me gesture (experimental condition) and the non-communicative hand configuration (control condition) were presented to the infants. In both conditions the stimulus included a hand (palm facing upward in the experimental condition and the same hand rotated 90° in the control condition). Stimuli were presented at random (with the constrains of maximum three repetitions of the same stimulus) and presented in the middle of a gray background for 1000 ms. Between each experimental stimulus; a fixation cross was presented for 100–300 ms (see **Figure 1**). Infants viewed the stimuli (20.7 × 16.5 visual degrees) on a 17-inch computer monitor at a viewing distance of 60 cm. The size of the hand was 5 horizontal and 16 vertical visual degrees. The stimuli were presented using the E-Prime 2.0, E-Studio software (Psychology Software Tools, Inc., Pittsburgh, PA, USA).

BEHAVIORAL TASK

Parents were asked if they have observed their child producing or responding to the give-me gesture outside of the laboratory. Subsequently, a researcher assessed infant's behavioral response to the give-me gesture using the Responding to Behavioral Request procedure from the ESCS (Mundy et al., 2003). The experimenter first familiarized the infant with three rubber toys (5 × 5 cm) and then placed the toys in front of the infant and waited (3 s) for the infant to give the toy back spontaneously. If the infant did not pass a toy, the experimenter verbally requested the toys with
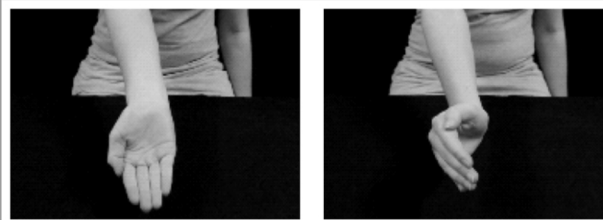
FIGURE 1 | Stimulus for the give-me gesture condition on the left and a control hand on the right.

the phrase: "give it to me." If after 3 s the infant did not respond to the verbal request, the experimenter used a combination of verbal request together with a non-verbal give-me gesture. The experimenter's gesture stopped within reach of the infant. The infant's behavior was video recorded and later assessed for the frequency of appropriate responses, that is, the number of times the child gave a toy to the experimenter at the request (verbal or verbal in combination with the give-me gesture). The total duration of this grasping test did not exceed 5 min.

PROCEDURE

During the lab visit, we first recorded infants' neural responses to the give-me gesture, followed by a behavioral task that measured the ability to respond to the give-me gesture. During the EEG recording, infants sat on their parent's lap approximately 60 cm from the stimulus monitor. The experimenter sat at a control computer separated from the parent and infant by a curtain and monitored the infant's behavior via a live camera. The researcher paused the experiment if the infant became inattentive and fussy. The stimulus monitor remained black for the duration of the pause. The experimenter terminated the study when the infant was no longer interested in the stimuli. After the EEG recording the parent and infant were given an approximate 5 min break before proceeding with the behavioral response task. This paper reports data from an ongoing longitudinal project looking at the neural correlates of social cognition and later language development.

EEG RECORDING AND ANALYSIS

We used 128-channel HydroCel Geodesic Sensor Nets to record infants' EEG. The recorded signal (250 Hz, vertex referenced) was amplified by an EGI Net Amps 300 amplifier (Electric Geodesic, Eugene, OR) and stored for off-line analysis. The EEG signal was digitally filtered (0.3–30 Hz) and segmented from 200 ms prior to the appearance of the hand to 1000 ms after the onset of the stimulus. Off-line inspection of video recordings ensured that only trials in which infants paid attention were further processed. The data was manually edited for artifacts (standard procedure for infant ERP studies, see Hoehl and Wahl, 2012). Trials with excessive noise levels (mostly due to movement artifacts) were rejected. Channels with moderate noise levels were reconstructed from an interpolation of surrounding electrodes. All included trials contained no more than 10% interpolated channels. The whole recording session did not exceed 10 min. The inclusion criterion for the final analysis was at least 10 artifact free trials per condition (standard inclusion criterion for infants ERP studies, see DeBoer et al., 2007; Stets et al., 2012). On average, an infant saw 90 trials across both conditions, with 44 trials for the give-me gesture condition and 46 for the control hand. After visual data inspection and manual data editing, a mean of 15 artifact free trials remained (range: 10–31) for the give-me gesture condition and a mean of 17 trails (range: 10–32) for the control hand. Finally, we baseline corrected and averaged all artifact free trials, as well as re-referenced to the average in order to create individual averages for each participant, as well as calculated grand averages from individual averages. Based on the visual inspection of the individual averages and grand average we selected 11 channels in the posterior area (62, 67, 70 71, 72, 74, 75, 76, 77, 82, 83) for statistical analyses. We captured three components in the ERP wave morphology after the stimulus onset, and performed the analysis in the following three time windows (see **Figure 2**): P1 (80–140 ms), N200 (150–250 ms) and P400 (300–600 ms). We conducted analyses of variance (ANOVAs) to compare the mean amplitudes between conditions (the give-me gesture and control) in all ERP components (P1, N200, P400) and to assess the effect of sex on ERP amplitude differences, respectively.

FIGURE 2 | Grand-average of ERP of the posterior area (channels of interest are marked in black). Black line represents the give-me gesture condition and grey line the control hand.
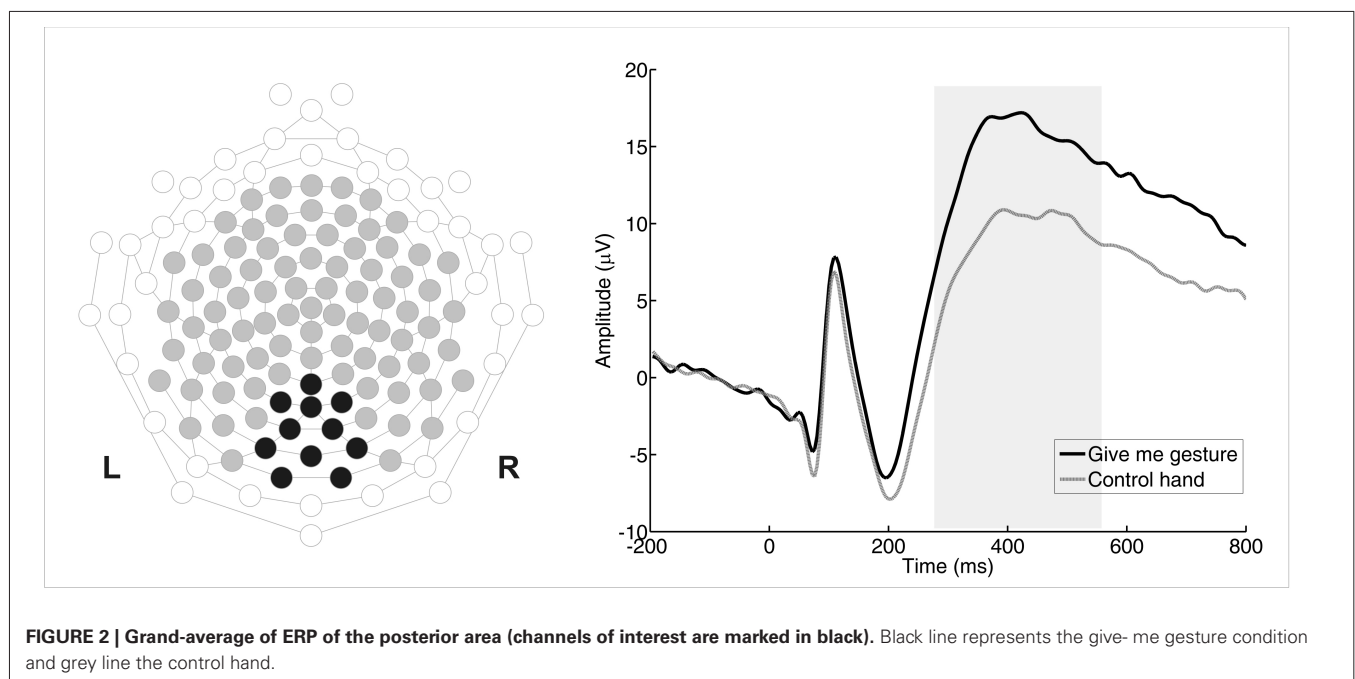
RESULTS

ERPs

Our first ERP analysis focused on the component of interest, the P400. In order to test the possible difference between the conditions as well as the effect of sex on the modulation of the P400 amplitude, we conducted a 2(sex) × 2(condition) mixed repeated measures ANOVA. Results revealed a main effect of condition $F_{(1,27)} = 40.12$, $p < 0.001$, $\eta^2 = 0.598$, with a mean amplitude of 15 μV (SD = 6 μV) in response to the give-me gesture and 9 μV (SD = 7 μV) in response to seeing the non-communicative hand configuration. Overall, 26 out of 29 infants demonstrated larger amplitudes for the give-me gesture compared to the non-communicative hand configuration. Additionally, there was a significant interaction between Condition and Sex [$F_{(1,27)} = 5.384$, $p = 0.028$, $\eta^2 = 0.166$; see **Figure 3**]. To inspect the condition by sex interaction, we performed planned comparison paired-samples t-tests (separately for each sex). Results revealed significant differences between conditions, both for girls [$t_{(27)} = 4.750$, $p < 0.001$] as well as for boys [$t_{(27)} = 4.360$, $p < 0.001$] with more positive mean amplitudes for the give-me gesture. As both boys and girls displayed a significant difference in their response to the two gestures, and as the direction of the difference was similar, it is possible that the interaction between Sex and Condition stems from differences in the size of the effect. To test this prediction, we further examined the difference between the sexes in their conditional amplitude difference scores. We performed an independent-samples t-test with the amplitude difference as a dependent variable and sex as a grouping variable. The analysis revealed a significant amplitude difference between the sexes [$t_{(27)} = 2.320$, $p = 0.028$], This clearly shows that the interaction is driven by the size of the difference between the conditions that is larger for girls (girls: M = 8 μV, SD = 6 μV; boys: M = 10 μV, SD = 8 μV).

To ensure that the effect between conditions as well as the interaction between Condition and Sex is specific to the P400 we performed a follow-up analysis for two other components visible in the ERP wave morphology, i.e., P1 and N200. We performed

a 2 × 2 mixed repeated measures ANOVAs with Condition as a within-subject factor and Sex as between-subject factor on the mean amplitudes of the P1 and N200. The analysis for the P1 component revealed no significant effects, neither for difference between conditions [$F_{(1,27)} = 2.297$, $p = 0.141$, $\eta^2 = 0.078$] nor an interaction between Condition and Sex [$F_{(1,27)} = 2.149$, $p = 0.154$, $\eta^2 = 0.074$]. The analysis for the N200 also failed to show significance, neither for differences between the conditions [$F_{(1,27)} = 2.808$, $p = 0.105$, $\eta^2 = 0.094$] nor for an interaction [$F_{(1,27)} = 0.077$, $p = 0.783$, $\eta^2 = 0.003$].

BEHAVIORAL TASK

On a behavioral level, none of the infants responded to the give-me gesture request as determined by the ESCS scale. Four infants responded by moving the hand with the object to the experimenter but did not release it. Two infants moved the hand away from the experimenter when seeing the request. None of the caregivers reported that their infant was able to produce or respond to the give-me gesture outside the laboratory. Therefore, no statistical analysis was performed.

DISCUSSION

This study investigated infants' neural correlates to the perception of the give-me gesture, a non-verbal communication. As predicted, we found that infants' P400 component increased in amplitude when infants were presented with the give-me gesture compared to a non-communicative hand configuration. This difference was significant despite the fact that most of the infants did not demonstrate an overt sensitivity to the give-me gesture (as measured with ESCS).

The current study is the first to demonstrate neural correlates to give-me gestures in 9-month-old infants. Furthermore, we demonstrate that the neural basis of non-verbal communication, as indexed by the sensitivity to the give-me gesture, develops before overt responses to other people's give-me gestures. It is possible that our results capture an early neural sensitivity that is a functional prerequisite of later overt behavior. As all intentional behavior must have its neural underpinnings, it is possible that the neural support networks must first be in place in order for overt behavior to emerge. For a more immediate connection between referential gesture communication and infants' own motor abilities in the case of grasping, see Bakker et al. (2014). Finally, as predicted, we demonstrate sex differences in the neural responses to the give-me gesture, with larger amplitude difference between conditions in girls than boys.

P400—NEURAL CORRELATE OF THE GIVE-ME GESTURE

In the current study we found that the give-me gesture elicits larger P400 amplitude than the non-communicative hand configuration in 9-month-old infants. This effect is highly similar to the neural response elicited while observing goal-directed pointing (Gredebäck et al., 2010) and grasping (Bakker et al., 2014). In those studies, the amplitude of the P400 was larger for typical and functional referential cues (i.e., give-me gesture, congruent pointing, congruent reaching) than for the control stimuli that were less communicative or functional. Here, we demonstrate similar differences in the amplitude of P400 for gestures directed

FIGURE 3 | Mean amplitude P400 separately for each condition and sex. Red and blue-dashed lines illustrate the interaction between Condition and Sex.
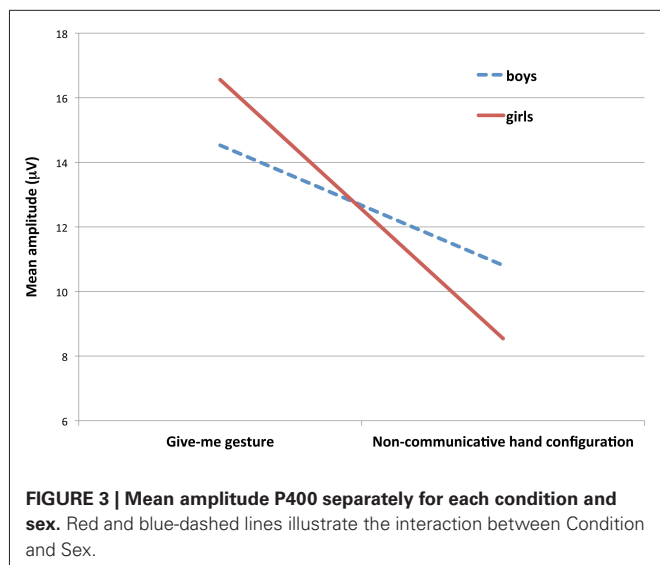
toward the infant. Together, these findings demonstrate that the P400 indexes a wide range of social gestures, comprising both gestures directed toward objects and those directed toward the observing infant.

In contrast to prior studies examining neural correlates in relation to behavioral response of pointing and grasping, we did not find a relation between P400 ERP to give-me gesture and infants' behavioral response to the same gesture. In the prior study on grasping perception (Bakker et al., 2014), 5–6 months old infants' own experience with grasping was closely connected to their ability to encode the relation between the presented object and the grasping hand. More specifically, a difference in the P400 between conditions (hand directed toward or away from the object location) was only evident in infants that were able to perform functional grasping. In the current study, however, infants that did not show a behavioral response to the give-me gesture showed a clear sensitivity in evoked ERPs to this gesture. It is possible that the neural correlates of basic action perception and action production develop simultaneously for actions that emerge early during infancy (like grasping). However, gestures like the give-me gesture are more complex and a proper behavioral response may require more understanding of properties of the gesture and turn-taking in social interactions.

More research is required to further examine the developmental trajectories of the perception and production of give-me gestures. Longitudinal designs investigating the relation between functional and behavioral aspects of give-me gesture perception could provide new perspectives on the development of non-verbal communication and infants' understanding of cooperative actions. Additionally, it would be valuable to gain an understanding on whether the give-me gesture relates to other referential gestures and referential cues on both a behavioral and neural level. The combination of neural and behavioral measures would expand our knowledge about infants' early communicative development, which so far has been limited to pointing, even though infants' gestural repertoire is more extensive.

INDIVIDUAL DIFFERENCES IN PERCEPTION OF GIVE-ME GESTURES

In the current study we found a significantly larger difference between conditions in P400 amplitudes for girls than for boys. This difference is interpreted as an indication that girls might be more sensitive to discriminating give-me gestures from other non-communicative hand configurations. To our knowledge there are no EEG studies that have reported sex differences in social perception in infancy. Some sex differences have, however, been observed in infant studies that used behavioral measures. For instance, differences between boys and girls have been demonstrated in the frequency of eye contact between the child and the mother, with girls making more eye contact than boys (Lutchmaya et al., 2002). It has also been suggested that infant girls may be more attracted to social stimuli than boys, for example when being presented with faces (Lutchmaya and Baron-Cohen, 2002) or abstract geometric shapes chasing each other (Frankenhuis et al., 2013) and faces (Lutchmaya and Baron-Cohen, 2002). In a meta-analytic review of sex differences in facial expression processing in infancy, McClure (2000) reported that females outperformed males in interpreting facial expressions

and other non-verbal cues. These advantages for females are visible both in infancy as well as in adulthood. A recent study that inspected brain activation during observation of biological motion revealed a difference between adult female and male participants, with females showing greater activation in brain regions that are involved in social perception (Anderson et al., 2013). The authors also found a similar trend in children (Anderson et al., 2013). Based on these findings it is likely that the sex differences found in the present study would replicate across a larger range of social perception studies examining neural processes targeting social stimuli. Furthermore, we speculate that the results from this study capture possible sex differences in processing of non-verbal cues. This is in line with previous research that reported females being more accurate in decoding non-verbal cues (Hall, 1978), joint attention and communicative skills (Olafsen et al., 2006). Additionally, Özçalışkan and Goldin-Meadow (2010) found that the onset of gesture and sentence production emerges later in boys than girls. In this context it is important to note that the current study captures sex differences in response to non-verbal social cues at an extremely early age, before the actual onset of gesture and speech production.

Taken together, we believe, that higher average P400 amplitude found in this study was generated by infants' encoding of more communicative intent in the give-me gestures in comparison to non-communicative hand configuration. It is worth mentioning that again that no differences were found in ERP components (P1) that often index pure visual differences in stimuli. Additionally, prior work has also conducted several controls that rule out affordance and visual attention as alternative explanations (Elsner et al., 2014). As a whole, the P400 literature suggests that infants from an early age perceive functional and goal-directed manual actions and gestures in a similar manner. These processes operate both during observation of manual gestures directed toward objects as well as toward the observing infant. All of these events result in larger amplitude modulation in comparison to non-goal directed or non-communicative hand configurations.

In conclusion, the current study is the first to examine neural underpinnings of the give-me gesture. Our findings contribute to the understanding of the P400 neural component suggesting an involvement in encoding social interactions and non-verbal communication. More specifically our study demonstrates that the P400 is sensitive to observation of the give-me gesture with 9-month-old girls demonstrating a larger difference between conditions than 9-month-old boys.

REFERENCES
Anderson, L. C., Bolling, D. Z., Schelinski, S., Coffman, M. C., Pelphrey, K. A., and Kaiser, M. D. (2013). Sex differences in the development of brain mechanisms for processing biological motion. *Neuroimage* 83, 751–760. doi: 10.1016/j.neuroimage.2013.07.040
Bates, E., Camaioni, L., and Volterra, V. (1975). The acquisition of performatives prior to speech. *Merril Palmer Q.* 21, 205–226.

Bakker, M., Daum, M. M., Handl, A., and Gredebäck, G. (2014). Neural correlates of action perception at the onset of functional grasping. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nsu119 [Epub ahead of print].

Butterworth, G. (2003). "Pointing is the royal road to language for babies," in *Pointing: Where Language, Culture, and Cognition Meet*, ed. S. Kita (Mahwah: Erlbaum), 9–33.

Butterworth, G. E., and Morissette, P. (1996). Onset of pointing and the acquisition of language in infancy. *J. Reprod. Infant Psychol.* 14, 219–231. doi: 10.1080/02646839608404519

Camaioni, L., Perucchini, P., Bellagamba, F., and Colonnesi, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy* 5, 299–308. doi: 10.1207/s15327078in0503_3

Carpendale, J. I. M., and Carpendale, A. (2010). The development of pointing: from personal directedness to interpersonal direction. *Hum. Dev.* 6, 110–126. doi: 10.1159/000315168

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Join social and communicative of age 9 to 15 months. *Mongr. Soc. Res. Child Dev.* 63, 1–143. doi: 10.2307/1166214

Caselli, M. C. (1990). "Communicative gestures and first words," in *From Gesture to Language in Hearing and Deaf Children*, eds V. Volterra and C. A. Erting (Springer-Verlag: Berlin Heidelberg), 56–67.

Crais, E., Douglas, D., and Campbell, C. (2004). The intersection of the development of gestures and intentionality. *J. Speech Lang. Hear Res.* 47, 678–695. doi: 10.1044/1092-4388(2004/052)

Daum, M. M., Ulber, J., and Gredebäck, G. (2013). The development of pointing pin Infancy: effects of communicative signals on covert shifts of attention. *Dev. Psychol.* 49, 1898–1908. doi: 10.1037/a0031111

DeBoer, T., Scott, L. S., and Nelson, C. A. (2007). "Methods for acquiring and analyzing infant event-related potentials," in *Infant EEG and Event-Related Potentials*, ed. M. de Haan (New York, NY: Psychology Press), 5–37.

Elsner, C., Bakker, M., Rohlfing, K., and Gredebäck, G. (2014). Infants' online perception of give-and-take interactions. *J. Exp. Child Psychol.* 126, 280–294. doi: 10.1016/j.jecp.2014.05.007

Falck-Ytter, T., von Hofsten, C., Gillberg, C., and Fernell, E. (2013). Visualization and analysis of eye movement data from children with typical and atypical development. *J. Autism. Dev. Disord.* 43, 2249–2258. doi: 10.1007/s10803-013-1776-0

Frankenhuis, W. E., House, B., Clark Barrett, H., and Johnson, S. P. (2013). Infants' perception of chasing. *Cognition* 126, 224–233. doi: 10.1016/j.cognition.2012.10.001

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

Hall, J. A. (1978). Hall gender effects in decoding non-verbal cues. *Psychol. Bull.* 85, 845–857. doi: 10.1037/0033-2909.85.4.845

Hoehl, S., and Wahl, S. (2012). Recording infant ERP data for cognitive research. *Dev. Neuropsychol.* 37, 187–209. doi: 10.1080/87565641.2011.627958

Liszkowski, U., Carpenter, M., and Tomasello, M. (2006). Pointing out new news, old news, and absent referents. *Dev. Sci.* 10, F1–F7. doi: 10.1111/j.1467-7687.2006.00552.x

Lutchmaya, S., and Baron-Cohen, S. (2002). Human sex differences in social and non-social looking preferences, at 12 months of age. *Infant Behav. Dev.* 25, 319–325. doi: 10.1016/S0163-6383(02)00095-4

Lutchmaya, S., Baron-Cohen, S., and Raggatt, P. (2002). Foetal testosterone and eye contact in 12-month-old infants. *Infant Behav. Dev.* 25, 327–335. doi: 10.1016/S0163-6383(02)00094-2

Masur, E. F. (1983). Gestural development, dual-directional signaling, and the transition to words. *J. Psycholinguist. Res.* 12, 93–109. doi: 10.1007/BF01067406

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychol. Bull.* 126, 424–453. doi: 10.1037/0033-2909.126.3.424

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., and Seibert, J. (2003). *A Manual for Early Social Communication Scales (ESCS)*. Coral Gables, FL: University of Miami

Mundy, P., Sigman, M. D., Ungerer, J., and Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *J. Child Psychol. Psychiatry* 27, 657–669. doi: 10.1111/j.1469-7610.1986.tb00190.x

Olafsen, K. S., Rønning, J. A., Kaaresen, P. I., Ulvund, S. E., Handegård, B. H., and Dahl, L. B. (2006). Joint attention in term and preterm infants at 12 months corrected age: The significance of gender and intervention based on a randomized controlled trial. *Infant Behav. Dev.* 29, 554–563. doi: 10.1016/j.infbeh.2006.07.004

Özçalışkan, Ş., and Goldin-Meadow, S. (2010). Sex differences in language first appear in gesture. *Dev. Sci.* 13, 752–760. doi: 10.1111/j.1467-7687.2009.00933.x

Shwe, H. I., and Markman, E. M. (1997). Young children's appreciation of the mental impact of their communicative signals. *Dev. Psychol.* 33, 630–636. doi: 10.1037/0012-1649.33.4.630

Stets, M., Stahl, D., and Reid, V. M. (2012). A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Dev. Neuropsychol.* 37, 226–252. doi: 10.1080/87565641.2012.654867

Thorgrimsson, G. B., Fawcett, C., and Liszkowski, U. (2014). Infants' expectations about gestures and actions in third-party interactions. *Front. Psychol.* 5:321. doi: 10.3389/fpsyg.2014.00321

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge: The MIT Press.

Tomasello, M., Carpenter, M., and Liszkowski, U. (2007). A new look at infant pointing. *Child Dev.* 78, 705–722. doi: 10.1111/j.1467-8624.2007.01025.x

von Hofsten, C., Dahlström, E., and Fredriksson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. *Infancy* 8, 217–231. doi: 10.1207/s15327078in0803_2

Shifting goals: effects of active and observational experience on infants' understanding of higher order goals

Sarah A. Gerson [1,2,3]*, Neha Mahajan [2,4], Jessica A. Sommerville [5], Lauren Matz [2] and Amanda L. Woodward [2,6]

[1] University of St Andrews, Saint Andrews, UK, [2] University of Maryland, College Park, MD, USA, [3] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands, [4] Portland State University, Portland, OR, USA, [5] University of Washington, Seattle, WA, USA, [6] University of Chicago, Chicago, IL, USA

Action perception links have been argued to support the emergence of action understanding, but their role in infants' perception of distal goals has not been fully investigated. The current experiments address this issue. During the development of means-end actions, infants shift their focus from the means of the action to the distal goal. In Experiment One, we evaluated whether this same shift in attention (from the means to the distal goal) when learning to produce multi-step actions is reflected in infants' perception of others' means-end actions. Eight-months-old infants underwent active training in means-end action production and their subsequent analysis of an observed means-end action was assessed in a visual habituation paradigm. Infants' degree of success in the training paradigm was related to their subsequent interpretation of the observed action as directed at the means versus the distal goal. In Experiment Two, observational and control manipulations provided evidence that these effects depended on the infants' active engagement in the means-end actions. These results suggest that the processes that give rise to means-end structure in infants' motor behavior also support the emergence of means-end structure in their analysis of others' goals.

Keywords: action understanding, action perception links, means-end actions, social cognition, motor learning, infant cognition

Introduction

Human infants are highly attentive and responsive to their social partners. They are also cognitively engaged with them. Research over the last decade has revealed that infants encode others' behavior not just as physical motions through space but rather as actions structured by goals (see Meltzoff, 2007; Woodward et al., 2009 for reviews). This sensitivity to the goal structure of action is a cornerstone of social cognition, providing the foundation for social learning (Tomasello, 1999; Baldwin and Moses, 2001) and theory of mind (Wellman et al., 2004, 2008) in early childhood. Given the importance of infants' goal sensitivity, recent research has investigated the factors that support its development during infancy. One insight from this research is the finding that infants' own experience acting in goal-directed ways seems to inform their sensitivity to others' action goals (e.g., Sommerville et al., 2005, 2008). In the studies reported here, we investigate this process, asking whether and how infants' own actions may inform their sensitivity to distal goals in others' actions.

At a basic level of analysis, adults understand actions as directed at the objects that are the proximal targets of the action. For example, imagine a man reaching across a crowded countertop to grasp a spoon. Adults view this action as organized by the relation between the man and the spoon, rather than in terms of its other perceivable attributes (e.g., the reach trajectory, speed of reach, etc.). Infants perceive this action in the same manner by the time they are 6 months of age. For example, when infants in a visual habituation experiment view a repeated goal-directed action (e.g., a person grasping a toy) they subsequently show selective recovery (longer looking) to test events in which the relation between the person and her goal is disrupted compared to trials on which the person's movements differ but her goal remains the same (e.g., Woodward, 1998; Biro and Leslie, 2006; Brandone and Wellman, 2010; Thoermer et al., 2013). Infants' selective attention to the goal structure of others' actions has also been revealed using measures of behavioral imitation, visual anticipation, and neural activity (e.g., Hamlin et al., 2008; Southgate et al., 2009; Cannon and Woodward, 2012; Krogh-Jespersen and Woodward, 2014).

Perceiving meaningful structure in others' actions requires more than the ability to encode single actions as goal-directed. Individual actions are often assembled in service of distal goals, and when this occurs, a simple action, like grasping a spoon, can be viewed as directed at a distal goal, such as stirring a pot of soup or feeding a baby. To analyze these downstream goals, the perceiver must shift focus from the proximal relations between agents and the objects they touch, to the distal relations between agents and their downstream goals. Recent findings have shown that infants engage in this kind of action analysis by 12 months of age. In one experiment, Sommerville and Woodward (2005) habituated 12-months-old infants to events like the ones depicted in the top panels of **Figure 1**. A woman grasped a cloth and pulled it toward her, thereby drawing near a toy that sat at its far edge. She then grasped the toy. The question of interest was whether infants viewed the woman's actions on the cloth as directed at the cloth or at the toy. To address this question, Sommerville and Woodward (2005) showed infants test events in which the toys' locations were reversed (see the lower panels of **Figure 1**) and the woman either

FIGURE 1 | Habituation paradigm used in Sommerville and Woodward (2005) **and current experiments.**
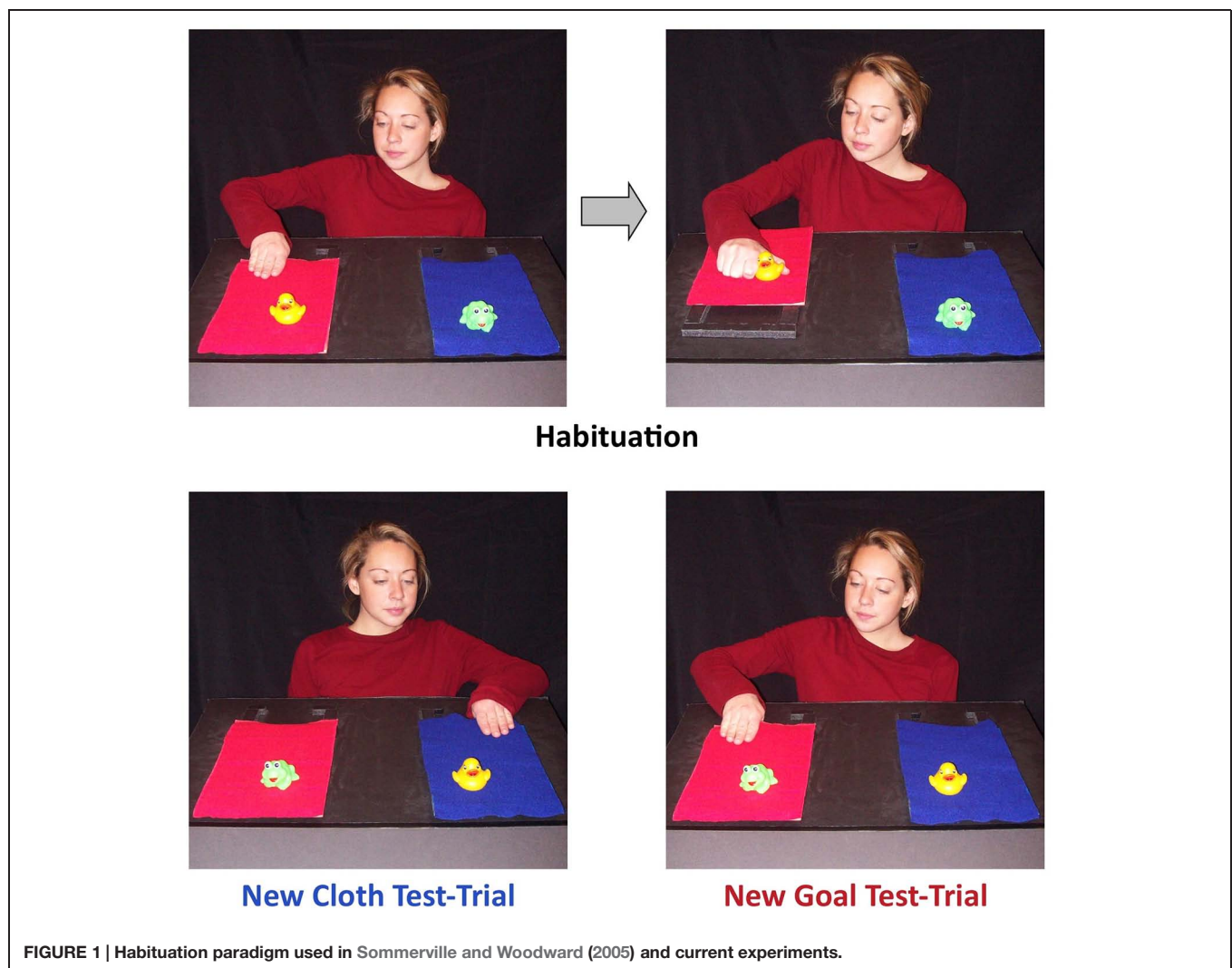
reached for the cloth on which she had previously acted which now held a new toy (*new-toy trials*), or the other cloth, which now held the toy she had previously attained (*new-cloth trials*). Twelve-months-old infants looked longer at new-toy trials than new-cloth trials, indicating that they interpreted the woman's actions on the cloth as directed at the toy; younger infants, 10-months-olds, did not respond systematically in this procedure (see Woodward and Sommerville, 2000; Biro et al., 2011 for related findings).

Critically, 12-months-old infants in this experiment (Sommerville and Woodward, 2005) used the causal structure of the event to interpret its means-end structure. A control group of infants was shown events that mimicked the surface structure of the events depicted in **Figure 1**, but which differed in causal structure because the toy sat next to the cloth rather than on it. In this condition, infants saw the experimenter grasp the cloth, pull the cloth and then grasp the toy, just as in the experimental condition. The act of pulling the cloth reliably preceded and was associated with grasping the toy, but nevertheless, infants in this condition did not interpret the cloth-grasp as directed at the toy (see Woodward and Sommerville, 2000 and Henderson and Woodward, 2011 for similar findings). That is, infants analyzed the same action, grasping the cloth, differently depending on whether it was causally related to attaining a distal goal. Thus, by 12 months, but possibly not before this time, infants are able to look beyond the proximal connections between agents and objects to discern distal goals.

Recent findings indicate that infants' sensitivity to the goal structure in others' actions is correlated with and affected by their own motor experience. These effects have principally been documented in studies of infants' production and perception of simple goal-directed actions, like reaching for a toy (Sommerville et al., 2005; Kanakogi and Itakura, 2010; Libertus and Needham, 2010; Daum et al., 2011; Loucks and Sommerville, 2012; Gerson and Woodward, 2014a,b). For example, Sommerville et al. (2005) found that 3-months-old infants who were trained to use Velcro-covered mittens to apprehend toys subsequently responded systematically to the goal structure of another person's reaching actions, but infants who did not undergo training did not (see also Libertus and Needham, 2010; Rakison and Krogh, 2011; Gerson and Woodward, 2014a). These behavioral findings are also consistent with recent neural evidence of shared representations between action production and perception in the brain (Rizzolatti and Craighero, 2004; Gerson et al., 2014).

In the case of simple actions, like grasping, motor experience may yield relatively concrete evidence about the way in which a particular action is organized with respect to goals. But understanding downstream goals requires a more flexible analysis of particular actions as potentially directed at distal goals rather than their proximal targets. Research regarding the role of experience in the understanding of means-end actions reflects this challenge. Sommerville and Woodward (2005) reported that, at 10 months, infants' skill at solving cloth-pulling problems correlated with their behavior in the above-described habituation paradigm: higher skill levels were associated with greater attention to the relation between the actor and the distal goal of the observed action, whereas lower levels of skill were associated with greater attention to the relation between the actor and the means. To gain clearer evidence as to the causal relations at play, Sommerville et al. (2008) conducted an intervention study in which 10-months-old infants were trained to use a cane as a means to obtain an out of reach toy. They were then tested in a habituation paradigm analogous to the one depicted in **Figure 1**. After being trained to use the cane, infants responded systematically to the means-end goal structure in the habituation events, looking longer on new-goal trials than on new-cane trials. In contrast, infants in control conditions who received no training or only observational exposure to cane events responded unsystematically on new-goal and new-cloth trials. Moreover, the effect in the active training condition was strongest for infants who had benefitted the most from training in their own actions. That is, infants who were better at performing the cane-pulling action at the end of training looked longer to new-goal (rather than new-cane) events in the habituation paradigm test-trials. These findings indicate that success on a means-end task engenders greater sensitivity to distal goals in others' actions. However, infants who were less successful in their own means-end actions responded randomly in the habituation task, rather than showing heightened attention to the means. Thus, it is not clear from these findings how infants perceive others' means-end actions during the initial stages of means-end learning.

A closer look at how infants develop the ability to produce means-end actions could shed light on this early stage of learning. Infants begin to engage in well-organized means-end actions by the end of the first year. For example, Willatts (1999), following on Piaget (1954) classic studies, reported that 8-months-old infants who were presented with cloth-pulling problems like the ones in **Figure 1** would *sometimes* produce clearly intentional solutions to the problem, visually fixating the toy while systematically drawing it within reach with the cloth (see also Bates et al., 1980; Chen et al., 1997; Munakata et al., 2002; Gerson and Woodward, 2012). Early in the acquisition of a means-end action, such as tool use, infants initially focus attention on the tool or means, rather than the distal goal (Willatts, 1999; Lockman, 2000; Keen, 2011). Learning to engage in efficient means-end actions requires exploratory behavior on the tool and the object retrieved by the tool and manipulating the relations between these two. Willatts (1999) described this as a transitional progression in means-end learning: infants first focus on the means (i.e., tool) while they are learning to perform the action and only later shift their focus to the object being acted upon by the tool. Given these patterns in infants' motor development, we might expect parallel effects on action perception. That is, when infants are at the early stages of means-end learning, their own attention to the means may lead them to focus on the relation between the actor and the means when viewing others' means-end actions. As they begin to produce well-organized means-end actions, infants may shift their attention from the relation between the actor and the tool to the relation between the actor and the distal goal both for others' actions and their own actions. In the current research, we examine whether this shift in focus from proximal elements of means-end problems to the distal goals of these actions seen in

motor learning is paralleled in infants' developing understanding of others' means-end actions.

In two experiments, we investigate the specific effects of different levels of motor and observational experience on infants' analysis of others' means-end goals. In Experiment 1, we measure infants' success in motor training and relate this to their action perception in a habituation paradigm in order to (1) further test the hypothesis that learning to engage in well-structured means-end actions leads to heightened attention to the relation between an actor and her distal goal, and (2) evaluate whether less successful training leads to heightened attention to the relation between an actor and the means on which she acts (i.e., the tool she first contacts). To this end, we implemented the approach developed by Sommerville et al. (2008) in their training condition, but we used a simpler means-end task (cloth-pulling) and tested younger infants (8-months-olds) in the hopes of finding greater variation in infants' success following the training. In Experiment 2, we evaluated infants' response to habituation events without training or with observational training as a point of comparison for the effects seen in Experiment 1.

Experiment One

Participants

Forty-eight 8-months-old infants (*M* age = 7.87 months; age range: 7.5–8.4 months) were included in this experiment. Infants had been born at full term (at least 37 weeks gestation) and resided in the Washington, DC, metropolitan area. All parents signed a written informed consent sheet that was approved by the University's Internal Review Board for this research and were told their participation in the research was voluntary. Parents identified their infants' racial group membership as follows: 46% Caucasian, 23% African American, 15% Hispanic, 8% multiracial, 2% Asian, and 6% unreported. Thirty additional infants began the procedure but were not included in the final sample due to experimenter error (*n* = 10), failure to complete the procedure due to distress (*n* = 11), failure to engage in activity during training

(*n* = 3), parental interference (*n* = 3), technical errors (*n* = 2), or because they had total looking times more than 3 SDs above the sample mean (*n* = 1). The attrition rate in this study is on par with similar paradigms used with infants (e.g., Király et al., 2003; Hofer et al., 2005; Biro and Leslie, 2006; Southgate et al., 2009).

Procedure

During training, infants sat, on a parent's lap, at a table adjusted to a height that allowed them to readily reach for and manipulate objects on its top (see **Figure 2**). Parents were asked to hold the infant securely but not to talk to the infant or influence the infants' actions in any way. An experimenter sat next to the infant and a camera facing the infant recorded the session for later coding. Following training, infants underwent a visual habituation paradigm designed to assess their understanding of another person's means-end action goals.

Active Training

The left panel in **Figure 2** depicts the events in the active training portion of the experiment. In this portion, infants were given the opportunity to act on a series of problems in which a toy was placed out of reach on the far side of a cloth that extended to within the infant's reach. First, the infant received four *pre-training trials*. On these trials, the infant was given the chance to act on the cloth-pulling problems but was given no guidance for doing so. The experimenter set up the problem in front of the infant and then looked down at the table. She drew the infant's attention to the cloth if necessary but did not provide more specific cues to prompt the infants' actions. The trial ended when the infant had obtained the toy or when 30 s had elapsed. Across successive pre-training trials, infants were presented with two cloths and two toys that matched the ones they would later see in the habituation paradigm. Each cloth was presented with each of the two toys on separate trials so that each infant was presented with all possible cloth-toy combinations. The order of presentation of each toy-cloth pairing was randomized. After pre-training, the infant received five *training trials*. On these trials, the experimenter set the cloth and toy in front of herself and then enacted

Training Demonstration
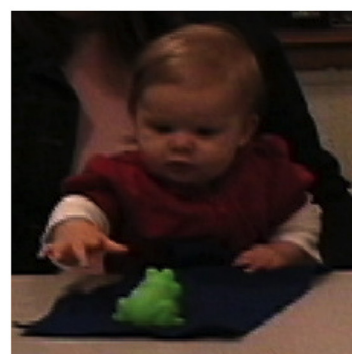
Infants' Pull
(Pre and Post)

FIGURE 2 | Training session demonstration and action.

a means-end solution: She pulled the cloth, watching the toy as it drew near, and then retrieved the toy, inspecting it and expressing interest by saying "Ooo" as she did so. The experimenter repeated these actions twice, and then set up the same problem in front of the infant, giving the infant a chance to respond without further prompting, as on pre-training trials. Each of the five training trials involved a unique cloth-toy combination that differed from the items used during pre-training. Finally, the infant received four post-training trials, which were identical to pre-training trials. Throughout training, infants received no hands-on guidance from the experimenter or parent. All successfully completed sequences were performed by the infant him or herself.

Coding of Infants' Actions

The training session was coded for the extent to which infants engaged in well-structured solutions. Infants' actions were scored as planful if the infant maintained visual contact with the toy while pulling the cloth in one continuous movement and then retrieved or touched the toy within 3 s of the completion of the pull (see Willatts, 1999; Sommerville and Woodward, 2005). Two independent coders, each unaware of the infants' responses in the visual habituation portion of the experiment, coded each infant's actions. The two coders agreed on infants' planfulness on 88% of trials (cohen's κ = 0.76). Additional frame-by-frame coding of attention to the experimenter's actions during training trials was assessed using a digital video coding program (Mangold, 1998). Coders measured the length of time infants attended to each aspect of the event (cloth, toy, or experimenter) during each portion of the pulling action (prior to touching the cloth, during the pull of the cloth, and during the grasp of the toy; reliability on duration of looking between two coders: rs > 0.95).

Habituation and Test

After the training procedure, infants were brought to a second testing room, equipped for the visual habituation procedure. Infants sat on a parent's lap facing a small stage 72 cm away. On the stage sat two cloths, side by side, on a table-top surface that sloped slightly down toward the infant (so as to be easily visible but not to cause objects to slide down the slope; see **Figure 1**). Each cloth supported a different toy (a frog or a duck). A presenting experimenter (henceforth, the presenter) sat behind the stage, facing the infant. A screen was raised to hide the stage from view between trials. Parents were instructed not to talk and to look down at the infant rather than at the experimental events. A camera mounted below the stage filmed infants as they watched the events. An observer in another room watched the infant on a video monitor and coded the infant's attention using a program that calculated looking times and habituation criteria (Casstevens, 2007). The observer could not see the experimental events and was not informed of the condition to which the infant had been assigned or the order of test trials.

At the start of each trial, the screen was lowered to reveal the stage and the presenter drew the infant's attention by saying "Hi" and making eye contact. During habituation trials, the presenter proceeded to look down toward one of the toys, pulled the cloth toward herself and then reached toward and grasped the toy that

had been drawn near. She remained still in this position, looking at the toy, until the trial ended. Infants' attention to the event was calculated beginning as soon as the presenter had stopped moving and the trial continued until the infant had looked away for 2 consecutive seconds. When the trial ended, the screen was raised, the cloth was returned to its original position, and then the screen was lowered for the presentation of the next habituation trial. Across habituation trials, the actor consistently reached for the same cloth and toy on the same side of the table. Habituation trials were continued until the infant's attention, summed over three consecutive trials, had declined to 50% of its initial level or for 14 trials.

Following habituation, the screen was raised and the positions of the toys on the cloths were reversed. Then the screen was lowered to allow infants to view the toys in their new positions for an infant-controlled familiarization trial. During this familiarization trial, the presenter looked down and did not look toward the stimuli. After this, the test trials were presented. On test trials, after saying "Hi" the presenter turned to grasp the near edge of one of the two cloths and look toward the toy at the end of the cloth. She then held still in this position for the duration of the trial, which was infant-controlled, as during habituation. It is important to note that, unlike in the habituation trials, during test trials the presenter never moved the cloth or touched the toy (matching the procedure used in Sommerville and Woodward, 2005). On *new-goal trials*, she grasped the same cloth that she had acted on during habituation, which now supported a new toy at its far end. On *new-cloth trials* she grasped the cloth she had not acted on during habituation, which held her prior goal toy at its far end. Three new-goal and new-cloth trials were presented in alternation. The type of test trial seen first, the side to which the presenter reached during habituation, and the toy that was the presenter's goal during habituation were counterbalanced across infants.

Each infant's video session was coded after the fact by a second independent observer. The online and reliability observers were counted as agreeing if they agreed on the point at which the infant looked away to end the trial. The two observers agreed on the endpoints of 95% of test trials. To evaluate potential observer bias, all disagreements were categorized as those that would indicate bias in favor of the hypothesis on the part of the on-line coder versus those that would indicate bias against the hypothesis. The observers' disagreements were randomly distributed (Fisher's Exact Test, *ns*).

Results
Training Session: Assessment of Quality of Motor Training

Coding of infants' attention to the experimenter's actions during training trials indicated that infants attended to the relevant aspect of the action during the majority of the experimenter's actions throughout training trials. That is, they attended to the cloth during the pulling action (90% of the time on average) and to the toy and experimenter during the grasping action (83% of the time on average).

On the 13 training trials (including pre- and post-training trials), infants produced planful actions on an average of 6.40

(SEM = 0.51) trials overall. As shown in **Figure 3**, infants increased their planfulness from pre- to post-training trials. A repeated measures analysis of variance (ANOVA) on the proportion of planful actions in the pre-training, training, and post-training trials revealed a significant increase in planful actions across these phases, $F(2,45) = 18.13$, $p < 0.001$, $\eta_p^2 = 0.45$. Pairwise comparisons of the estimated marginal means indicated that infants' planfulness increased significantly from pre-training to training (mean difference = 0.23, SEM = 0.048; $p < 0.001$) and from training to post-training (mean difference = 0.14, SEM = .05; $p = 0.013$). Age did not correlate reliably with infants' degree of planfulness in any of the three phases or with infants' degree of improvement from pre-training to post-training (all rs < 0.12, ps > 0.42). Thus, the active training procedure reliably increased the extent to which infants engaged in well-organized means-end actions. Even so, infants' responses to training varied. Only half of the infants were highly planful after training—25 of the 48 infants achieved planful scores on 3 or 4 out of the 4 post-training trials. Thus, a median split of planfulness (Quality of Training) corresponded with a theoretically meaningful cutoff.

Planful and unplanful infants did not differ from one another in age ($p = 0.89$). We further assessed whether planful and unplanful infants differed in their general attention to assure that unplanful infants were not simply less alert in general. Infants in the two groups did not differ in amount of attention (i.e., the length of looking to each trial) at the beginning ($p = 0.33$) or end of habituation trials ($p = 0.98$). Further, the number of habituation trials needed to reach habituation criterion (often thought of as a measure of speed of processing and known to be related to later intelligence; Fagan, 1992) did not differ between the two groups ($p = 0.38$). Planful and unplanful infants did not differ in overall amount of attention during test trials (collapsed across two different kinds of test-trials) during test-trials ($p = 0.39$). Finally, we also assessed infants' attention during the training session in order to assure that infants had the same opportunity to learn from training trials. Infants in the two groups spent

comparable amounts of time attending to the relevant aspects of the experimenter's actions during training trials during both the pulling ($p = 0.24$) and grasping ($p = 0.33$) portions of training trials. Across groups, no significant correlation was found between attention to relevant aspects of training and post-training planfulness (rs < 0.23, ps > 0.13). Thus, there was no evidence that variations in infants' attentiveness during the procedure, or in their age, accounted for their ability to benefit from training (see **Table 1** for a summary of means and SDs). Subsequent analyses took the variation in the extent to which infants benefited from training into account, as described below.

Habituation Session: Relative Attention to Cloth and Goal Relations

Preliminary analyses assessed infants' attention during the habituation trials. A repeated measures ANOVA with habituation trial (the first three and last three trials for each infant) as the repeated measure revealed a main effect of trial, $F(1,47) = 65.11$, $p < 0.001$, $\eta_p^2 = 0.58$, reflecting a decline in attention across trials. Infants required ~9 trials on average to reach habituation criteria.[1]

The focal analysis concerned infants' differential attention to the change in relation between the agent and the means she used (i.e., new-cloth test events) or her distal goal (i.e., new-goal test events) and whether differential responses to these test events varied as a function of the success of training. A repeated-measures ANOVA was conducted with average looking time to the new-goal and new-cloth events as the repeated measure (Type). In order to take into account the variability in training success, a median split of infants' planfulness at the end of training (Training Success) was included as a between-subjects factor. As discussed above, approximately half of the infants were successful in planfully carrying out the means-end action in at least three of the four post-training trials and these two groups did not differ in age, attention during habituation, or attention during training trials. This analysis revealed a significant Type X Training Success interaction, $F(1,46) = 14.50$, $p < 0.001$, $\eta_p^2 = 0.24$. The main effects of Type and Training Success were not significant (Fs < 0.75).[2] Pairwise comparisons of the estimated marginal means (see **Figure 4** for raw means and standard errors) revealed that infants below the median in planfulness at post-training looked significantly longer to new-cloth than to new-goal trials (mean difference = 2.66, SEM = 0.96; $p = 0.008$) whereas infants above the median in planfulness looked significantly longer to new-goal than new-cloth trials (mean difference = 2.43, SEM = 0.92; $p = 0.012$).

Relations Between Training Success and Action Perception

Given the differences found based on the success of training as reflected in the median split of post-training activity

FIGURE 3 | Infants in the active condition increased in planfulness from pre-training to post-training (box plot median, quartiles, minimum, and maximum proportion of trials within each portion of training that infants were planful) *p < 0.02.

---
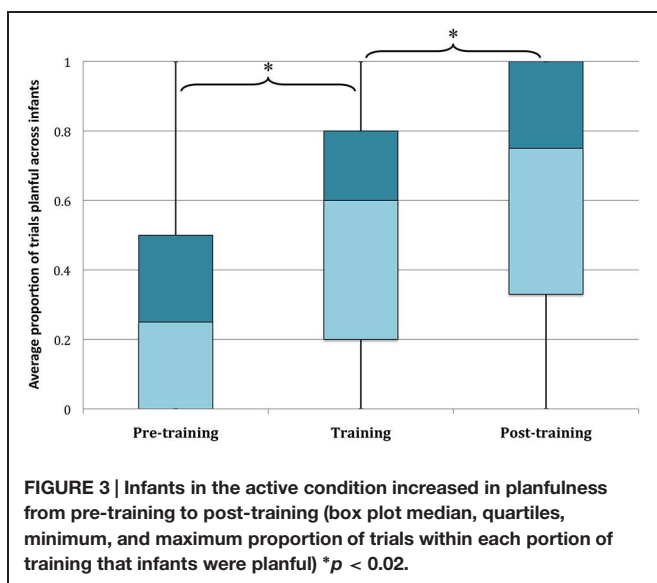
[1]Four infants reached 14 habituation trials without meeting the habituation criterion. When these infants were removed from the sample, the principle findings were unchanged. Therefore the analyses are reported for the full sample.

[2]We conducted this critical analysis with a randomly selected subset of 24 subjects (in order to match the sample size of infants in the observational and control conditions of Experiment 2) and saw a nearly identical pattern: Type X Training success: $F(1,22) = 10.20$, $p = 0.004$; no significant main effects (Fs < 2.70).

TABLE 1 | Similarity in infants' attentional patterns across active groups.

	Age	Attn to beg of Hab	Attn to end of Hab	# of Hab trials	Total attn to test trials	Pre-training planfulness	Attn to pull (in training)	Attn to grasp (in training)
Below median planful M(SEM)	7.86 mos (0.06)	46.69 s (6.83)	14.04 s (1.87)	8.35 (0.58)	40.66 s (4.91)	0.23 (0.06)	2.48 s (0.09)	2.33 s (0.15)
Above median planful M(SEM)	7.85 mos (0.05)	39.08 s (3.54)	13.97 s (1.39)	9.08 (0.59)	35.39 s (3.74)	0.33 (0.06)	2.63 s (0.07)	2.53 s (0.14)
t-test p-value	0.89	0.33	0.98	0.38	0.39	0.24	0.33	0.36
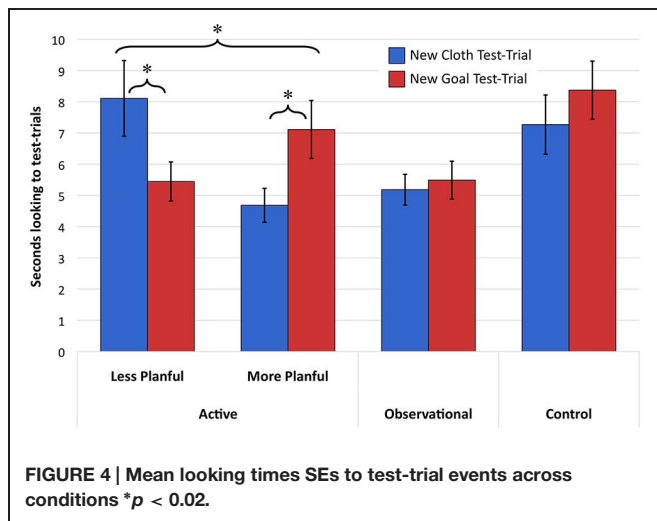
FIGURE 4 | Mean looking times SEs to test-trial events across conditions *$p < 0.02$.

TABLE 2 | Hierarchical multiple regression: effect of post-training planfulness on new-goal preference.

Variables	Model 1	Model 2
Intercept	0.51**	0.42**
Pre-training planfulness	0.02	−0.01
Post-training planfulness		0.15*
Adjusted R^2	−0.022	0.054
Model F	0.021	2.31
Change R^2	<0.001	0.95*
Incremental F		4.6*

*$p = 0.037$, **$p < 0.001$.

(training success), we further explored the relation between planfulness in different phases of training and looking time differences in the habituation paradigm. Infants' planfulness in post-training was unrelated to their planfulness in pre-training ($r = 0.15$, $p = 0.32$), suggesting that individual differences in post-training planfulness were not a function of motor abilities prior to training. We also examined whether attention to different aspects of the experimenter's actions during training trials related to infants' new-goal preference in the habituation phase, but no aspect of attention was significantly related ($rs < 0.24$, $ps > 0.11$).

To examine the unique contribution of pre versus post-training on infants' differential looking to new- versus old-goal test events, we performed a hierarchical multiple regression analysis. For each infant, we calculated a difference score reflecting his or her relative visual preference for new-goal trials compared to new-cloth trials (average looking time on new-goal trials minus average looking time on new-cloth trials; see Sommerville et al., 2005 for a similar measure) and entered this as the dependent variable. Pre- and post-training planfulness were entered in two steps. In step 1, pre-training planfulness was the independent variable. In step 2, post-training planfulness was added to the step 1 equation.

The results of step 1 indicated that the variance accounted for by pre-training planfulness was not significantly different from zero (see **Table 2**). Adding post-training planfulness as a predictor significantly improved the model; post-training planfulness was positively related to new-goal preference ($B = 0.15$,

whereas pre-training planfulness was unrelated ($B = -0.01$). Thus, infants' learning from the training session (as evidenced in post-training planfulness), rather than their starting means-end abilities, predicted their subsequent responses to the observed actions in the habituation paradigm.

Discussion

The findings of Experiment 1 converge with those of Sommerville et al. (2008) in showing that training in a means-end action supports infants' sensitivity to others' means-end actions. Infants who benefited from means-end training, in that they became able to organize their own actions in service of a distal goal, subsequently responded to the higher-order goal structure of another person's means-end actions. Extending beyond the findings of Sommerville et al. (2008), the current findings also indicate that infants who were less successful in organizing their means-end actions attended, instead, to the relation between the observed agent and the means she acted on initially (the cloth). Importantly, there was no evidence that infants who did less well in organizing their own actions were less alert, attentive, or engaged than infants who were more successful. Further, both planful and unplanful infants responded systematically on test trials, showing that they had encoded and remembered the habituation events. The differences in infants' responses to the test events were not predicted by infants' ability to perform the task prior to training, but rather were predicted by infants' skill following training. Thus, the differential findings based on planfulness seemed to reflect infants' experiences during the training phase, rather than differential abilities they brought with them into the laboratory.

The coding scheme by which infants' actions were identified as planful gives some clues as to the reason infants who did

not benefit showed a bias toward interpreting the means-end action as directed at the simple one-action-step goal (i.e., the cloth). Action task trials in which the infant looked at the cloth during a pull rather than the toy or in which the child pulled the cloth but then failed to retrieve the toy within a short amount of time were qualified as unplanful (and the child often did not attain the toy at the end of the trial). Those infants who produced more of these actions during training may have spent more time attending to the cloth than other infants in that they concentrated their attention on the cloth in service of attempting to successfully coordinate their actions on the cloth. These patterns are consistent with developmental patterns in infant motor development: When initially learning new actions, infants seem first to attend to the means of the action and as they gain proficiency, they shift attention to the goal (Willatts, 1999).

Our findings raise the question of why infants varied in their learning from the training manipulation. This variation was not accounted for by age. It is likely instead that infants' ability to benefit from motor training depended on their existing motor abilities and developmental readiness for learning (see Piaget, 1954; Lockman, 2000; Keen, 2011). The current findings do not provide evidence for fully evaluating this issue. Further research is needed to investigate the developmental predictors of motor learning and their relation to generalizing motor information to the perception of others' actions.

What is it that self-produced experience provided for infants in this experiment? One interesting possibility is they learned about the goal of the action through simply observing their own actions. We know that infants this age can learn about the goals of tool-use actions without active experience organizing their own actions on tools in certain circumstances (see, for example, Gerson and Woodward, 2012, 2014c). On the other hand, there is reason to think that the act of producing an action, rather than simply observing it, may be particularly informative because it yields shared action perception representations.

Recent research has investigated the unique effects of active, relative to observational, experience on action perception (Sommerville et al., 2008; Gerson and Woodward, 2014a; Gerson et al., 2014), face perception (Libertus and Needham, 2011, 2014), and spatial perception (Frick and Wang, 2014). If the mirror system plays a role in the link between motor experience and action perception (see Hunnius and Bekkering, 2014; Woodward and Gerson, 2014, for discussion), benefits should (at least initially) be unique to active experience as this widens the motor repertoire of the infant.

Accordingly, recent studies have examined the effects of observational experience with novel actions in training studies with infants matching those described above. Gerson and Woodward (2014a) found that active, but not observational, training with reaching actions led 3-months-old infants to recognize the goal of a grasping action in a habituation paradigm. Similarly, Sommerville et al. (2008) included an observational training condition in their means-end training experiment, in which 10-months-old infants watched an experimenter produce planful actions using the same canes and for the same number of trials as infants in the active training condition. Infants in this condition did not respond systematically in the habituation paradigm. Thus, these findings indicate that self-produced actions provide stronger support for viewing others' actions as goal-directed than do observed actions. In the Sommerville et al. (2008) experiment, all infants in the observational condition received a set amount of experience observing means-end actions (the mean of that produced by actively trained infants), so effects of variability in the amount of observational experience received could not be assessed (but see Sommerville et al., 2011).

In Experiment 2, we matched the amount of experience watching means-end actions to that of a yoked infant in the active condition from Experiment 1. Given the differential effects of training found in Experiment 1, we aimed to examine whether different amounts of observational experience would differentially influence action perception. That is, would the amount of observational experience received influence infants' relative attention to the relation between the actor and the means or goal of her actions, as it did in Experiment 1? This matched variability allows us to assess potential correlations between observed activity and action perception. We also included a control condition in which infants had the chance to explore each cloth and toy prior to the habituation paradigm but never got to act on one in relation to the other and did not observe this action prior to the habituation paradigm. This allowed us to compare observational experience with infants' action perception when they received no means-end training.

Experiment Two

Participants

Forty-eight 8-months-old infants (M age = 7.9 months; age range: 7.27–8.43 months) participated in one of two conditions in this experiment: observational or control. Infants had been born at full term (at least 37 weeks gestation) and resided in the Washington, DC, metropolitan area. Parents identified their infants' racial group membership as follows: 60% Caucasian, 3% Asian, 17% African American, 10% Hispanic, and 10% multiracial. Twenty-nine additional infants began the procedure but were not included in the final sample due to experimenter error (n = 13), failure to complete the procedure due to distress (n = 12), parental interference (n = 1), or because they had total looking times more than 3 SDs above the sample mean (n = 3).

Procedure

Infants underwent a "training" period prior to participation in the habituation paradigm. During this session, as in Experiment One, infants sat on a parent's lap at a table adjusted to a height that allowed them to readily reach for and manipulate objects on its top. Parents were asked to hold the infant securely but not to talk to the infant or influence the infants' actions in any way. An experimenter sat next to the infant and a camera facing the infant recorded the session for later coding.

Observational Training

Infants in the observational condition were shown the same series of cloth-pulling problems as infants in the active condition from Experiment One, but they observed the experimenter solving each problem and were not given the opportunity to act on the toys themselves. To equate, as much as possible, the duration of training in this condition to the amount of experience received in the active condition from Experiment One, the duration of each trial for infants in the active condition was coded, and the session from each infant in the active condition was used to generate a script for an infant in the observational condition that specified the duration of each observation trial. Because the experimenter's actions were generally more well-organized than those of the infants, this meant that the experimenter sometimes repeated the problem several times in order to keep the infant engaged for the full trial duration. Thus, infants in the observational condition had equivalent durations of exposure to the problems as did infants in the active condition from Experiment One, and they viewed more instances of well-organized solutions than did infants in the active condition (see below for details).

Control "Training"

Infants in the control condition were given the opportunity to explore each cloth and each toy that were involved in the active and observational training, but they saw each cloth and each toy presented independently (i.e., sequentially), rather than in the context of a means-end problem. The order of presentation paralleled the order in the active and observational conditions, with infants first being given each of the four items involved in the pre- and post-training phase for 15 s each, then each of the 10 items from the training phase for 30 s each, and then the four pre- and post-training items again for 15 s each.

Coding of Training Session

Videos of the observational condition were coded for infants' attention during each phase of the experimenter's movements–grasping the cloth, pulling the cloth, and retrieving the toy—to identify the number of complete means-end actions that each infant viewed. To assess reliability, a second independent coder coded the sessions for 25% of infants. The two coder's judgments of the number of planful actions infants observed in each phase of the training session were highly correlated, $r = 0.99$. As in the training trials from Experiment One, additional frame-by-frame coding of attention to the experimenter's actions during observational training was assessed using a digital video coding program (Mangold, 1998; reliability: rs > 0.95).

Habituation and Test

The habituation procedure in this experiment was identical to that of Experiment One. Reliability of the online coders was assessed and coders agreed on the end of the trials for 94% of test trials (cohen's $\kappa = 0.88$). To evaluate potential observer bias, all disagreements were categorized as those that would indicate bias in favor of the hypothesis on the part of the on-line coder versus those that would indicate bias against the hypothesis. The observers' disagreements were randomly distributed (Fisher's Exact Test, ns).

Results

Training Session: Assessment of Amount of Observational Training

In the observational condition, we examined the number of planful pulls observed by each infant. Because the experimenter repeated planful pulls for the duration of each trial in the observational condition, infants in this condition had the opportunity to view more planful actions than infants in Experiment One produced (mean number of planful pulls in Experiment One was 6.4, ranging from 0 to 12). Coding of infants' attention to the pulls revealed that infants in the observational condition viewed 24 planful pulls on average (range = 16–29). Further, frame-by-frame coding of infants' attention to the experimenter's actions indicated that they attended to the relevant aspect of the action the majority of the time: to the cloth during pulling actions (88% of the time) and to the toy and experimenter during the grasping action (77% of the time). Infants in the observational condition did not differ from infants in the active condition from Experiment One in their attention to any of these aspects (ps > 0.10).

Habituation Session: Relative Attention to Cloth and Goal Relations

Preliminary analyses assessed infants' attention during the habituation trials. A repeated-measures ANOVA with the first three and last three trials of habituation as repeated measures and condition (observational versus control) as a between subjects factor revealed a main effect indicating a significant decrease in attention across conditions, $F(1,46) = 97.04$, $p < 0.001$, $\eta_p^2 = 0.68$, no interaction between condition and trials ($p > 0.57$), and no main effect of condition ($p > 0.49$). When the active condition from Experiment One was included in this analysis there was again no interaction between condition and trial. Infants in Experiment Two habituated in approximately eight trials on average.

The main analysis concerned whether infants in either the observational or control condition showed preferential looking to the new-goal or new-cloth test-trials and whether they differed from each other and/or infants in the active condition from Experiment 1 who were more or less planful at the end of training. We first examined only the infants in the control and observational conditions (see **Figure 4**). A repeated-measures ANOVA with test-trial type as the repeated measure (new-goal or new-cloth) and condition as the between subjects factor (observational or control) revealed no main effect of Type [$F(1,46) = 1.58$, $p = 0.22$, $\eta_p^2 = 0.03$] and no interaction between Condition and Type [$F(1,46) = 0.51$, $p = 0.48$, $\eta_p^2 = 0.01$]. A main effect of condition [$F(1,46) = 7.10$, $p = 0.01$, $\eta_p^2 = 0.13$] indicated that infants in the control condition looked longer across test trials than did infants in the observational condition.

Relations Between Amount of Training and Action Perception

As a measure of experience in the observational condition, we also examined whether the number of planful pulls observed differentially influenced looking times to different test trials. As a

measure of the possible continuous relation between experience observing planful pulls and new-goal preference (as found in Experiment 1), we calculated the difference between average new-goal and new-cloth (*average difference*) trials and examined its relation to number of pulls observed. No significant relation was found ($r = 0.085$, $p = 0.69$). As in Experiment One, we also examined whether any aspect of attention to the experimenter's actions during training trials related to new-goal preference in the habituation paradigm. No significant relations were found ($rs < 0.28$, $ps > 0.20$).

Because the number of pulls presented (and thus the maximum possible number to observe) was randomly assigned to infants based on scripts from activity of infants in Experiment One, we also took into account individual differences created by the infants themselves by dividing the number of trials observed by the number of trials presented. On average, infants observed 89% of the actions produced by the experimenter (range: 63–100%). The relation between proportion of pulls observed and new-goal preference was not significant ($r = -0.25$, $p = 0.24$).

As in Experiment 1, we created a median split of experience (Amount of Training: more or fewer than 15 planful pulls observed). In order to compare the effects of experience in the observational condition with infants from the active condition in Experiment One directly, we conducted a repeated-measures ANOVA with test-trial type as the repeated measure and condition (active or observational) and Amount of Training as between subjects factors. This revealed no main effects of Type ($F < 0.05$, $p > 0.85$), Condition ($F < 1.4$, $p > 0.24$), or Amount of Training ($F < 1.3$, $p > 0.27$), and no interaction between Type and Condition [$F(1,68) = 0.16$, $p = 0.69$, $\eta_p^2 = 0.002$]. A significant interaction between Type and Amount of Training [$F(1,68) = 5.74$, $p = 0.019$, $\eta_p^2 = 0.078$] was qualified by a three-way interaction between Type, Condition, and Amount of Training [$F(1,68) = 5.67$, $p = 0.02$, $\eta_p^2 = 0.077$]. Comparisons of estimated marginal means again revealed that the three-way interaction was a function of the significant effect of Type that was in opposite directions for infants above and below the median in active experience in Experiment One (as described above; $ps \leq 0.005$) but no significant differences between Type in either infants above (estimated marginal means for new-cloth trials, 5.69. SEM = 1.24, and new-goal trials, 6.00. SEM = 1.16) or below the median in experience (estimated marginal means for new-cloth trials, 4.82. SEM = 1.05, and new-goal trials, 5.12. SEM = 0.98) in the observational condition (mean differences < 0.32, $ps > 0.79$).

Given the lack of effects of experience in the observational condition and the lack of difference between observational and control conditions, we then collapsed across these conditions to examine whether responses in the habituation paradigm differed between these conditions and the more and less planful infants from the active condition in Experiment One. We conducted a univariate ANOVA with proportion of attention to the new-goal test trials (new-goal)/(new-goal + new-cloth) as the dependent variable and condition group (active-high, active-low, or observational and control) as the between-subjects factor. The condition groups significantly differed from one

another in new-goal preference [$F(1,2) = 5.72$, $p = 0.005$, $\eta_p^2 = 0.11$], and we followed up with planned comparisons between conditions. Pairwise comparisons of the estimated marginal means indicated that infants in the observational and control conditions had significantly higher new-goal preferences than unplanful infants in the active condition (mean difference = 0.08, SEM = 0.04, $p = 0.03$) and had marginally lower new-goal preferences than planful infants in the active condition (mean difference = 0.06, SEM = 0.04, $p = 0.095$; see **Figure 4**).

Discussion

The results of Experiment 2 did not reveal effects of observational experience with means-end on infants' action perception. This result is consistent with findings indicating that observation of means-end training is not as beneficial as active training at 10 months of age (Sommerville et al., 2008). In this experiment, we expanded on prior research to explore individual differences in the amount of observational experience received. Given the importance of the amount of planful actions produced in the active condition in Experiment 1, we allowed infants the opportunity to observe planful actions for the same range of time as infants in the active condition. This way of matching infants meant that infants in the observational condition experienced more instances of well-formed, planful actions than did infants in the active condition. Even so, infants in the observational condition did not demonstrate a benefit from training at a group level or show any of the same patterns of individual differences as infants in Experiment 1.

It is important to note that the difference in effects between conditions cannot be due to a difference in opportunities to observe or infant attentiveness. In creating yoked observational scripts, we erred on the side of allowing infants in the observational condition to view more demonstrations than their partners in the active condition. Infants in the observational condition always saw more planful actions during the training phase than their matched partner in the active condition from Experiment 1 (in fact, infants in this condition saw almost four times more exemplars of the cloth-pulling action than their active training counterparts). They attended to cloth-pulling actions for as long as infants in the active condition and thus received equal exposure to the toys and cloths. The physical causal information (pulling the cloth makes the toy move) was identical to infants across the active and observational conditions. Looking times during both the habituation and test phases of the looking time procedure did not indicate any differences in overall attention to the events between the two studies. Additionally, the patterns found in the active condition were a result of attention to a specific relation between particular actions, means, and objects, rather than general attention to the event or a particular toy or cloth.

Despite controlling for attention, there may be other differences in readiness to learn that we could not assess in our observational condition. The yoking procedure we used randomly assigned infants in the observational condition to a script duration based on an active infants' timing. It is possible that seeing a greater number of demonstrations at a faster rate than

infants in the active condition could have hindered some infants' ability to make sense of the viewed action, but we had no way to take into consideration the length of time that each individual infant in the observational condition may have needed to benefit from observation. Thus, it is possible that infants tested under conditions in which readiness to learn is taken into consideration may show greater benefits from observation.

Critically, however, the design of the current experiment parallels a real-world difference between active and observational learning. In active learning, infants' experience is self-generated and thus can be readily calibrated to their current learning state (e.g., infants can continue to act on the world until they have all the information that they need to learn relevant information). In contrast, when learning via observation, infants are at the behest of the caregiver, adult or more advanced peer who is doing the demonstrating. During observational learning, it is the demonstrator that decides how much information to give infants and for how long; given that demonstrators do not have direct access to infants' knowledge base or learning state (although infants may provide implicit cues to this state), information accrued via observation may be less well suited to an infant's learning state than is information accrued via active learning. Indeed, this distinction between active and observational learning may be one of the factors driving the potential benefits of active versus observational learning. Future work can directly assess this possibility.

The observational and control conditions provided a point of comparison for the active training groups, allowing us to examine whether both the low and high planful groups differed from how infants might respond to the habituation events spontaneously. The fact that both groups of active infants differed from the observational and control conditions in opposite directions suggests that initial, unsuccessful attempts at means-end problems push attention to the proximal agent-means relation whereas more successful training pushes attention to the distal agent-goal relation.

General Discussion

Prior findings have shown that active motor experience affects infants' sensitivity to the goal structure of others' simple actions (Sommerville et al., 2005; Gerson and Woodward, 2014a,b). Our question, in the current studies, was whether active motor experience also supports infants' emerging sensitivity to others' distal goals. Understanding distal goals requires that the perceiver look beyond the actor's immediate motor interactions in order to consider his or her potential distal goals, and this raises the question of whether and how concrete motor experiences could contribute to this aspect of goal analysis. Our findings provide evidence that active motor experience supports infants' analysis of distal goals, and further, provide new insight into the influence of infants' motor experiences on their analysis of others' actions.

In the current experiments, infants saw a chain of interrelated actions in the habituation trials of the looking time paradigm. The presenter first reached for and grasped a cloth. After pulling

on it, she then reached for and grasped the toy at the end of the cloth. Test trials assessed whether infants viewed the experimenter's actions on the cloth as directed at the cloth itself, or instead as directed at the toy. The findings of Experiment 1 indicated that infants' active experience in a cloth-pulling task predicted which of these interpretations they adopted. Infants who benefited from training and became highly organized in their own actions viewed the experimenter's action on the cloth as directed at the toy. Infants who were less successful in their training activities viewed her actions as directed at the cloth. Compared to infants in Experiment 2, who underwent observational training or no training, infants in Experiment 1 showed systematic differences in each response pattern. Thus, at a first level of analysis, the current findings contribute support to the conclusion that infants' interpretation of distal goals is influenced by their own motor experience (Sommerville and Woodward, 2005; Sommerville et al., 2008).

The current findings go beyond prior work in demonstrating that variation in infants' success in means-end activities leads to systematic variation in their analysis of others' actions. Infants who benefited from active training showed the higher-level interpretation of the events in the habituation paradigm, consistent with findings from older infants (Sommerville et al., 2008). But infants who engaged in ineffective means-end actions showed just the opposite response, interpreting the observed actions in terms of the proximal goal (the cloth) rather than the distal goal. These distinct patterns of response mirror the patterns that occur during developments in infants' own means-end actions (Willatts, 1999). This result suggests that the processes that give rise to means-end structure in infants' motor behavior also support the emergence of means-end structure in their analysis of others' goals.

We can conclude, then, that there is a specific relation between organizing means-end action toward the goal and understanding others' means-end actions as organized toward a goal. The individual differences found in Experiment 1 suggest that infants may at first concentrate and learn about the means of a multi-step action and then change their focus to the goal once they gain proficiency with a new action. Active experience seems to focus infants' attention on relevant relations and, depending on the nature of their own actions, this could be the relation between the cloth (i.e., proximal goal) and the agent or the goal (i.e., distal goal) and the agent.

Importantly, this shift in focus was not seen in Experiment 2, when infants observed an adult engage in repeated, well-structured means-end actions, nor was there any indication that variations in observational experience related to variations in infants' responses to the habituation events. Infants' failure to benefit from the observational training is striking. In the observational training, infants were witness to critical information about the goal-structure of the cloth-pulling action. They viewed the causal relation between acting on the cloth and attaining the toy, and they saw the experimenter express interest in the toy. Infants were highly attentive to these events, and yet seemed not to recover meaningful information from them regarding the goal structure of cloth-pulling events. This finding, in conjunction with previous research (Sommerville et al., 2008;

Gerson and Woodward, 2014a), suggests that active experience provides a particularly potent, and possibly unique, source of evidence for understanding others' actions during early development.

Even so, open questions remain concerning the nature of the benefit conferred by active experience. It is possible that self-produced actions yield information about goal structure that infants cannot glean from observation alone. Alternatively, it remains possible that infants can glean goal information from observational experience, but were unable to demonstrate it given the demands of the current task. The training and habituation sessions were conducted in different rooms and involved different people, and infants have difficulty carrying goal information across contexts (Sommerville and Crane, 2009). Thus, active experience may create particularly robust or "portable" representations, as compared to observational experience (see Gerson and Woodward, 2010 for further discussion).

The current findings indicate that infants' own actions render changes in their sensitivity to the goal structure of others' actions. Recent findings in infants (van Elk et al., 2008; Southgate et al., 2009; Saby et al., 2012; Gerson et al., 2014; Cannon et al., 2015) suggest that the motor system is active during, and may play a role in, infants' perception of others' actions. Although the current findings do not provide direct evidence concerning the neural mechanisms at work, they raise the question of whether shared neurocognitive representations support infants' analysis of higher-order goals. Mirror neurons in primates and mirror systems in humans are modulated, not only by the goals of simple actions, but also by overarching goals that structure action

sequences (Fogassi et al., 2005; Iacoboni et al., 2005). For example, Fogassi et al. (2005) found mirror neurons in macaque monkeys that fired differentially to grasping actions that preceded eating versus placing of the grasped object when there were contextual cues to support one of these two analyses of the grasp. In this way Rizzolatti and Craighero (2004) suggest that "chains" of neurons in the inferior parietal lobe could facilitate action understanding through linking sequences of actions and goals (see also Sinagaglia, 2009). Similar results have been found with human adults (Iacoboni et al., 2005). These findings suggest that there might be shared representations at higher-order levels that could play a role in linking active experience and action understanding. Thus, it is plausible that these representations may emerge in development and support early developments in action understanding. Clearly, further research is needed to investigate this possibility.

These open issues aside, the current findings support the notion that self-produced experience is uniquely beneficial for action perception in the first year of life. They shed light on the nature of information gained from active experience with means-end actions, indicating that the shift in one's own attention to the means or distal goal when learning to produce multi-step actions is similarly reflected in infants' perception of others' means-end actions.

Acknowledgments

References

Baldwin, D. A., and Moses, L. J. (2001). Links between social understanding and early word learning: challenges to current accounts. *Soc. Dev.* 10, 309–329. doi: 10.1111/1467-9507.00168

Bates, E., Carlson-Luden, V., and Bretherton, I. (1980). Perceptual aspects of tool using in infancy. *Infant Behav. Dev.* 3, 127–140. doi: 10.1016/S0163-6383(80)80017-8

Biro, S., and Leslie, A. M. (2006). Infants perception of goal-directed actions: development through cue-based bootstrapping. *Dev. Sci.* 3, 379–398. doi: 10.1111/j.1467-7687.2006.00544.x

Biro, S., Verschoor, S., and Coenen, L. (2011). Evidence for a unitary goal concept in 12-month-old infants. *Dev. Sci.* 14, 1255–1260. doi: 10.1111/j.1467-7687.2011.01042.x

Brandone, A. C., and Wellman, H. M. (2010). You can't always get what you want: infants understand failed goal-directed actions. *Psychol. Sci.* 20, 85–91. doi: 10.1111/j.1467-9280.2008.02246.x

Cannon, E. N., Simpson, E. A., Fox, N. A., Vanderwert, R. E., Woodward, A. L., and Ferrari, P. F. (2015). Relations between infants' emerging reach-grasp competence and event-related desynchronization in EEG. *Dev. Sci.* doi: 10.1111/desc.12295 [Epub ahead of print].

Cannon, E. N., and Woodward, A. L. (2012). Infants generate goal-based action predictions. *Dev. Sci.* 15, 292–298. doi: 10.1111/j.1467-7687.2011.01127.x

Casstevens, R. M. (2007). jHab: Java Habituation Software (version 1.0.0). [Computer Software]. Chevy Chase, MD.

Chen, Z., Sanchez, R. P., and Campbell, T. (1997). From beyond to within their grasp: the rudiments of analogical problem solving in 10- and 13-month-olds. *Dev. Psychol.* 33, 790–801. doi: 10.1037/0012-1649.33.5.790

Daum, M. M., Prinz, W., and Aschersleben, G. (2011). Perception and production of object-related grasping in 6-month-old infants. *J. Exp. Child Psychol.* 108, 810–818. doi: 10.1016/j.jecp.2010.10.003

Fagan, J. F. (1992). Intelligence: a theoretical viewpoint. *Curr. Dir. Psychol. Sci.* 1, 82–86. doi: 10.1111/1467-8721.ep10768727

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667. doi: 10.1126/science.1106138

Frick, A., and Wang, S. (2014). Mental spatial transformations in 14-and 16-month-old infants: effects of action and observational experience. *Child Dev.* 85, 278–293. doi: 10.1111/cdev.12116

Gerson, S. A., Bekkering, H., and Hunnius, S. (2014). Short-term Motor training, but not observational training, alters neurocognitive mechanisms of action processing in infancy. *J. Cogn. Neurosci.* doi: 10.1162/jocn_a_00774 [Epub ahead of print].

Gerson, S., and Woodward, A. L. (2010). "Building intentional action knowledge with one's hands," in *Neo-constructivism*, ed. S. P. Johnson (Oxford: Oxford University Press), 295–313.

Gerson, S., and Woodward, A. L. (2012). A claw is like my hand: comparison supports goal analysis in infants. *Cognition* 122, 181–192. doi: 10.1016/j.cognition.2011.10.014

Gerson, S. A., and Woodward, A. L. (2014a). Learning from their own actions: the unique effects of producing actions on infants' action understanding. *Child Dev.* 85, 264–277. doi: 10.1111/cdev.12115

Gerson, S. A., and Woodward, A. L. (2014b). The joint role of trained, untrained, and observed actions at the origins of goal recognition. *Infant Behav. Dev.* 37, 94–104. doi: 10.1016/j.infbeh.2013.12.013

Gerson, S. A., and Woodward, A. L. (2014c). Labels facilitate infants comparison of action goals. *J. Cogn. Dev.* 15, 197–212. doi: 10.1080/15248372.2013.777842

Hamlin, J. K., Hallinan, E. V., and Woodward, A. L. (2008). Do as I do: 7-month-old infants selectively reproduce others' goals. *Dev. Sci.* 11, 487–494. doi: 10.1111/j.1467-7687.2008.00694.x

Henderson, A. M. E., and Woodward, A. L. (2011). "Let's work together": what do infants understand about collaborative goals? *Cognition* 121, 12–21. doi: 10.1016/j.cognition.2011.05.008

Hofer, T., Hauf, P., and Aschersleben, G. (2005). Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behav. Dev.* 28, 466–480. doi: 10.1016/j.infbeh.2005.04.002

Hunnius, S., and Bekkering, H. (2014). Review article: what are you doing? How active and observational experience shape infants' action understanding. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130490.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3:529–535. doi: 10.1371/journal.pbio.0030079

Kanakogi, Y., and Itakura, S. (2010). The link between perception and action in early infancy: from the viewpoint of the direct-matching hypothesis. *Jpn. Psychol. Res.* 52, 121–131. doi: 10.1111/j.1468-5884.2010.00429.x

Keen, R. (2011). The development of problem solving in young children: a critical cognitive skill. *Annu. Rev. Psychol.* 62, 1–21. doi: 10.1146/annurev.psych.031809.130730

Király, I., Jovanovic, B., Prinz, W., Aschersleben, G., and Gergely, G. (2003). The early origins of goal attribution in infancy. *Conscious. cogn.* 12, 752–769. doi: 10.1016/S1053-8100(03)00084-9

Krogh-Jespersen, S., and Woodward, A. L. (2014). Making smart social judgments takes time: infants' recruitment of goal information when generating action predictions. *PLos ONE* 9:e98085. doi: 10.1371/journal.pone.0098085

Libertus, K., and Needham, A. (2010). Teach to reach: the effects of active versus passive reaching experiences on action and perception. *Vision Res.* 50, 2750–2757. doi: 10.1016/j.visres.2010.09.001

Libertus, K., and Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Dev. Sci.* 14, 1355–1364. doi: 10.1111/j.1467-7687.2011.01084.x

Libertus, K., and Needham, A. (2014). Encouragement is nothing without control: factors influencing the development of reaching and face preference. *J. Motor Learn. Dev.* 2, 16–27. doi: 10.1123/jmld.2013-0019

Lockman, J. J. (2000). A perception–action perspective on tool use development. *Child Dev.* 71, 137–144. doi: 10.1111/1467-8624.00127

Loucks, J., and Sommerville, J. A. (2012). The role of motor experience in understanding action function: The case of the precision grasp. *Child Dev.* 83, 801–809. doi: 10.1111/j.1467-8624.2012.01735.x

Mangold, P. (1998). *Interact [Computer Software]*. Arnstorf: Mangold International. Available at: www.mangold-international.com

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Dev. Sci.* 10, 126–134. doi: 10.1111/j.1467-7687.2007.00574.x

Munakata, Y., Bauer, D., Stackhouse, T., Landgraf, L., and Huddleston, J. (2002). Rich interpretation vs. deflationary accounts in cognitive development: the case of means-end skills in 7-month-old infants. *Cognition* 83, B43–B53. doi: 10.1016/S0010-0277(02)00007-0

Piaget, J. (1954). *The Construction of Reality in the Child*. New York, NY: Basic Books, Inc. doi: 10.1037/11168-000

Rakison, D. H., and Krogh, L. (2011). Does causal action facilitate causal perception in infants younger than 6 months of age? *Dev. Sci.* 15, 45–53. doi: 10.1111/j.1467-7687.2011.01096.x

Rizzolatti, G., and Craighero, L. (2004). The mirror neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi: 10.1146/annurev.neuro.27.070203.144230

Saby, J. N., Marshall, P. J., and Meltzoff, A. N. (2012). Neural correlates of being imitated: an EEG study in preverbal infants. *Soc. Neurosci.* 7, 650–661. doi: 10.1080/17470919.2012.691429

Sinagaglia, C. (2009). Mirror in action. *J. Conscious. Stud.* 16, 309–334. doi: 10.1038/nrn2805

Sommerville, J. A., Blumenthal, E. J., Venema, K., and Sage, K. D. (2011). "The body in action: the impact of self-produced action on infants action perception and understanding," in *Early Development of Body Representations*, eds V. Slaughter and C. Brownwell (Cambridge: Cambridge University Press), 247–266.

Sommerville, J. A., and Crane, C. C. (2009). Ten-month-old infants use prior information to identify an actor's goal. *Dev. sci.* 12, 314–325. doi: 10.1111/j.1467-7687.2008.00787.x

Sommerville, J. A., Hildebrand, E., and Crane, C. C. (2008). Experience matters: the impact of doing versus watching on infants' subsequent perception of tool use events. *Dev. Psychol.* 44, 1249–1256. doi: 10.1037/a0012296

Sommerville, J. A., and Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition* 95, 1–30. doi: 10.1016/j.cognition.2003.12.004

Sommerville, J. A., Woodward, A. L., and Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96, B1–B11. doi: 10.1016/j.cognition.2004.07.004

Southgate, V., Johnson, M. H., Osborne, T., and Csibra, G. (2009). Predictive motor activation during action observation in human infants. *Biol. Lett.* 5, 769–772. doi: 10.1098/rsbl.2009.0474

Thoermer, C., Woodward, A., Eisenbeis, H., Kristen, S., and Sodian, B. (2013). To get the grasp: seven-month-olds encode and reproduce goal-directed grasping. *J. Exp. Child Psychol.* 116, 499–509. doi: 10.1016/j.jecp.2012.12.007

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

van Elk, M., van Schie, H. T., Hunnius, S., Vesper, C., and Bekkering, H. (2008). You'll never crawl alone: neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage* 43, 808–814. doi: 10.1016/j.neuroimage.2008.07.057

Wellman, H. M., Lopez-Duran, S., LaBounty, J., and Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Dev. Psychol.* 44, 618–623. doi: 10.1037/0012-1649.44.2.618

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., and LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Dev. Sci.* 7, 283–288. doi: 10.1111/j.1467-7687.2004.00347.x

Willatts, P. (1999). Development of means-end behavior in young infants: pulling a support to retrieve a distant object. *Dev. Psychol.* 35, 651–667. doi: 10.1037//0012-1649.35.3.651

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34. doi: 10.1016/S0010-0277(98)00058-4

Woodward, A. L., and Gerson, S. A. (2014). Review article: mirroring and the development of action understanding. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130181. doi: 10.1098/rstb.2013.0181

Woodward, A. L., and Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychol. Sci.* 11, 73–76. doi: 10.1111/1467-9280.00218

Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M. E., and Buresh, J. S. (2009). "The emergence of intention attribution in infancy," in *The Psychology of Learning and Motivation*, Vol. 51, ed. B. Ross (Waltham: Academic Press), 187–222.

Variability in social reasoning: the influence of attachment security on the attribution of goals

Kristen A. Dunfield[1]* and Susan C. Johnson[2]

[1] Department of Psychology, Center for Research in Human Development, Concordia University, Montreal, QC, Canada, [2] Department of Psychology, The Ohio State University, Columbus, OH, USA

Over the last half decade there has been a growing move to apply the methods and theory of cognitive development to questions regarding infants' social understanding. Though this combination has afforded exciting opportunities to better understand our species' unique social cognitive abilities, the resulting findings do not always lead to the same conclusions. For example, a growing body of research has found support for both universal similarity and individual differences in infants' social reasoning about others' responses to incomplete goals. The present research examines this apparent contradiction by assessing the influence of attachment security on the ability of university undergraduates to represent instrumental needs versus social-emotional distress. When the two varieties of goals were clearly differentiated, we observed a universally similar pattern of results (Experiments 1A/B). However, when the goals were combined, and both instrumental need and social-emotional distress were presented together, individual differences emerged (Experiments 2 and 3). Taken together, these results demonstrate that by integrating the two perspectives of shared universals and individual differences, important points of contact can be revealed supporting a deeper, more nuanced understanding of the nature of human social reasoning.

Keywords: social-cognitive development, social development, cognitive development, social evaluation, attachment, prosocial behavior

Introduction

Humans are unique in the integral role that social relationships play in our success as a species (e.g., Brewer and Caporael, 2006). As a result, there is considerable interest in understanding how individuals come to understand, engage with, and navigate their social environment. Though historically social development and cognitive development were viewed as integrally intertwined (e.g., Piaget, 1945/1995; Vygotsky, 1978; see also, Dweck, 2013), for decades these two lines of inquiry have been pursued largely independently; with social developmentalists typically examining how variability in experiences leads to differences in well-being while cognitive developmentalists typically examine commonalities in the content and development of children's minds. Recently these two perspectives have been reunited (e.g.,Olson and Dweck, 2008, 2009; Dweck, 2013). This integration has helped build important points of contact between a variety of sub-disciplines of psychology, however in doing so, it has become apparent that sometimes studies which appear to address highly similar questions may lead to quite different conclusions.

Variability in the Consistency of Social Reasoning

One area where similar studies have led to different conclusions is within the domain of social reasoning. Specifically, a recent explosion of interest in children's reasoning about others—particularly within the domain of other-oriented behavior—has led to a body of literature in which infants' representation of positive versus negative interactions (e.g., Premack and Premack, 1997), preferences for helpers versus hinderers (e.g., Hamlin et al., 2007), and expectations following prosocial versus antisocial interactions (e.g., Kuhlmeier et al., 2003; Johnson et al., 2007) appear to support both universal consistency and individual differences (e.g., Johnson et al., 2013).

Universal Expectations of Helpers and Hinderers

One line of research utilizes the "helper/hinderer paradigm" to examine infants' reasoning about others' responses to instrumental needs and finds a single pattern of common expectations. In these studies, infants watch a brief animation of small ball (the "Climber") trying and failing to reach the top of a steep hill. On alternating trials, one of two similarly sized shapes (typically a triangle and square) comes down and either pushes the Climber to the top of the hill (the "Helper") or pushes the Climber to the bottom of the hill (the "Hinderer"). Across a variety of dependent measures, infants appear surprisingly consistent in their expectations of, and preferences for, helpful versus hindering characters.

In the original version of the helper/hinderer paradigm, after infants were habituated to the climb, they were shown the three characters interacting in a novel context. By 12 months, infants differentiated between scenes in which the Climber approached the Helper versus the Hinderer and preferred the video in which the Climber approached the Helper (Kuhlmeier et al., 2003). This preference was consistent with pilot adult participants' tendency to report seeing "the ball as 'liking' or 'preferring' the helper object" (Kuhlmeier et al., 2003, p. 402). And, although the participants varied in the degree to which they differentiated between the two types of approach, infants who showed the largest difference in attention to the generally preferred (approach Helper) over non-preferred (approach Hinderer) outcome showed more advanced theory of mind at 4 years than infants who show smaller, or reversed, differences in attention (Yamaguchi et al., 2009); suggesting that this preference was not only shared across individuals but was also associated with relatively more mature social cognitive development.

More recent research finds that infants not only differentiate between these two varieties of approach, but also actively predict them. Using eye-tracking methodology, 12-month-old infants' anticipatory looks were recorded while they observed the Climber ambiguously approaching the Helper or Hinderer. Twelve out of 17 infants (70.5%) predicted that the Climber would approach the Helper as opposed to the Hinderer (Fawcett and Liszkowski, 2012). Moreover, when given the opportunity to choose between the Helper and Hinderer, 12 out of 12 (100%) 6-month-olds and 14 out of 16 (87.5%) 10-month-olds preferred the Helper (Experiment 1, Hamlin et al., 2007; see also Hamlin, 2014 for a replication of this finding). Together, these studies converge to suggest that when evaluating others' responses to instrumental needs, most infants prefer helpers to hinderers and expect others to feel similarly. Indeed, these results are so striking that they have been used as evidence in support of the existence of a universal, innate moral core (Hamlin, 2013).

Individual Differences in Expectations of Caregivers

In contrast, when infants' reasoning about others' responses to social emotional distress have been investigated utilizing a "caregiver paradigm", the results appear to support robust individual differences in expectations (Johnson et al., 2007, 2010). Utilizing a similar experimental design (i.e., visual habituation), and strikingly similar abstract, animated agents (i.e., a small ball struggling to climb a steep hill) studies find that, around their first birthday, infants' expectations of and preferences for responsive versus unresponsive caregivers reflect multiple distinct patterns of expectations rooted in personal caregiving experiences.

In these studies, infants are habituated to a large "Mommy" ball climbing a steep hill and leaving her "Baby" at the bottom, crying and unable to follow. Despite clear similarities to the previously described studies, infants' expectations of, and preferences for, responsive versus unresponsive caregivers varied as a function of personal attachment style. Securely attached infants expected the Caregiver to return to the Baby, while insecurely attached infants expected the Caregiver to ignore the distressed Baby (Johnson et al., 2007; Study 1, Johnson et al., 2010). When the infants were subsequently presented with a video of the Baby alternately approaching a responsive versus unresponsive Mommy, securely attached infants expected the Baby to prefer the responsive Mommy whereas insecurely attached infants expected the Baby to prefer the unresponsive Mommy (Study 3, Johnson et al., 2010). Finally, when infants were shown a partially responsive Mommy (who comes part-way back down the hill to meet the distressed Baby) securely attached infants expected that the Baby would approach the Mommy while insecurely attached infants differed in their expectations based on their unique variety of attachment insecurity. Like securely attached infants, insecure-resistant infants were surprised when the Baby moved further away from the partially responsive Mommy, whereas insecure-avoidant infants were surprised when the Baby approached a partially responsive Mommy (Study 2, Johnson et al., 2010). Together, these findings suggest that relatively stable, early emerging individual differences exert an important influence on the representation and processing of valenced social interactions (Johnson et al., 2007, 2010).

As these two lines of research address common theoretical questions using similar methodologies and stimuli, yet produce different patterns of empirical findings, we are left with an important question regarding how to integrate these results. One explanation is that we only see what we are looking for. It is possible that the helper/hinderer paradigm (e.g., Kuhlmeier et al., 2003; Hamlin et al., 2007) finds universal similarity in reasoning simply because sub-groups were not analyzed. This seems unlikely given that, where counts are available, between 70.5% (Fawcett and Liszkowski, 2012) and 100% (Hamlin et al., 2007) of infants showed similar expectations and preferences in helper/hinderer paradigm yet, only about half of infant samples are securely attached (e.g., 10 out of 21 infants in Johnson et al., 2007; 14 out of

30 infants in Johnson et al., 2010, Study 2; and 20 out of 35 infants in Johnson et al., 2010, Study 3). Moreover, when the responses of the securely and insecurely attached infants were collapsed in the caregiver paradigm, results were not distinguishable from chance (see Johnson et al., 2013). This suggests that the different patterns of results across the two varieties of studies are not simply a reflection of different analytical choices.

An alternative explanation is that these two sets of studies, though superficially similar, actually tap into different underlying representations. Although both sets of studies show a small ball struggling to achieve a goal that is either supported (i.e., when the Helper pushes the Climber up the hill or the Mommy responds to the Baby's distress) or thwarted (i.e., when the Hinderer pushes the Climber down the hill or the Mommy ignores the Baby's distress), the two sets of studies may require the attribution of different varieties, or at least complexities, of goals leading to differences in subsequent representations and expectations. In the helper/hinderer studies (e.g., Kuhlmeier et al., 2003), one must represent the Climber's instrumental goal (i.e., "get up the hill") in order to interpret and evaluate the subsequent social interactions (i.e., helping versus hindering). In contrast, in the caregiving paradigm (e.g., Johnson et al., 2007), one must represent the Baby's social-emotional goal (i.e., "get to Mommy") in order to interpret and evaluate the subsequent social interaction (i.e., responsive versus unresponsive caregiving). Thus, the similar social expectations assessed at test may rely on different initial goal representations and it is this initial willingness or ability to represent the target ball's goal that may affect the consistency with which social reasoning occurs.

Although both lines of research aim to understand how individuals reason about agents in their environment, it is possible that the different patterns of results reflect important asymmetries in the way individual differences influence our representations of, and expectations about, the goals pursued by others (see also, Johnson et al., 2013) with one set of studies (those employing the helper/hinderer paradigm) relying on the ability to first represent the instrumental goal of an agent (the Climber) acting on an object (the hill) and the other (those employing the caregiver paradigm) relying on the ability to represent the social-emotional goal of an agent (the Baby) acting on another agent (the Mommy; see Spelke, 2014, for a similar distinction). In other words, these two varieties of stimuli may require that participants first represent two different types of goals (i.e., instrumental versus social-emotional) before they can reason about the subsequent social interactions and it is at the level of goal representation that the participants may vary.

Reconciling Differences

The present research attempts to understand why there appears to be both universal similarities and individual differences in the way individuals reason about those who respond positively versus negatively to others' unfulfilled goals. By utilizing methods and theory from developmental, social, and cognitive psychology we will examine the extent to which these apparent differences can be understood by examining the types of goals individuals are initially representing when observing these abstract, stylized, animated interactions. Critically, because human infants are

limited in the types of responses they can provide, and previous research suggests that there is likely continuity in the way these videos are perceived across the lifespan (see Kuhlmeier et al., 2003), we will examine this question in a much older participant population, namely university undergraduates. In a series of three experiments, utilizing both free-response and eye tracking methodologies, we will examine how attachment security affects the way university undergraduates represent and discuss two varieties of incomplete goals: instrumental need (e.g., agents acting on objects; e.g., Kuhlmeier et al., 2003) and social-emotional distress (e.g., agents acting with agents; e.g., Johnson et al., 2010). In doing so, we aim to demonstrate how the observation of apparent contradictions can help us to develop a more nuanced understanding of the nature of social reasoning.

Attachment Security

Early experiences within the caregiver–child dyad are thought to result in internal working models of relationships that organize and bias subsequent social-emotional processing (Bowlby, 1969/1982). Individuals who are securely attached readily approach relationship partners, openly share their needs, and expect that close others will accept and respond appropriately to their distress. In contrast, individuals who are insecurely attached typically avoid or resist their relationship partners and expect close others to reject their needs or respond unpredictably (Ainsworth et al., 1978; De Wolff and van Ijzendoorn, 1997; Cassidy and Shaver, 2008). For decades, researchers have examined the influence of these internal working models of attachment on individuals' social development, demonstrating that from infancy to adulthood, early experiences in the caregiver–child dyad exert reliable and robust influences on social processing, representation, and behavior (e.g., Main et al., 1985; Dykas and Cassidy, 2011).

Though it is clear that variability in internal working models of attachment affect a number of social-emotional outcomes, what is less clear is where these differences originate. Historically, researchers have examined the link between objective differences in caregiving and attachment security. From this perspective, researchers have found that infants who receive responsive caregiving are considerably more likely to be securely attached than infants who do not (De Wolff and van Ijzendoorn, 1997). Yet, despite the reliability of this finding, variability in the quality of parenting received only accounts for a minority of the observed variance in attachment security. It has recently been proposed that the objective quality of the parenting that the infant receives is less important than the infant's subjective construal of their experiences (Johnson and Chen, 2011). Specifically, individual differences in attachment can be thought to result not from *objective* differences in social emotional experiences, but from the way these experiences are *subjectively* construed as a function of individual differences in oxytocin receptor (OXTR) genes. In human infants, it appears as though variability in OXTR influences *both* attention to emotionally salient stimuli and attachment security (Johnson and Chen, 2011), suggesting that stable differences in attachment are not simply related to the experiences one has, but to the manner in which one perceives those experiences. Though both accounts posit that variability

in attachment security reflects differences in the experience of early caregiving, they differ in regards to whether it is the care received or the construal of the care that biases subsequent social-emotional information processing.

Prosocial Behavior

In a related line of research examining the development of other-oriented behavior, there is growing consensus that humans recognize and respond to a variety of problems experienced by others, ranging from relatively simple, emotion-neutral instrumental needs to relatively complex, highly emotional distress (e.g., Dunfield, 2014; Eisenberg et al., 2015). The ability to respond to each of these different types of problems appears to emerge at different ages (e.g., Dunfield et al., 2011) and develop independently of each other (e.g., Svetlova et al., 2010; Dunfield and Kuhlmeier, 2013; Paulus et al., 2013). Together, these findings have led to the proposal that recognizing instrumental need relies on different underlying representations than recognizing emotional distress (e.g., Warneken and Tomasello, 2009; Svetlova et al., 2010; Dunfield, 2014).

Acting effectively on behalf of another requires the ability to represent the problem that the individual is facing, the ability to recognize the required intervention, and the motivation to help alleviate the problem. Recent research supports this position finding that early helping is dependent on children's abilities to represent stable, abstract goals in others (Hobbs and Spelke, 2015). Yet not all goals are represented with equal ease. Infants represent action goals such as reaching before they understand more mentalistic goals such as using a point to direct attention (Woodward et al., 2001). Relatedly, when examining the literature on the development of the different types of evaluations that may underlie different varieties of prosocial behavior, the ability to represent and reason about others' instrumental goals appears to emerge earlier than the ability to reason about others' emotional distress (see Dunfield, 2014, for a review). Moreover, these two varieties of goal attributions are not only dissociable at the developmental level, but appear to be supported by two distinct neural systems. While the mirror neuron system supports the representation of familiar, frequently executed actions based on low-level behavioral input, the metalizing system appears to support the representation of others' thoughts and beliefs on the basis of social intelligence (Van Overwalle and Baetens, 2009). Finally, these differences in underlying representations affect the ease with which children respond to others' needs. Although children begin engaging in instrumental help as early as 14 months (Warneken and Tomasello, 2007), social-emotional helping (i.e., getting another's attention on behalf of a third-party) develops much later (closer to 3 years) and is less frequent and robust (i.e., 16 out of 32 toddlers helping in social tasks versus 29 out of 32 toddlers helping instrumental tasks, Experiment 1; Beier et al., 2014). Together, it is clear that there is considerable heterogeneity in the ability to represent the problems that others face and that these differences affect when and how individuals act on behalf of others.

Critically, attachment security should not necessarily bias the representation of all goals equally. While securely attached individuals have a positive self-construal and feel confident in their ability to accept others' needs for closeness, sympathy, and support, insecurely attached individuals typically do not. As such, variations in attachment security should exert a greater influence on tasks that require the interpretation of more emotionally laden social stimuli than less emotional instrumental stimuli (see Dykas and Cassidy, 2011, for a review). Because instrumental needs are based on the ability to reason about agents acting on objects, while social-emotional distress requires the ability and willingness to represent another's negative emotions and social relationships, the ability and willingness to reason about social emotional distress should be uniquely affected by internal working models of attachment. Thus the apparent contradiction in the developmental literature investigating social reasoning may reflect the fact that representing instrumental need is distinct from representing social-emotional distress and the latter shows more variability because it activates, and is influenced by, the social schema that underlie attachment security (e.g., Johnson et al., 2013). However, because attachment security affects attention to, processing of, and the ability to discuss emotionally laden social stimuli, the mechanism through which attachment security will exert its influence is not presently clear.

Current Study

In order to better understand variability in social reasoning and provide explanatory insight into the apparent contradiction between universal similarity and individual differences in social cognition, we asked university undergraduates to describe a variety of abstract, animated social interactions that were based on the two original hill stimuli (e.g., Kuhlmeier et al., 2003; Johnson et al., 2007). Specifically, we created three brief videos in which a small yellow ball interacted with a large yellow ball and a hill. To disentangle the role of attachment security on the processing of different types of goals, we systematically varied the interaction between the two balls and the hill in order to afford participants the opportunity to discuss both the instrumental (the ball is trying to get up the hill), and social-emotional (the ball is trying to get the attention of, or in proximity to, a social partner) aspects of the interaction. We predicted that if attachment security differentially biases the processing of instrumental needs versus social-emotional distress, then: (1) both securely and insecurely attached participants will discuss instrumental goals similarly (Study 1A); (2) insecurely attached participants will tend to avoid discussing social goals (Study 1B), particularly when the stimuli are complex or ambiguous (Study 2); and (3) any variability in the tendency to report social goals across the two groups will be associated with an attentional bias that is consistent with the underlying attachment representations (Study 3).

Study 1A

The goal of Study 1 was to determine if individual differences in attachment security affected participants' recognition of instrumental need versus social-emotional distress. Across two studies, two groups of participants watched as a small ball struggled to complete either an instrumental "hill" goal (Study 1A) or an emotional "social" goal (Study 1B). In both videos, the small ball was separated from a larger ball. However, the

videos varied regarding what the little ball was attempting to do. Specifically, in Study 1A (instrumental), the small ball tries but fails to climb the hill, whereas in Study 1B (social), the small ball tries but fails to get the larger ball's attention. By presenting two separate groups of participants with the two types of goals independently, we can begin to determine the extent to which attachment security imposes an absolute limit on the processing of social stimuli.

Method

The Office of Responsible Research Practices at the Ohio State University approved all of the research reported in this manuscript.

Participants

Ninety-one undergraduate students (39 female) enrolled in an Introductory Psychology course participated for partial course credit.

Measures

Participants were shown a brief (20 s) animated video in which a small yellow ball attempts to climb a relatively steep hill while a larger ball looks on (**Figure 1A**). The small ball makes two attempts at ascent separated by a "sigh" in which the small ball expands and contracts while darkening in color. Both balls had faces but maintained a neutral expression. Following the video, participants were given a small piece of paper and asked to briefly describe what they thought the video was about.

After the participants described the video, they completed the Experiences in Close Relationships questionnaire (ECR; Brennan et al., 1998), which measures attachment security along two dimensions, namely anxiety and avoidance. Attachment anxiety refers to the concern that others will be unavailable in times of need (e.g., "I worry about being abandoned"), while attachment avoidance refers to the tendency to avoid potential pain by keeping others at a distance (e.g., "I feel comfortable sharing my private thoughts and feelings with my partner"). Participants were asked to think about their close relationships in general, without focusing on a specific partner, and rate the extent to which each statement accurately reflects their feelings.

Coding

To determine if there were individual differences in the types of goals that the participants attributed, we developed a single coding scheme that we applied consistently across all three free-response studies. First we coded for the presence of any goal directed language. Participants were given a general "goal" code if they used agentive language such as "trying," "wanting," "attempting," or "failing." Next, we categorized the specific types of goals that the participants identified. Of particular interest was the participants' tendency to discuss the instrumental (hill) goal and the social (reunion) goal. Hill goals were coded when the participant indicated that the small ball was trying to get up the hill (e.g., "a small circle tried to go up a hill but failed"). Social goals were coded when participants explicitly referred to either a social partner (e.g., a mother, parent, or friend) or a social behavior
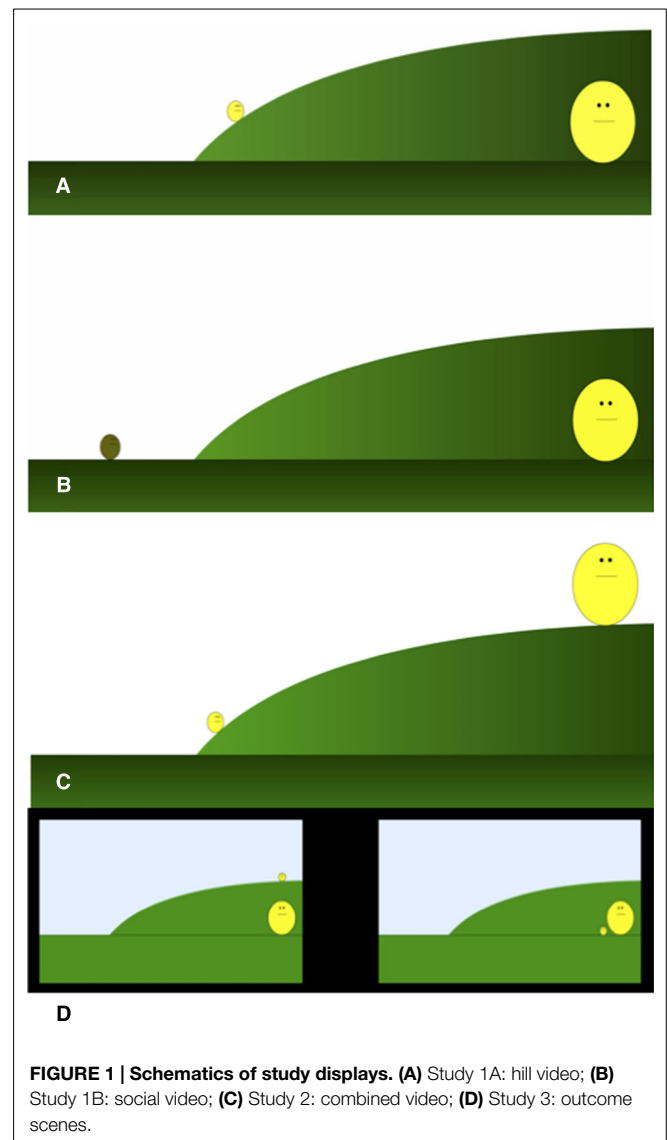
FIGURE 1 | Schematics of study displays. (A) Study 1A: hill video; **(B)** Study 1B: social video; **(C)** Study 2: combined video; **(D)** Study 3: outcome scenes.

(e.g., "get attention") as the small ball's target. To allow for a more nuanced understanding of the effect of attachment security on the types of goals people represent, these codes were not mutually exclusive. Participants who discussed both goals were given both codes (e.g., "a baby trying to climb the hill to reach his parent"). Some participants discussed the small ball's behavior in terms of goals that were not related to either the hill or other agent (e.g., "trying to get what you want is not as easy as you think"). These participants received a goal code, but neither of the specific codes. A secondary coder, blind to attachment status and the purpose of the study, coded all of the responses. Agreement was near perfect for goals (96%, κ = 0.92), hills (96%, κ = 0.87), and social goals (94% κ = 0.86).

Results and Discussion
Attachment Classification

To examine whether individual differences in attachment security affects the attribution of goals to others, and to allow comparison

to the developmental literature, we created two groups of participants based on their ECR scores (Brennan et al., 1998). Though there are many ways to classify attachment security, we chose to create two groups for our main analysis because the most comparable infant condition found that expectations regarding the Mommy's behavior (responsive versus unresponsive) differed between securely and insecurely attached infants but did not differ between varieties of insecurely attached infants (see Johnson et al., 2007; Study 1 Johnson et al., 2010). In our sample, the secure group includes individuals who were low on both attachment anxiety and avoidance ($N = 22$, 24.2%, 11 female) and represents individuals who are likely to process both instrumental and social information in an open and relatively accurate manner. In contrast, the insecure group includes participants who are high on one, or both, of the dimensions of attachment insecurity ($N = 69$, 75.8%, 28 female). These individuals are hypothesized to have more negative expectations regarding others' tendency to seek and accept comfort and are expected to interpret social information in a biased and selective manner. Both attachment anxiety and attachment avoidance were significantly higher in the insecure group than the secure group [anxiety, $t(89) = 6.72$, $p < 0.001$; avoidance, $t(89) = 6.00$, $p < 0.001$].

Verbal Reports

When presented with unambiguous instrumental goals, there were no group-based differences in the tendency to report goals [$\chi^2(1, N = 91) = 0.05$, $p = 0.82$, $\varphi = 0.02$], nor in the types of goals reported [hill: $\chi^2(1, N = 91) = 1.12$, $p = 0.29$, $\varphi = 0.11$; social: $\chi^2(1, N = 91) = 1.69$, $p = 0.19$, $\varphi = 0.14$; **Figure 2A**][1].

Consistent with our predictions, we observed universal similarity in the ability to represent and discuss instrumental goals. When an agent appears to be unsuccessfully acting on an object (in this case a steep hill), attachment security exerts little influence on the ability to represent the underlying goal. Study 1B extends this finding by examining individual differences in the representation and reporting of social goals.

Study 1B

To determine if variability in the universality of social reasoning is related to differences in the underlying goal representations, particularly when the goals are social, Study 1B presented participants with a video designed to reflect a purely social problem (in this case, a Mommy abandoning her baby).

Method
Participants

Ninety undergraduate students (50 female) enrolled in an Introductory Psychology course, who did not participate in Study 1A, participated for partial course credit. Three additional participants were tested but excluded from analysis for failure to complete all measures.

---

[1] We conducted the same analyses treating the three varieties of attachment insecurity separately and observe the identical pattern of results: Goals: $\chi^2(3, N = 91) = 1.2$, $p = 0.75$, $\varphi = 0.11$; Hill: $\chi^2(3, N = 91) = 3.90$, $p = 0.27$, $\varphi = 0.21$; Social: $\chi^2(3, N = 91) = 3.15$, $p = 0.37$, $\varphi = 0.19$.

FIGURE 2 | Results of Studies IA–2. The bars represent the proportion of participants reporting any goals (left panel), and the specific goals of interest (right panel). **(A)** Study 1A: hill video; **(B)** Study 1B: social video; **(C)** Study 2: combined video. The "*" indicates the difference is significant at $p < 0.05$.
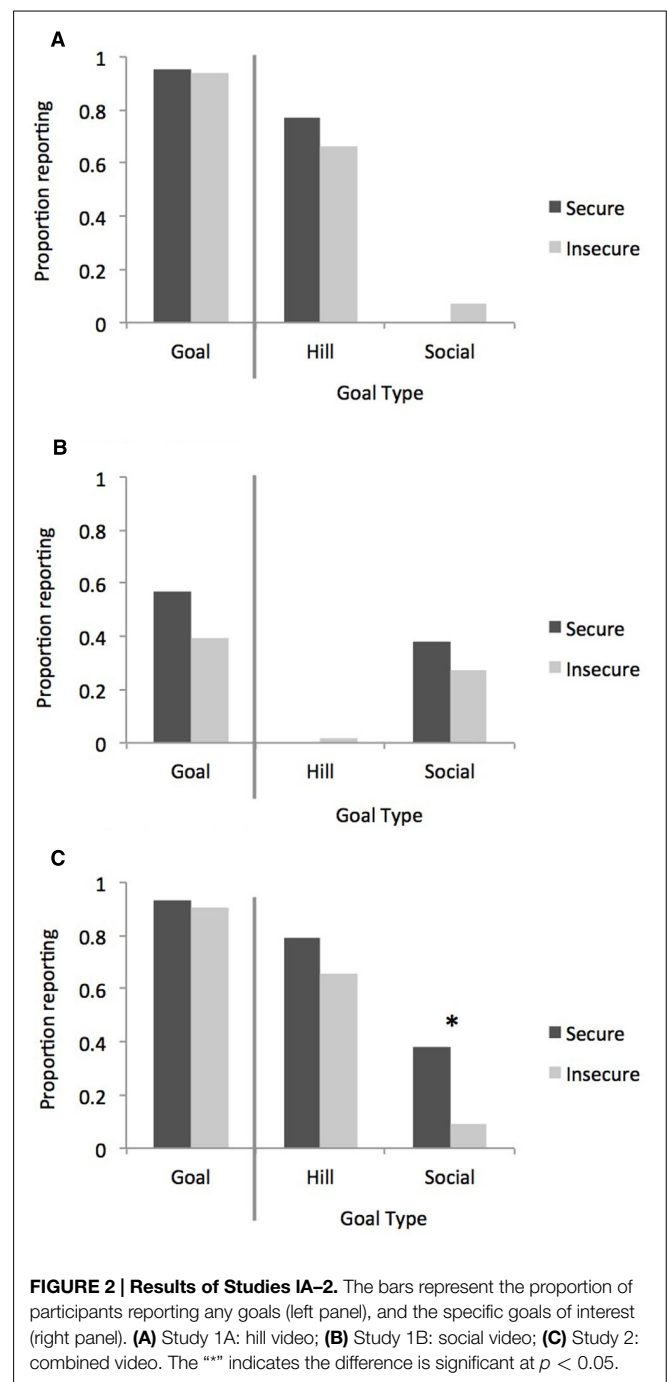
Measures

Study 1B was largely identical to Study 1A with the exception that the video was modified to reflect a single social-emotional goal (**Figure 1B**). Instead of attempting to climb the hill, the small ball turned to look at the larger ball and then engaged in a series of expansions and contractions, associated with a darkening of color, intended to represent distress. Goals were coded as described in Study 1A. A secondary blind coder coded all reports; agreement was near perfect for goals (100%, $\kappa = 1$), hill (98%, $\kappa = 0.98$), and social (96%, $\kappa = 0.83$). After watching and describing the videos, participants completed the ECR.

Results and Discussion
Attachment Classification

Relative to secure participants ($N = 21$, 23.3%, 12 female), insecure participants ($N = 69$, 76.7%, 38 female) had lower attachment anxiety and avoidance: anxiety [$t(88) = 5.47$, $p < 0.001$] and avoidance [$t(88) = 6.52$, $p < 0.001$].

Verbal Reports

Despite being presented with a purely social interaction, there were no attachment related differences in the tendency to report goals [$\chi^2(1, N = 90) = 2.13$, $p = 0.15$, $\varphi = 0.15$], nor in the specific goals reported [hill: $\chi^2(1, N = 90) = 0.31$, $p = 0.58$, $\varphi = 0.05$; social: $\chi^2(1, N = 90) = 0.85$, $p = 0.35$, $\varphi = 0.10$; **Figure 2B**][2].

Study 1B replicates and extends the findings of Study 1A by demonstrating that when presented with pure and unambiguous goals, individuals make the same attributions regardless of goal type or attachment categorization. Though these findings appear to suggest that differences in attachment security do not differentially influence the ability to represent instrumental versus social goals, because social schemas, such as internal working models of attachment, are particularly likely to bias processing when stimuli are complex or ambiguous (e.g., Baldwin, 1992; Johnson et al., 2013) it is possible that separating the two goals and presenting them independently and unambiguously diluted the effect.

Consistent with the proposal that schemas have a greater influence on the representation of ambiguous stimuli, Johnson et al. (2007, 2010) first documented individual differences in social reasoning when both the hill and social goal were presented together. Unlike our pure videos, the original caregiver paradigm showed the Mommy distressing the Baby by climbing up a steep hill, affording both an instrumental (the baby simply cannot get up the hill) and social-emotional (the baby is distressed because it cannot get to its Mommy) problems. Given this design, it is possible that different participants were attending to different aspects of interaction. To address this consideration, and explore the extent to which attachment security affects the interpretation of *complex/ambiguous* problems, we modified our videos to make them more similar to Johnson et al. (2007). Specifically, we created a new video in which both the hill and social goals were equally salient.

Study 2

Study 2 aimed to determine if individual differences in attachment security affected participants' recognition of instrumental need versus social-emotional distress in complex scenes. To that end, participants watched a video that included *both* the instrumental "hill" goal of Kuhlmeier et al. (2003), and the social "reunion" goal of Johnson et al. (2010). Because the video was complex and included both an instrumental and social goal, we predicted that although all participants should be able to recognize goal directed

behavior, and both groups of participants should be equally likely to discuss the instrumental goal, insecurely attached individuals will avoid reporting the social goal because this video, unlike the pure social video, affords this option.

Method
Participants

Ninety-three undergraduate students (45 female) enrolled in an Introductory Psychology course participated for partial course credit. One additional student participated in the study but failed to complete all the measures and was removed from subsequent analysis.

Measures

Largely identical to the previous two studies, the only modification was the content of the videos. Specifically, we moved the large ball from the bottom of the hill to the top thus combining the small ball's instrumental and social goals (**Figure 1C**). In order to make both varieties of goals equally salient, and comparable to Studies 1A/B, the small ball attempts to climb the hill once, expands and contracts once, then, at the bottom of the hill, expands and darkens in color, appearing to cry. The larger ball remains motionless at the top of the hill for the duration of the video. Consistent with the previous videos, both balls had faces but maintained a neutral expression. Following the video participants completed the ECR. Again, all reports were coded by a secondary, blind coder and agreement was high (97%, $\kappa = 0.79$), hill (94%, $\kappa = 0.84$), and social (98%, $\kappa = 0.93$).

Results and Discussion
Attachment Classification

Both attachment anxiety and avoidance were lower in the secure group ($N = 29$, 31.2%, 11 female) than the insecure group [$N = 64$, 68.8%, 34 female; anxiety, $t(91) = 5.74$, $p < 0.001$; avoidance, $t(91) = 5.98$, $p < 0.001$].

Verbal Reports

Both groups of participants were equally likely to discuss the ball's behavior in agentive, goal-directed language [$\chi^2(1, N = 93) = 0.16$, $p = 0.69$, $\varphi = 0.04$; **Figure 2C**]. Moreover, both groups were equally likely to recognize and report the instrumental "hill" goal [$\chi^2(1, N = 93) = 1.78$, $p = 0.18$, $\varphi = 0.14$]. However, consistent with our hypotheses, the groups differed in their tendency to report the "social" goal [$\chi^2(1, N = 93) = 10.89$, $p = 0.001$, $\varphi = 0.34$][3]; specifically, insecurely attached participants were significantly less likely than securely attached participants to report the Baby's social goal of reuniting with the Mommy.

To determine whether it was attachment insecurity in general or one of the continuous attachment dimensions in particular that affected participant's tendency to report the social goal, we conducted a logistic regression with attachment anxiety, avoidance, and their interaction as continuous, independent

---

[2]Again, the pattern of results remains the same when the three varieties of attachment insecurity are treated as separate groups: Goals: $\chi^2(3, N = 90) = 2.31$, $p = 0.51$, $\varphi = 0.16$; Hill: $\chi^2(3, N = 90) = 3.32$, $p = 0.34$, $\varphi = 0.19$; Social: $\chi^2(3, N = 90) = 1.25$, $p = 0.74$, $\varphi = 0.12$.

[3]We analyze the three varieties of attachment insecurity separately the pattern of results is identical: Goals: $\chi^2(3, N = 93) = 1.50$, $p = 0.68$, $\varphi = 0.13$; Hill: $\chi^2(3, N = 93) = 3.49$, $p = 0.32$, $\varphi = 0.19$; Social: $\chi^2(3, N = 93) = 11.33$, $p = 0.01$, $\varphi = 0.35$.

predictors. The overall model accounted for a significant portion of the variance [$\chi^2(3, N = 93) = 8.17$, $p = 0.04$]; however, none of the independent predictors were significant [p's > 0.154]. Indicating that it is attachment insecurity broadly, as opposed to either of the specific dimensions that influenced the participants' tendency to report the social goal.

Study 2 examined the extent to which individual differences in attachment security affected the representation of two varieties of goals that varied in their social-emotional content. Differences in attachment security exerted a greater influence on the processing of social-emotional than instrumental goals, but only when the two types of goals were presented together. This result appears consistent with past research suggesting that insecure attachment categorically biases social representations irrespective of the nature of the insecurity (i.e., avoidance or anxiety; see Johnson et al., 2007, for a similar result). Moreover, these findings support the proposal that an individual's pre-existing relationship schema exerts a greater influence on representations when evaluating ambiguous, as opposed to clear, stimuli (e.g., Baldwin, 1992).

Though Study 1B rules out the possibility that these results reflect a general unwillingness of insecurely attached individual to discuss social emotional needs, it is not clear from these descriptions whether attachment security is biasing the way participants are attending to and representing the interaction or simply the way participants are discussing the interaction. Study 3 uses eye-tracking methodology to determine the extent to which the differential discussion of social-emotional goals observed in Study 2 is driven by differences in underlying attention and representation.

Study 3

Study 3 presented participants with the same stimuli as Study 2, but instead of having them provide a written description, we presented two outcomes intended to represent the successful completion of either the hill or social goal. Because infants have limited verbal abilities, methodologies for assessing mental representations that do not require verbal responses have become an invaluable tool to developmental psychologists (see Oakes, 2010, for a comprehensive review of this methodology). Though visual attention varies greatly across the lifespan (Colombo, 2001), gaze duration has previously been used in adult populations to examine attention to, and expectations of, similarly social stimuli (e.g., Guastella et al., 2008). Further, although it is less common to utilize looking time methodologies to assess the social cognitive representations of adults, doing so allows for a more direct comparison to the developmental literature that motivated the current research. Following the logic of infant looking time designs (e.g., Spelke, 1985), we expect that participants who have an expectation regarding the ball's goal will show greater attention to, and spend more time looking at, the outcome they find relatively unexpected.

Method
Participants
Two-hundred and twenty-nine undergraduate students (126 female) received partial course credit for participation. Two

additional students participated in the study but failed to complete the ECR and were removed from subsequent analysis. Participants whose Tobii capture rate was less than 75% were also excluded from further analysis. The final sample included 192 participants (103 females).

Measures/Procedure
Participants were 5-point calibrated on a Tobii T60 XL eye tracker. Once calibrated, participants watched the complex video from Study 2. Participants then saw a flashing central fixation point followed by two static outcomes presented simultaneously (**Figure 1D**). Because the complex video affords two accurate goal attributions (i.e., hill *and* social), we created two resolution scenes that dissociated these two outcomes. In both scenes the large ball was moved to the bottom of the hill, however, the location of the small ball varied. In the *hill* outcome, the small ball was seated atop the hill, physically separated from the large ball. In contrast, in the *social* outcome, the small ball was at the bottom of the hill beside the large ball. areas of interest (AOIs) were created around both of the outcomes and total fixation duration during the first 5 s of presentation was analyzed. Direction of motion (left versus right) and location of outcome (left versus right) were counterbalanced between participants. Following the eye-tracking portion of the study participants completed the ECR.

Results and Discussion
Attachment Classification
Consistent with the previous studies, participants were split into a secure ($N = 67$, 34.9%, Female = 34) and insecure group ($N = 125$, 65.1%, Female = 68). The securely attached group had significantly lower anxiety and avoidance than the insecure group [anxiety, $t(190) = 9.26$, $p < 0.001$; avoidance, $t(190) = 9.60$, $p < 0.001$].

No main effects or interactions of gender were observed; thus it was removed from subsequent analyses. A 2 (outcome: hill, mom) $\times$ 2 (security: secure, insecure) $\times$ 2 (motion: left, right) $\times$ 2 (location: left hill, left mom) mixed model analysis of variance was conducted to determine if the two groups of participants differentially attended to the two outcomes. There was a significant main effect of outcome [$F(1,184) = 13.47$, $p < 0.001$, $\eta_p^2 = 0.07$], and location [$F(1,184) = 5.286$, $p < 0.023$, $\eta_p^2 = 0.03$], and an interaction between outcome, motion, and location [$F(1,184) = 22.75$, $p < 0.001$, $\eta_p^2 = 0.11$].

Of particular interest to our research question was the effect of attachment security on attention to the two outcomes. As predicted by an attentional bias account, we found a significant interaction between security and outcome [$F(1,184) = 6.795$, $p < 0.01$, $\eta_p^2 = 0.04$; **Figure 3**]. Securely attached participants spent significantly more time looking at the hill outcome ($M = 1.89$, SD = 0.77) than the social outcome ($M = 1.42$, SD = 0.63) whereas, the insecurely attached participants looked equally long at both the hill ($M = 1.72$, SD = 0.58) and social outcomes ($M = 1.65$, SD = 0.57). Finally, security and outcome interacted with location [$F(1,184) = 6.22$, $p < 0.01$, $\eta_p^2 = 0.03$] such that, securely attached participants showed a main effect of outcome [$F(1,65) = 10.47$, $p = 0.002$, $\eta_p^2 = 0.14$] regardless of location
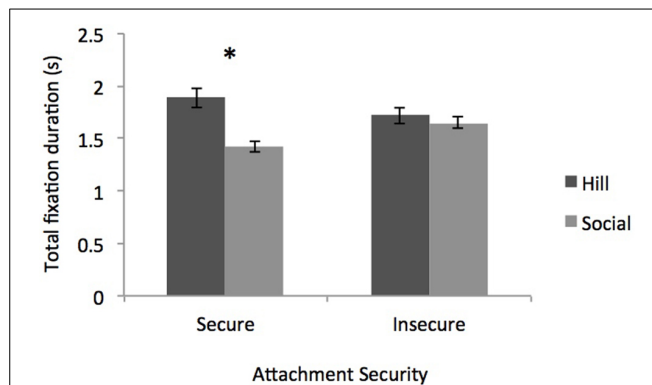
FIGURE 3 | Participant's total fixation duration to the instrumental (hill) versus social (reunion) outcome by attachment security.
The "*" indicates the difference is significant at $p < 0.05$.

$[F(1,65) = 3.12, p < 0.08, \eta_p^2 = 0.05]$. In contrast, insecurely attached participants showed a marginal interaction between outcome and location $[F(1,123) = 3.75, p < 0.06, \eta_p^2 = 0.03]$ but no main effect of outcome $[F(1,123) = 0.92, p = 0.34, \eta_p^2 = 0.007][4]$. Simply put, though securely attached participants tended to look longer to the hill outcome regardless of where it was located, insecurely attached participants tended to look longer at whichever outcome that was closer to the ball's initial movement, regardless of its content. A Wilcoxon Signed-rank test revealed a significant effect of attachment security on preferred outcome ($z = -2.12, p = 0.034$), suggesting that the pattern of results is not simply a function of averaging but instead holds across individual group members.

Taken together, this pattern of results suggests that securely attached participants have automatic, robust implicit expectations regarding the ball's ultimate goal that supersede task demands. Consistent with, though stronger than, the verbal responses observed in Study 2, securely attached participants appear to represent the hill as a means of achieving the overarching social goal. Securely attached participants devoted more attention (operationalized as longer looking) to the hill outcome, where the ball has successfully overcome a physical barrier but remains separated from its social partner, than to the social outcome, where the ball remains at the bottom of the hill but is reunited with its social partner, suggesting that participants found the instrumental outcome less expected and requiring of more attention. In contrast, insecurely attached participants did not differentiate their attention to either of the outcomes suggesting that they are either not automatically prioritizing social over instrumental goals, or are doing so in a manner that is weak and quickly overwhelmed by the surface characteristics of the stimuli. Insecurely attached participants did not appear to have a strong expectation regarding either resolution. Together with Study 2, these results support the proposal that variability in attachment security can influence the way we represent others' goals. When

---

[4]The pattern of results largely replicates when attachment insecurity is separated into three groups, however the interaction between ECR and outcome falls from significant to trending $[F(1,176) = 2.40, p < 0.07, \eta_p^2 = 0.04]$.

participants process complex social interactions that afford a number of different construals, the ease with which an individual approaches and interacts with their social environment can bias the representation of social-emotional goals particularly when the social goals are ambiguous and paired with a less emotionally evocative instrumental goal.

General Discussion

The overarching goal of the present research was to begin to address the question of why some early social reasoning appears universal, while some shows marked individual differences. Specifically, using both free-response and eye-tracking methodologies, we attempted to bridge two related domains of literature examining attachment security and other-oriented behavior in order to determine if this apparent contradiction could serve as a starting point for future research.

Across a series of three studies we demonstrate that the individual difference variable of attachment security affects the representation of instrumental needs differently than social-emotional distress (Studies 2 and 3). However, this was only the case when the stimuli were complex and afforded multiple potential interpretations (Studies 1A, B). Together these results suggest that attempting to understand and integrate divergent findings within a single theoretical framework can lead to more nuanced understanding.

Though these studies approach the question of social reasoning from a novel perspective, the findings are largely consistent with existing literature. As predicted by attachment theory we observed an influence of attachment security on the representation of social-emotional stimuli (Dykas and Cassidy, 2011), particularly when the stimuli were complex and afforded multiple construals (Baldwin, 1992). In addition, these findings are consistent with a growing body of literature examining the social cognitive constraints on early other-oriented behaviors; particularly that the ability to recognize and respond to instrumental needs emerges prior to, and independent from, the ability to respond to emotional distress (see Dunfield, 2014, for a review). Further, these results may help to explain why the ability to provide instrumental help appears more robust, and earlier emerging, than the ability to offer social help (Beier et al., 2014). Finally, these results are consistent with the finding that infants appear to universally evaluate helpers positively and hinderers negatively (e.g., Kuhlmeier et al., 2003; Hamlin et al., 2007), while showing individual differences in their expectations of responsive versus unresponsive caregivers (Johnson et al., 2007, 2010). Indeed, by considering both the underlying task demands and bodies of related research, we can gain insight and support for the perspective that attributing instrumental goals to agents acting on objects requires different underlying representations than attributing social emotional goals to agents acting on other agents (e.g., Spelke, 2014).

By taking a broad approach to social-cognitive development, and attempting to integrate a diversity of findings into a single theoretical account, we demonstrate an important role for examining how an individual difference variable, such as attachment security, can influence *both* similarities and

differences in social reasoning across individuals. Importantly, this work represents a first step toward integrating two approaches that have previously been largely pursued independently. While inspired by developmental theory and research, these studies examined university undergraduates, leaving open the question of when and how early social experiences influence cognitive development; however, based on existing findings (e.g., Johnson et al., 2007, 2010) it appears that the reciprocal relations emerge early.

Though these findings are consistent with previous research, the mechanisms underlying these differences are unclear. Based on these studies it is not currently possible to determine whether it is the content or complexity of the underlying representation that drives the observed differences. Future work is required to determine how these different goal construals interact with social evaluations in order to support behavioral outcomes. For example, it is possible that the infants in the initial caregiver studies (Johnson et al., 2007, 2010) were differentially evaluating the responsive versus unresponsive caregiver because they differed in the initial goal representation (i.e., some infants attended to the Baby's distress while others were attending to the Mommy's climb). It is also possible that the goals were construed similarly (i.e., reunion), however the participants differed, as proposed, in the type of caregiving responses they expected.

The goal of this paper was to examine more directly the apparent contradiction between research supporting universal similarities and individual differences in social reasoning. Although these two perspectives were often examined separately by researchers interested in either innate, early-emerging, universal components of cognition, or researchers interested in variable social outcomes affected by experience, there is a growing move to bring these two perspectives back together (Olson and Dweck, 2008, 2009). While this recombination may lead to the appearance of contradiction and inconsistency, we have demonstrated that by addressing the conflict head on, and using points of tension as starting points for further investigation, we can work toward a more nuanced and accurate understanding of the nature of human social cognition.

Acknowledgments

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., and Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale, NJ: Erlbaum.

Baldwin, M. W. (1992). Relational schema and the processing of social information. *Psychol. Bull.* 112, 461–484. doi: 10.1037/0033-2909.112.3.461

Beier, J. S., Over, H., and Carpenter, M. (2014). Young children help others to achieve their social goals. *Dev. Psychol.* 50, 934. doi: 10.1037/a0033886

Bowlby, J. (1982). *Attachment and Loss: Vol. 1. Attachment*. Middlesex: Pelican Books. [Original work published 1969].

Brennan, K. A., Clark, C. L., and Shaver, P. R. (1998). "Self-report measurement of adult romantic attachment: an integrative overview," in *Attachment Theory and Close Relationships*, eds J. A. Simpson and W. S. Rholes (New York: Guilford Press), 46–76.

Brewer, M. B., and Caporael, L. R. (2006). "An evolutionary perspective on social identity: Revisiting groups." In *Evolution and social psychology*, eds M. Schaller, J. A. Simpson, and D. T. Kendrick (New York: Psychology Press), 143–161.

Cassidy, J., and Shaver, P. R. (2008). *Handbook of Attachment, Second Edition: Theory, Research, and Clinical Applications*. New York: Oxford University Press.

Colombo, J. (2001). The development of visual attention in infancy. *Annu. Rev. Psychol.* 52, 337–367. doi: 10.1146/annurev.psych.52.1.337

De Wolff, M. S., and van Ijzendoorn, M. H. (1997). Sensitivity and attachment: a meta-analysis on parental antecedent of infant attachment. *Child Dev.* 68, 571–591. doi: 10.1111/j.1467-8624.1997.tb04218.x

Dunfield, K. A. (2014). A construct divided: prosocial behaviour as helping, sharing, and comforting subtypes. *Front. Psychol.* 5:958. doi: 10.3389/fpsyg.2014.00958

Dunfield, K. A., and Kuhlmeier, V. A. (2013). Classifying prosocial behavior: helping, sharing, and comforting subtypes. *Child Dev.* 84, 1766–1776. doi: 10.1111/cdev.12075

Dunfield, K. A., Kuhlmeier, V. A., O'Connell, L. J., and Kelley, E. A. (2011). Examining the diversity of prosocial behavior: helping, sharing, and comforting in infancy. *Infancy* 16, 227–247. doi: 10.1111/j.1532-7078.2010.00041.x

Dweck, C. S. (2013). "Social-cognitive development," in *Navigating the Social World: What Infants, Children, and Other Species Teach Us*, eds M. R. Banaji and S. G. German (New York: Oxford University Press), 3–5.

Dykas, M. J., and Cassidy, J. (2011). Attachment and the processing of social information across the life span: theory and evidence. *Psychol. Bull.* 137, 19–46. doi: 10.1037/a0021367

Eisenberg, N., Spinrad, T. L., and Knafo, A. (2015). "Prosocial development," in *Handbook of Child Psychology*, 7th Edn, eds M. Lamb and C. Garcia-Coll, Vol. ed. R. M. Lerner (New York: Wiley), 610–656.

Fawcett, C., and Liszkowski, U. (2012). Infants anticipate others' social preferences. *Infant. Child. Deve.* 21, 239–249.

Guastella, A. J., Mitchell, P. B., and Dadds, M. R. (2008). Oxytocin increases gaze to the eye region of human faces. *Biol. Psychiatry* 63, 3–5. doi: 10.1016/j.biopsych.2007.06.026

Hamlin, J. K. (2014). The case for social evaluation in preverbal infants: gazing toward one's goal drives infants' preferences for Helpers over Hinderers in the hill paradigm. *Front. Psychol.* 5:1563. doi: 10.3389/fpsyg.2014.01563

Hamlin, J. K. (2013). Moral judgment and action in preverbal infants and toddlers: evidence for an innate moral core. *Curr. Dir. Psychol. Sci.* 22, 186–193. doi: 10.1177/0963721412470687

Hamlin, J. K., Wynn, K., and Bloom, P. (2007). Social evaluation by preverbal infants. *Nature* 450, 557–559. doi: 10.1038/nature06288

Hobbs, K., and Spelke, E. (2015). Goal attributions and instrumental helping at 14 and 24 months of age. *Cognition* 142, 44–59. doi: 10.1016/j.cognition.2015.03.014

Johnson, S. C., and Chen, F. S. (2011). Socioemotional information processing in human infants: from genes to subjective construals. *Emot. Rev.* 3, 169–178. doi: 10.1177/1754073910387945

Johnson, S. C., Dweck, C., and Chen, F. S. (2007). Evidence for infants' internal working model of attachment. *Psychol. Sci.* 18, 501–502. doi: 10.1111/j.1467-9280.2007.01929.x

Johnson, S. C., Dweck, C., Chen, F. S., Ok, S. J., Stern, H. L., and Barth, M. E. (2010). At the intersection of social and cognitive development: internal working models of attachment in infancy. *Cogn. Sci.* 34, 807–825. doi: 10.1111/j.1551-6709.2010.01112.x

Johnson, S. C., Dweck, C. S., and Dunfield, K. A. (2013). "How universals and individual differences can inform each other: the case for social expectations in infancy," in *Navigating the Social World: What Infants, Children, and Other Species Teach Us*, eds M. R. Banaji and S. G. German (New York: Oxford University Press), 44–48.

Kuhlmeier, V., Wynn, K., and Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychol. Sci.* 14, 402–408. doi: 10.1111/1467-9280.01454

Main, M., Kaplan, N., and Cassidy, J. (1985). Security in infancy, childhood, and adulthood: a move to the level of representation. *Monogr. Soc. Res. Child Dev.* 50, 66–104. doi: 10.2307/3333827

Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *J. Cogn. Dev.* 11, 255–268. doi: 10.1080/15248371003699977

Olson, K. R., and Dweck, C. S. (2008). A blueprint for social cognitive development. *Perspect. Psychol. Sci.* 3, 193–202. doi: 10.1111/j.1745-6924.2008.00074.x

Olson, K. R., and Dweck, C. S. (2009). Social cognitive development: a new look. *Child Dev. Perspect.* 3, 60–65. doi: 10.1111/j.1750-8606.2008.00078.x

Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B., and Meinhardt, J. (2013). Neural correlates of prosocial behavior in infancy: different neurophysiological mechanisms support the emergence of helping and comforting. *Neuroimage* 66, 522–530. doi: 10.1016/j.neuroimage.2012.10.041

Piaget, J. (1995). *Logical Operations and Social Life*. New York: Routledge. [Original work published in 1945].

Premack, D., and Premack, A. J. (1997). Infants attribute value ± to the goal-directed actions of self-propelled objects. *J. Cogn. Neurosci.* 9, 848–856. doi: 10.1162/jocn.1997.9.6.848

Spelke, E. S. (1985). "Preferential looking methods as tools for the study of cognition in infancy," in *Measurement of Audition and Vision in the First Year of Postnatal Life*, eds G. Gottlieb and N. Krasnegor (Norwood, NJ: Ablex), 323–363.

Spelke, E. S. (2014). Core social cognition. *Paper Presented at the Concordia Science College Public Lecture Series*, Montreal, QC.

Svetlova, M., Nichols, S. R., and Brownell, C. A. (2010). Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Dev.* 81, 1814–1827. doi: 10.1111/j.1467-8624.2010.01512.x

Van Overwalle, F., and Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48, 564–584. doi: 10.1016/j.neuroimage.2009.06.009

Vygotsky, L. (1978). "Mind in society: the development of higher psychological processes," eds M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Cambridge, MA: Harvard University Press).

Warneken, F., and Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy* 11, 271–294. doi: 10.1111/j.1532-7078.2007.tb00227.x

Warneken, F., and Tomasello, M. (2009). Varieties of altruism in children and chimpanzees. *Trends Cogn. Sci.* 13, 397–402. doi: 10.1016/j.tics.2009.06.008

Woodward, A. L., Sommerville, J. A., and Guajardo, J. J. (2001). "How infants make sense of intentional action," in *Intentions and Intentionality: Foundations of Social Cognition*, eds B. F. Malle, L. J. Moses, and D. A. Baldwin (Cambridge, MA: MIT Press), 149–169.

Yamaguchi, M., Kuhlmeier, V. A., Wynn, K., and VanMarle, K. (2009). Continuity in social cognition from infancy to childhood. *Dev. Sci.* 12, 746–752. doi: 10.1111/j.1467-7687.2008.00813.x

Gendered race: are infants' face preferences guided by intersectionality of sex and race?

Hojin I. Kim, Kerri L. Johnson and Scott P. Johnson*

Department of Psychology, UCLA Baby Lab, University of California, Los Angeles, Los Angeles, CA, USA

People occupy multiple social categories simultaneously (e.g., a White female), and this complex intersectionality affects fundamental aspects of social perception. Here, we examined the possibility that infant face processing may be susceptible to effects of intersectionality of sex and race. Three- and 10-month-old infants were shown a series of computer-generated face pairs (5 s each) that differed according to sex (Female or Male) or race (Asian, Black, or White). All possible combinations of face pairs were tested, and preferences were recorded with an eye tracker. Infants showed preferences for more feminine faces only when they were White, but we found no evidence that White or Asian faces were preferred even though they are relatively more feminized. These findings challenge the notions that infants' social categories are processed independently of one another and that infants' preferences for sex or race can be explained from mere exposure.

Keywords: sex and race categorization, infant face preference, social development, social cognitive development, sex differences

Introduction

Studies of infant face preference represent an important opportunity to inform theories of social cognitive development, in particular the means by which infants determine critical features of social categories (Ramsey et al., 2005) and the means by which social context influences recognition of individuals from specific groups (Scott et al., 2007). Moreover, understanding infant face preferences may help reveal the developmental origins of stereotypes and prejudice (given that preverbal infants lack direct knowledge of group characteristics), if these origins are at least partly perceptual in nature. Contemporary theories of face processing appeal to results from studies of prototype formation, prototype preference, and intermodal matching in providing important evidence for an asymmetry in the development of infants' categorical knowledge of female and male faces, such that knowledge of the female category becomes relatively more advanced early in development than the male category (Ramsey et al., 2005). Performance in these studies requires formation of prototypes from prior experiences with multiple exemplars of faces before a preference for a prototypic (i.e., attractive) exemplar could be observed. Importantly, it appears that infants have difficulty forming prototypes of male faces and matching a series of male faces and voices together in laboratory setting; performance in prototype-formation and face-matching studies is facilitated by the use of female faces.

In addition, infants take more time to process male than female faces. These effects may stem from a relative lack of experience with male faces, given that most infants spend more time with females (e.g., their mothers), and they show perceptual preferences for female faces (Quinn et al., 2002). At the same time, male faces may be more variable in their features and in spacing between features (Hopper et al., 2014), and if the range of perceptual differences among category

members is relatively wide, infant categorization is impaired (Mareschal et al., 2000). This too may make it more difficult for infants to categorize male faces. Thus differential experience and variability may systematically affect infant categorization of male and female faces. Infants may first learn to recognize and discriminate the mother's face from other female faces, and subsequently these discrimination abilities may extend to other female faces. As a result of experience with various faces, infants should begin to form a representation of faces and a rudimentary category for faces. This representation should be most heavily weighted on the mother's face and therefore specific to the human species, most representative of the mother's race, and primarily female-like, so that it guides infants' attention toward other, similar faces.

Similar effects may be operational in face race processing, which may function in accord with the contact hypothesis of social perception (Sporer, 2001). This hypothesis argues that contact with individuals from specific social groups fosters the ability to extract visual cues or invoke processing strategies that support recognition of individuals within these groups. Exposure to faces within one's own race compared to faces of other races, for example, may lead to less practice recognizing other-race faces (cf. Johnson, 2010). Research on face categorization in infancy is consistent with this possibility. Nine-month-olds categorized faces from own- and other-races (White and Asian, respectively), yet appeared to recognize only own-race individuals (Anzures et al., 2010). This so-called "other-race effect" in recognition has been attributed to our differential experience with different categories of faces (Kelly et al., 2007).

Studies of infant and adult processing of sex and race of faces generally isolate a single social category while holding other categories constant (e.g., manipulating race while holding sex constant; e.g., Kelly et al., 2007). Research that investigates face processing when identities intersect (e.g., when a target is both Asian and male) remains relatively rare in the adult literature, and to our knowledge this issue remains largely unexplored with infants. Recent attempts to reach a more nuanced understanding of these complexities are noteworthy, and suggest that the perception of various social categories may be interdependent. For example, Quinn et al. (2008) reported that 3-month-old White infants exhibited a preference for female faces only when the faces were White, but not Asian. That is, race category may bias sex categorizations, and *vice versa*, due to common facial cues to which infants may be sensitive. (For older individuals, overlapping cognitive stereotypes may also bias race and sex categorization, but we do not expect such effects in infants.)

Such research is vital for a full account of face categorization because in the real world, people occupy multiple social categories simultaneously, and this complex intersectionality affects fundamental aspects of social perception (to be distinguished from the sociological use of the term; e.g., Browne and Misra, 2003). For example, Black men and women were judged by (predominantly White) adults as more masculine and race stereotypical than same-sex White targets. In addition, sex categorization errors, although rare, were more common for Black women than any other race/sex combination (Goff et al., 2008). There are effects of intersectionality involving emotion

as well: The perceived onset and duration of happiness and anger appears to depend on both race and age. Specifically, adult observers detected anger earlier and judged it to endure longer for younger, relative to older, Black men, and observers judged happiness to disappear earlier and to be shorter lived for younger Black men. The opposite occurred for perceptions of White men (Kang and Chasteen, 2009). In addition, neutral male faces are perceived as relatively more angry than female faces, neutral White faces resemble angry expressions more than do Black or Asian faces, and neutral Black faces resemble happy expressions more than do White faces (Zebrowitz et al., 2010).

The effects of intersectionality also impact the efficiency of social categorization. Recently, Johnson et al. (2012) tested the possibility that face race will bias sex categorization through common cues and/or overlapping stereotypes, both leading to similar predictions. They theorized that a race category associated with phenotypes or stereotypes that align with the target's sex category membership should facilitate sex categorization. This is because, for instance, Asian faces share phenotypes and/or activate stereotypes that are also common to women, and Black faces share phenotypes and/or activate stereotypes that are also common to men. A race category associated with phenotypes or stereotypes that are at odds with the target's sex category membership, in contrast, should impair sex categorization. Specifically, male categorizations were predicted to be more efficient for Black faces, but less efficient for Asian faces, relative to White faces, and female categorizations were predicted to be less efficient for Black faces, but more efficient for Asian faces, relative to White faces. These predictions were supported in an experiment in which adult participants categorized the sex of computer-generated Asian, Black, or White faces. In a second study, the computer program employed to create these faces (FaceGen Modeler) was used to quantify the degree to which sex-typed cues covaried with race in 166 photographs of real faces. This analysis revealed that Black faces were overall more masculinized in appearance relative to Asian and White faces. Additional experiments tested implicit stereotypes held by adult observers, and confirmed that Blacks were considered to have more stereotypically male attributes (e.g., aggressive, assertive, dominant) and Asians were considered to have more stereotypically female attributes (e.g., considerate, dependent, modest). More recently, experiments have confirmed that the reciprocal relation is also true: Black categorizations were facilitated for male/masculine faces but White and Asian categorizations were facilitated for female/feminine faces (Carpinella et al., 2015). Thus sex-race intersectionality in these studies was found to operate in both a "bottom-up" (from facial characteristics) and a "top-down" (from stereotyped attributes) fashion.

The present study examines the possibility that infant face perception, likewise, is susceptible to intersectionality of sex and race. We reasoned that the tendency for infants to prefer female faces could be leveraged to examine the extent to which different face races comprise facial features that are relatively more feminine, and we tested an age range spanning important developments in race categorization to better understand how

infants' emerging sensitivity to characteristics of own- and other-races may alter such preferences. Notably, given that we test preverbal populations, the effects we report necessarily operate in the absence of stereotypes. We presented 3- and 10-month-old infants a series of face pairs in which one member of the pair was presumed to appear more feminine, and we predicted that infants would generally prefer the more feminized face. We tested this prediction in two ways. First, we manipulated facial features in FaceGen Modeler to appear explicitly female, androgynous, or male, and tested infants' preferences for the more feminized face in pairs of Asian, Black, or White faces. Second, we manipulated face race in androgynous faces (Asian, Black, or White), and examined preferences for the face from the more "feminized" race in each pairing (Asian → White → Black). Finally, we presented blank faces (featureless ovals) which were the same average color as the androgynous faces of each race to test for inherent color preferences, to address the possibility that the hypothesized female preference might be confounded by the color of the face. Three- and 10-month-olds were chosen for observation because these age groups bracket important developments in, for example, the other-race effect (Kelly et al., 2007) and face recognition (Nelson, 2001). We reasoned that effects of intersectionality in face preference likewise might be experience-dependent, such that these effects would be stronger in the older infants. Ten-month-olds have substantially more exposure to faces than 3-month-olds, and thus more exposure to social categories and their overlapping features. We also examined differences in performance between infants from different racial groups, and effects of race of the primary caregiver, which was the mother for all infants we observed in this study.

Materials and Methods

Participants

Thirty-two 3-month-olds (17 boys, 15 girls; $M = 3.2$ months, $SD = 0.28$) and thirty-two 10-month-olds (16 boys, 16 girls; $M = 10.0$ months, $SD = 0.27$) composed the final sample. All infants were full term and had no known developmental difficulties. Infants were recruited from lists of birth records provided by Los Angeles County. Parents were contacted by letter and telephone, and were provided with a small thank-you gift (a toy or a T-shirt with the lab logo) for participation. An additional 34 infants were observed but excluded due to excessive fussiness or inattention (six 3-month-olds), eye-tracking calibration failures (twenty 3-month-olds, four 10-month-olds), or inability of the eye tracker to consistently track the point of gaze (one 3-month-old, three 10-month-olds).

Materials

A total of nine computer-generated face stimuli were created using commercial software (FaceGen Modeler) with three levels of face race (Asian, Black, or White) and three levels of face femininity (female, androgynous, or male; see **Figure 1**). To produce these stimuli, we first created an average androgynous White face using the Random Generation Feature of the software and by setting the gender level at the center of the 80-point femininity-masculinity scale. Using the same gender scale, we then created comparable

FIGURE 1 | Stimuli used in the present experiment. Top–bottom rows: Asian, Black, and White faces. Left–right columns: female, androgynous, male, and blank faces.
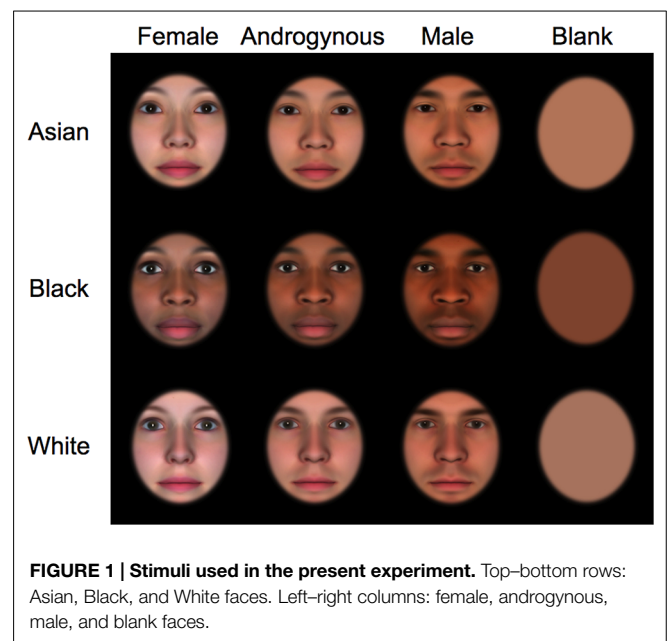
male and female White faces, setting the scale at 60 and 20, respectively. Subsequently, Black and Asian counterparts were created by systematically manipulating the apparent race of each face using the Race Morphing Control feature of the software. Once all nine faces were generated, we used Adobe Photoshop to edit the faces. An oval-shaped outline was superimposed on each of the nine faces to expose only the internal facial features. This was necessary because we wanted to minimize the effect of external facial features on infants' preference for a particular gender or race.

Adobe Photoshop was used to construct the final set of the stimuli, a series of side-by-side comparisons. Each visual stimulus measured 25 cm × 22.5 cm (23.5° × 21.2° visual angle) and was separated by a gap of 1.5 cm (1.4°). Each face measured approximately 14 cm × 10.5 cm (13.3° × 10.0°). For each of the three face races (i.e., White, Black, Asian), three within-race gender comparison trials were constructed (i.e., female vs. male, female vs. androgynous, and androgynous vs. male) to test infants' preference for the more feminine faces in each comparison for all three races. In addition, three between-race comparison trials were constructed using only the androgynous faces (i.e., White vs. Black, White vs. Asian, Black vs. Asian) to examine infants' preference for a particular face race while minimizing the potential effect of the female face preference. Furthermore, we created three additional between-race comparison trials (i.e., White vs. Black, White vs. Asian, Black vs. Asian) using blank faces (i.e., colored ovals) to examine the effect of skin tones on infants' preference for a particular race, as noted previously. To perform such a comparison, a total of three additional blank faces were created (one per face race). The blank faces contained no facial features, and represented the average skin tone of the androgynous face.

The final stimuli of the present study consisted of two blocks of 15 side-by-side trials: nine within-race gender comparisons, three between-race comparisons using androgynous faces, and

three between-race comparisons using blank faces. The left-right presentation of the faces was counterbalanced by presenting two blocks of identical trials. The second block of the trials consisted of mirror images of those in the first block. Each trial lasted for 5 s.

Procedure

Research protocols were approved by the UCLA Institutional Review Board. Prior to testing, parents filled out a demographic questionnaire that requested information about the primary and secondary caregivers' race. The primary caregiver for all infants in our sample was the mother. There were 33 self-identified White mothers, 8 Asians, 18 Hispanic/Latina, 2 Black, and 2 Middle Eastern. Twenty-five of the infants were categorized as White (two white parents), and 39 as mixed-race.

Each infant was observed while seated on his or her parent's lap approximately 60 cm from a 24-inch TFT widescreen monitor (resolution set at 1900 × 1200 pixels) surrounded by black curtains to minimize distractions. Eye movements were recorded with a Tobii T60-XL eye tracker at 60 Hz with a spatial accuracy of approximately 0.5°–1°. The lights in the experimental room were dimmed and the only source of illumination came from the monitor.

Prior to their participation in the study, infants' point of gaze was calibrated by repeated presentations of a dynamic target-patterned ball undergoing continuous contraction and expansion. The calibration stimulus was presented briefly at each of five locations on the monitor (the four corners and the center) while infants tracked it with their eyes. The Tobii eye tracker provides information about calibration quality for each point; if there were no data for one or more points or if calibration quality was poor, calibration at those points was repeated. Calibration was followed immediately by presentation of faces as described previously. Prior to each trial a small audiovisual attention-getting stimulus was shown to reorient infants' attention to the center of the monitor.

Results

The goal of our first set of analyses was to establish the extent to which intersectionality of sex and race characteristics influences infant preferences. We predicted greater looking toward the more feminine face in each face pairing, which we operationalized as dwell times (accumulated fixations as recorded by the eye tracker) to each of the two faces; the dependent variable for these analyses, therefore, was "femininity preference." Our principal questions were first, whether the hypothesized female face preference would be modulated by the race of the face (i.e., Asian, Black, or White), and second, whether the female preference would be modulated by the comparisons represented by each pairing (i.e., female–male pairings, female–androgynous pairing, and androgynous–male pairings). We also examined age differences in performance to assess the possibility that infants' preferences and potential intersectionality effects may emerge in parallel with other key face processing skills (Nelson, 2001; Kelly et al., 2007). We computed a 3 (Comparison: female–male, female–androgynous, or androgynous–male) × 3 (Face Race:

FIGURE 2 | Dwell times toward feminine vs. masculine faces in Asian, Black, and White face pairs. Error bars = SEM.
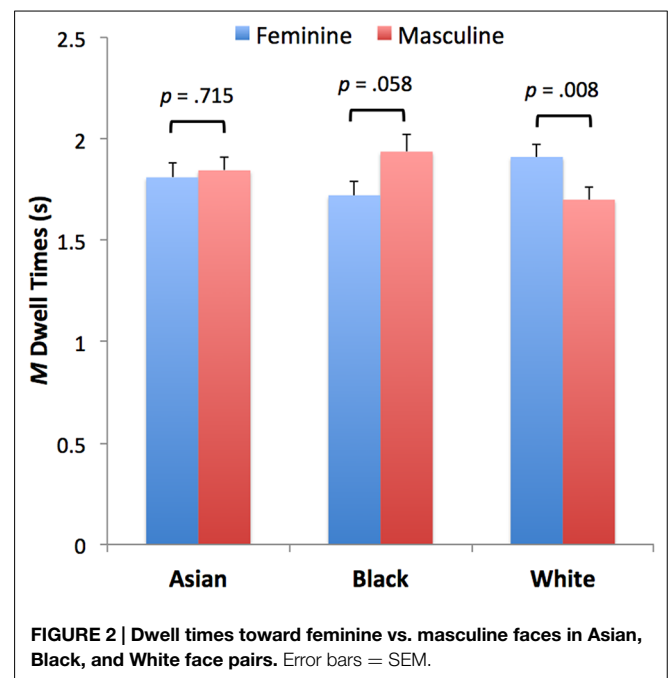
Asian, Black, or White) × 2 (Age Group) × 2 ("Femininity" Preference in each pairing) mixed ANOVA with repeated measures on the last factor. This analysis yielded a statistically significant main effect of Comparison, $F_{(2,124)} = 3.20$, $p = 0.044$, $\eta_p^2 = 0.049$, stemming from longer overall looking at female–male than at androgynous–male pairings and somewhat more at female–androgynous than at androgynous–male pairings (the reasons for these effects are unclear). More importantly, we found a significant Face Race × Femininity Preference interaction, $F_{(2,124)} = 5.06$, $p = 0.008$, $\eta_p^2 = 0.075$ (see **Figure 2**). Tests for simple effects revealed that infants looked longer at the more feminine face when faces were White, $F_{(1,63)} = 7.61$, $p = 0.008$, but not when faces were Asian, $F_{(1,63)} = 0.13$, $p = 0.715$. When faces were Black, in contrast, there was a trend toward a *male* preference, $F_{(1,63)} = 3.72$, $p = 0.058$. Additional simple effects tests revealed that the preference for feminine faces was reliably stronger for White vs. Black faces, $F_{(1,62)} = 8.80$, $p < 0.01$, and marginally stronger for White vs. Asian faces, $F_{(1,62)} = 3.47$, $p = 0.067$. There were no other significant main effects or interactions. These data, therefore, demonstrate that the female face preference reported in earlier studies (e.g., Quinn et al., 2002) is contingent on face race, having been observed under tested conditions only when infants viewed White faces (cf. Quinn et al., 2008).

The next analysis examined the possibility that infants perceive Asian, Black, and White race faces to be gendered, as has been reported for different race face morphologies in photographs of real faces (Johnson et al., 2012). If so, we predicted that, when gender cues have been equated except for face race (i.e., in androgynous faces), faces that are relatively more feminized would be preferred by virtue of the hypothesized intersectionality of face race and gender. Specifically, we predicted that Asian androgynous faces will be preferred to White and Black, and White androgynous faces will be preferred to Black because faces

FIGURE 3 | Correlations between *M* color preference (i.e., preference for darker colors) and preference for feminized faces in comparisons of female vs. male (red), female vs. androgynous (blue), and androgynous vs. male (green). Corresponding colored trend lines are superimposed.
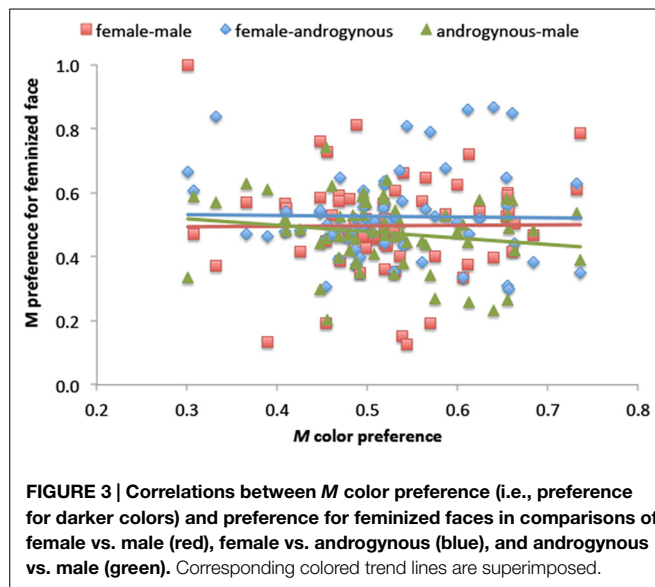
may be gendered by race rather than by direct manipulation of sex-typed facial features within FaceGen Modeler. A 3 (Comparison: Asian vs. White, Black vs. Asian, Black vs. White) × 2 (Femininity Preference in each pairing) × 2 (Age Group) mixed ANOVA revealed a statistically significant main effect of Age Group, $F(1,62) = 6.27$, $p = 0.015$, $\eta^2_p = 0.092$, due to overall longer dwell times by 10-month-olds ($M = 11.64$ s, SD = 2.51) vs. 3-month-olds ($M = 10.18$ s, SD = 2.15). There were no other significant main effects or interactions. These analyses provide evidence against the likelihood that different race faces appear differently gendered to infant observers.

A third set of analyses examined the possibility that differences in skin color between Asian, Black, and White faces may have influenced infants' preferences. To achieve this goal, we examined preferences for the darker blank face in Asian–White, Black–Asian, and Black–White pairings with a 3 (Comparison) × 2 (Skin Tone Preference: darker vs. lighter face) × 2 (Age Group) mixed ANOVA, which revealed a statistically significant main effect of Skin Tone Preference, $F(1,62) = 4.66$, $p = 0.035$, $\eta^2_p = 0.070$, the result of longer looking overall at darker faces ($M = 1.55$ s, SD = 0.50) relative to lighter faces ($M = 1.42$ s, SD = 0.40), and a main effect of Age Group, $F(1,62) = 8.55$, $p = 0.005$, $\eta^2_p = 0.121$, due to overall longer dwell times by 10-month-olds ($M = 9.81$ s, SD = 2.26) vs. 3-month-olds ($M = 8.02$ s, SD = 2.63). There was also a reliable three-way interaction, $F(2,124) = 6.45$, $p = 0.002$, $\eta^2_p = 0.094$, stemming from somewhat stronger dark preferences by 3-month-olds viewing the Black–Asian comparison and by 10-month-olds viewing the Black–White comparison. In addition, we used correlation analyses to examine relations between the *M* color preference for individual infants and femininity preferences in face pairs (female–male, female–androgynous, and androgynous–male) in light of our previously described results showing that infants' female face preference is modulated by face race. These analyses revealed no statistically reliable effects, ps > 0.172 (see **Figure 3**). Taken together, these analyses reveal that infants tended to prefer darker colors, but this preference did

not interact with face race, and there was no consistent way in which color preference *per se* was related to female preference.

Finally, we examined differences in face preference as a function of the mother's and the infant's race with a series of Bonferroni-corrected *t*-tests. There were no statistically significant differences in any of the preferences we reported previously in this section between infants of White mothers ($n = 33$) vs. infants of Asian, Hispanic, Black, or Middle Eastern mothers ($n = 31$), nor were there any reliable differences in preference between infants from White ($n = 25$) vs. mixed-race ($n = 39$) families.

Discussion

We examined the hypothesis that facial features specifying race and gender may overlap to the extent that infants perceive faces to be gendered (as do adults; Johnson et al., 2012) by capitalizing on the previously-reported tendency of infants to prefer female faces (Quinn et al., 2002, 2008). We tested infants' visual preferences for female vs. male in Asian, Black, and White computer-generated face pairs, we tested preferences for Asian vs. White, Asian vs. Black, and White vs. Black in pairs of androgynous faces, and we tested for preferences for oval patches of color that represented the average hue of Asian, Black, and White androgynous faces. Specifically, we tested the possibilities that (a) infants' purported female face preference would vary as a function of face race, and (b) that race faces are inherently gendered due to phenotypic overlap in facial features that are characteristic of sex differences. The first hypothesis, but not the second, was supported, and we interpret these two findings in turn.

Consider first the results of analyses of the female preference in different race face pairs. The female preference we predicted was observed in White face pairs, but not in Asian or Black face pairs. This result replicates and extends findings of Quinn et al. (2008), who discovered that White 3-month-olds preferred female faces only when the faces were White, but not Asian. Here, we found the same result across the sample, even among infants who were not White or who came from mixed-race families. Interestingly, we found also that the female preference is actually reversed to an extent when infants view Black faces. We showed also that 10-month-olds' visual preferences were not statistically different from those of 3-month-olds. This result implies that developments in the other-race effect (Kelly et al., 2007), which likely stem from growing experience with same-race faces during the first year after birth (cf. Scott et al., 2007) had little bearing on infants' behavior under tested circumstances; nor did races of household members seem to matter, in contrast to the findings reported by Quinn et al. (2002).

It may be that infants did not exhibit the female preference in Asian face pairs because sexual dimorphism in Asian faces is reduced relative to White faces—that is, the differences between female and male facial features is greater in Whites. Hopper et al. (2014) used multidimensional (MDS) scaling to place 40 photographs of Asian and White women and men (10 photos each) into a "face space," so that different facial attributes (dimensions of facial features and distances between features) corresponded to distinct dimensions within the MDS scaling space. Gender was found to vary more for White faces, resulting

in a negative or positive correlation between gender and race when only considering male or only considering female faces. Female and male Asian faces, therefore, are relatively more similar in appearance, and this may mean it is somewhat less likely that infants can discriminate female from male, or that they are not sufficiently distinct in appearance such that females attract more attention. It may be that infants did not exhibit the female preference in Black face pairs because of superficial similarities between the characteristics of neutral male and neutral Black faces to happy expressions in general, as revealed by outputs of connectionist models trained to recognize facial metrics of angry, happy, and surprise expressions in White male and female faces (Zebrowitz et al., 2010). If phenotypic characteristics of Black faces (in particular, Black male faces) overlap with positive expressions, which are known to attract infants' attention relative to other emotions (e.g., Kim and Johnson, 2013, 2014), then a reduction in female preferences in Black faces (indeed, nearly to the point of statistical significance in the other direction) may stem from a latent tendency for Black faces to convey positive emotions, even though the computer-generated faces used in the present experiment were explicitly neutral with respect to emotional expression. It remains for future research to examine more carefully the possibility of intersectionality of race and emotion in infant face perception.

Consider next our second principal question in the present study, the possibility that face race is inherently gendered, again due to purported overlap in facial features that convey information for attributes specifying race and sex. We found no evidence under tested circumstances that infants perceived Asian and White faces to be relatively more feminized, or Black faces to be relatively more masculinized, as has been reported from experiments with adult observers and from detailed measurements of facial features in photographs (Johnson et al., 2012). We observed no age differences between 3- and 10-month-olds in infant female preferences, nor did we observe differences in visual preferences as a function of infant race (White or non-White). It may be that the conditions we employed to test this question involved distinctions in face race that were too fine-grained to be detected by infants in androgynous faces, or it may be that this kind of race-gender overlap awaits developments in perceptual skills that occur beyond infancy. Notably, adult responses to race-gender intersectionality are highly sensitive to both bottom-up (feature overlap) and top-down (stereotypicality) effects, as observed with reaction time, mouse tracking, and verbal judgments (Johnson et al., 2012). Given the importance of the top-down effects that Johnson et al. (2012) reported, however, it is possible that sensitivity to some subtle facial cues supporting race and gender distinctions emerge in tandem with the cognitive representations that underlie stereotypes, in-group preferences, and racial and gender biases (cf. Hugenberg et al., 2010; Young et al., 2012). For example, it has been proposed that attributes that distinguish among social groups attain "psychological salience" in childhood (Bigler and Liben, 2007), and this may tune the visual system toward certain physical characteristics that then become perceptually salient (cf. Scott et al., 2007).

References

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., and Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' perception of face race. *Dev. Sci.* 13, 553–564. doi: 10.1111/j.1467-7687.2009.00900.x

Bigler, R. S., and Liben, L. S. (2007). Developmental intergroup theory: explaining and reducing children's social stereotyping and prejudice. *Curr. Dir. Psychol. Sci.* 16, 162–166.

Browne, I., and Misra, J. (2003). The intersection of gender and race in the labor market. *Annu. Rev. Sociol.* 29, 487–513. doi: 10.1146/annurev.soc.29.010202.100016

Carpinella, C. M., Chen, J., Hamilton, D., and Johnson, K. L. (2015). Gendered facial cues influence race categorizations. *Pers. Soc. Psychol. Bull.* 41, 405–419. doi: 10.1177/0146167214567153

Goff, P. A., Thomas, M. A., and Jackson, M. C. (2008). "Ain't I a woman?" Towards an intersectional approach to person perception and group-based harms. *Sex Roles* 59, 392–403. doi: 10.1007/s11199-008-9505-4

Hopper, W. J., Finklea, K. M., Winkielman, P., and Huber, D. E. (2014). Measuring sexual dimorphism with a race-gender face space. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1779–1788. doi: 10.1037/a0037743

Hugenberg, K., Young, S., Bernstein, M., and Sacco, D. F. (2010). The Categorization-Individuation Model: an integrative account of the cross race recognition deficit. *Psychol. Rev.* 117, 1168–1187. doi: 10.1037/a0020463

Johnson, K. L., Freeman, J. B., and Pauker, K. (2012). Race is gendered: how covarying phenotypes and stereotypes bias sex categorization. *J. Pers. Soc. Psychol.* 102, 116–131. doi: 10.1037/a0025335

Johnson, S. P. (2010). How infants learn about the visual world. *Cogn. Sci.* 34, 1158–1184. doi: 10.1111/j.1551-6709.2010.01127.x

Kang, S. K., and Chasteen, A. L. (2009). Beyond the double-jeopardy hypothesis: assessing emotion on the faces of multiply-categorizable targets of prejudice. *J. Exp. Soc. Psychol.* 45, 1281–1285. doi: 10.1016/j.jesp.2009.07.002

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., and Pascalis, O. (2007). The other-race effect develops during infancy. *Psychol. Sci.* 18, 1084–1089. doi: 10.1111/j.1467-9280.2007.02029.x

Kim, H. I., and Johnson, S. P. (2013). Do young infants prefer an infant-directed face or a happy face? *Int. J. Behav. Dev.* 37, 125–130. doi: 10.1177/0165025413475972

Kim, H. I., and Johnson, S. P. (2014). Detecting "infant-directedness" in face and voice. *Dev. Sci.* 14, 621–627. doi: 10.1111/desc.12146

Mareschal, D., French, R. M., and Quinn, P. C. (2000). A connectionist account of asymmetric category learning in early infancy. *Dev. Psychol.* 36, 635–645. doi: 10.1037/0012-1649.36.5.635

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant Child Dev.* 10, 3–18. doi: 10.1002/icd.239

Quinn, P. C., Uttley, L., Lee, K., Gibson, A., Smith, M., Slater, A. M., et al. (2008). Infant preference for female faces occurs for same-but not other-race faces. *J. Neuropsychol.* 2, 15–26. doi: 10.1348/174866407X231029

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., and Pascalis, O. (2002). Representation of the gender of human faces by infants: a preference for female. *Perception* 31, 1109–1121. doi: 10.1068/p3331

Ramsey, J. L., Langlois, J. H., and Marti, N. C. (2005). Infant categorization of faces: ladies first. *Dev. Rev.* 25, 212–246. doi: 10.1016/j.dr.2005.01.001

Scott, L. S., Pascalis, O., and Nelson, C. A. (2007). A domain general theory of the development of perceptual discrimination. *Curr. Dir. Psychol. Sci.* 16, 197–201. doi: 10.1111/j.1467-8721.2007.00503.x

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: an integration of theories. *Psychol. Public Policy Law* 7, 170–210.

Young, S. G., Hugenberg, K., Bernstein, M. J., and Sacco, D. F. (2012). Perception and motivation in face recognition: a critical review of theories of the cross-race effect. *Pers. Soc. Psychol. Rev.* 16, 116–142. doi: 10.1177/1088868311418987

Zebrowitz, L. A., Kikuci, M., and Fellous, J. (2010). Facial resemblance to emotions: group differences, impression effects, and race stereotypes. *J. Pers. Soc. Psychol.* 98, 175–189. doi: 10.1037/a0017990

Asian infants show preference for own-race but not other-race female faces: the role of infant caregiving arrangements

Shaoying Liu[1], Naiqi G. Xiao[2], Paul C. Quinn[3], Dandan Zhu[1], Liezhong Ge[1]*, Olivier Pascalis[4] and Kang Lee[2]*

[1] Zhejiang Sci-Tech University, Hangzhou, China, [2] University of Toronto, Toronto, ON, Canada, [3] University of Delaware, Newark, DE, USA, [4] Laboratoire de Psychologie et Neurocognition – Université Grenoble Alpes, Centre National de la Recherche Scientifique, Grenoble, France

Previous studies have reported that 3- to 4-month-olds show a visual preference for faces of the same gender as their primary caregiver (e.g., Quinn et al., 2002). In addition, this gender preference has been observed for own-race faces, but not for other-race faces (Quinn et al., 2008). However, most of the studies of face gender preference have focused on infants at 3–4 months. Development of gender preference in later infancy is still unclear. Moreover, all of these studies were conducted with Caucasian infants from Western countries. It is thus unknown whether a gender preference that is limited to own-race faces can be generalized to infants from other racial groups and different cultures with distinct caregiving practices. The current study investigated the face gender preferences of Asian infants presented with male versus female face pairs from Asian and Caucasian races at 3, 6, and 9 months and the role of caregiving arrangements in eliciting those preferences. The results showed an own-race female face preference in 3- and 6-month-olds, but not in 9-month-olds. Moreover, the downturn in the female face preference correlated with the cumulative male face experience obtained in caregiving practices. In contrast, no gender preference or correlation between gender preference and face experience was found for other-race Caucasian faces at any age. The data indicate that the face gender preference is not specifically rooted in Western cultural caregiving practices. In addition, the race dependency of the effect previously observed for Caucasian infants reared by Caucasian caregivers looking at Caucasian but not Asian faces extends to Asian infants reared by Asian caregivers looking at Asian but not Caucasian faces. The findings also provide additional support for an experiential basis for the gender preference, and in particular suggest that cumulative male face experience plays a role in inducing a downturn in the preference in older infants.

Keywords: infant, gender preference, caregiving arrangements, other-race effect, age-related

Introduction

The development of face processing is greatly shaped by visual experience (Le Grand et al., 2001; Pascalis et al., 2002, 2005; Kelly et al., 2007a,b; Cassia et al., 2009; see Lee et al., 2013). Previous

studies have reported that differential face gender experience provided by caregiving interactions influences the formation of face gender categories at a very early stage of life (Quinn et al., 2002, 2008, 2010; Rennels and Davis, 2008). Infants tend to prefer the face gender of their primary caregiver. Due to the fact that previous studies only focused on face gender preference before 4 months of age, it is unclear how caregiving practices affect face gender preference in later infancy (cf. Quinn, 2002). To bridge this gap in the literature, the current study investigated the influence of caregiving involvements on face gender preference in 3-, 6-, and 9-month-old infants. Moreover, all of the previous studies examined the face gender preferences of Caucasian infants from Western countries in Europe or North America. It is therefore unclear whether the female face preference can be extended to infants from other racial groups and different cultures with distinct child rearing practices. The present study specifically recruited Chinese infants to examine the role of Asian caregiving arrangements in the development of the female face preference.

Extensive studies suggest that face experience plays an important role in shaping the development of face processing in infancy (e.g., Pascalis et al., 2002; Kelly et al., 2007a,b, 2009). Among the various kinds of face experience, the face interactions provided by caregivers are one of the earliest face experiences that infants obtain. As indicated by recent studies, infants spend approximately 70% of their time with female faces, inclusive of an approximate 50% contribution from a female caregiver and an approximate 20% contribution from other female faces, as estimated by parental reports during the first year of life (Rennels and Davis, 2008). The 70% estimate of female face experience has been corroborated by data obtained with a head mounted camera during the first 3 months of life (Sugden et al., 2014). Most relevant to the present paper is the finding that how infants represent gender information in faces is influenced by the gender of the primary caregiver by 3–4 months of age. Quinn et al. (2002) presented infants with male versus female face pairings and reported that the infants who were primarily raised by female caregivers preferred the female faces. In contrast, infants raised mostly by male caregivers showed a reliable preference for male faces. Also, consistent with an experiential account, a more recent study found that newborns did not show any gender preference, and that the female face preference is limited to own-race faces at 3 months of age (Quinn et al., 2008). In addition, Quinn et al. (2010) found 3- to 4-month-olds were able to generalize their female face preference across age and display a preference for a girl over boy prototype face. Taken together, the results suggest that differential female versus male face experience in caregiving contributes to the differences in how infants respond to female versus male faces.

According to Quinn et al. (2002), the face gender preferences of infants reflect caregiving arrangements, which are determined by differential involvements of female and male caregivers. This account implies that day-to-day caregiving interactions play a crucial role in how infants come to represent face gender information. Extensive experience with a female or male primary caregiver helps infants form a representation of face gender that favors the gender of the more frequently experienced caregiver at a relatively early stage of life (i.e., 3–4 months). Some prior studies suggest that

face gender categories formed in early infancy undergo further development during later infancy (e.g., 7–10 months, Leinbach and Fagot, 1993; Younger and Fearing, 1999). However, it is presently unclear how caregiving practices influence responsiveness to face gender categories over 4 months of age.

The female face preference suggests that the female face representation is formed earlier than the male face representation. Studying age related development in face gender preference offers an extension of the current literature, and can provide evidence relevant to how gender representations may change in the first year of life. One possible developmental model is a ratio or proportional experience model. By this model, a consistent difference between the female and male caregiving involvement could result in a constant reliable difference in female and male face representation, and we would expect a comparable female face preference across ages. Another possible developmental model is a threshold model. That is, when male experience reaches some threshold amount, we may observe a decreased female face preference with age. This is because with increased age, infants might eventually acquire sufficient male face experience to form a representation for male faces comparable to that of female faces. The later formed male face representation would presumably interfere with the initial female face preference, which was driven by the earlier-formed female representation. If the latter is the case, we may find a female face preference in the younger infant participants in our sample. By contrast, the older infants in our sample may not show a differential preference for female or male faces.

The present study was designed to examine the influence of caregiving on the face gender preference during infancy. Specifically, it was conducted with Chinese infants at 3, 6, and 9 months of age, who were primarily reared by female caregivers. Like caregiving in North America, caregiving in China is female dominant. In addition, the relative proportions of female and male involvement in infant's caregiving are relatively stable in Asian countries. China thus offers an ideal environment to study the influence of caregiving on the development of face gender representation in the first year of life.

We used a visual preference task to examine the face gender preferences of infants. In addition, we investigated each infant participant's caregiving arrangement, as determined by the amount of female and male involvement in caregiving. The caregiving arrangement allows us to directly examine the relationship between an infant's face gender preference and the involvement of female and male caregivers.

There is one additional rationale for the current study. To our knowledge, all of the prior studies on gender preference were performed with predominantly Caucasian participants. This makes it unclear whether the finding of Quinn et al. (2008) that the female face preference is limited to own-race faces indicates that the face gender preference is limited to faces from the infant's own race group (i.e., own-race faces) or to a specific type of faces (i.e., Caucasian faces). The present study for the first time investigated whether the female face preference can be observed in Asian infants, and whether it will hold for Asian faces, but not for Caucasian faces. We therefore used both own-race Asian faces and other-race Caucasian faces to examine the face gender preferences of Asian infants. If the face gender preference is

limited to own-race faces, we should observe it only in own-race Asian faces, but not in other-race Caucasian faces.

Materials and Methods

Participants

We recruited infant participants through advertisements posted on community news boards. Eighty-four Chinese infants participated in the current study, and their parents gave consent for the infants to participate in the current study. There were 11 participants in the 3-month-old group (age range = 86–110 days; 4 females, 7 males), 40 participants in the 6-month-old group (age range = 178–196 days; 15 females, 25 males), and 33 participants in the 9-month-old group (age range = 268–289 days; 17 females, 16 males). The relatively small sample size in the 3-month-old group was partially due to the fact that several previous gender preference studies have reported a robust gender preference with small samples of 3-month-olds (e.g., Quinn et al., 2002). By contrast, the older age groups have not previously been tested for gender preference. We therefore needed stronger statistical power to evaluate the performance of these groups.

All participants were healthy, full-term infants. Thirteen additional infants also participated in the current study, but were excluded from final analysis because of extreme side bias ($n_{3\text{-month}} = 6$; $n_{6\text{-month}} = 6$) or fussiness ($n_{9\text{-month}} = 1$). Extreme side bias occurred when a participant spent more than 90% of their looking time on one side of the display for each presentation. Based on parental report, no infants had direct other-race face experience before the current study.

Materials

The stimuli were face pairs comprised of female and male frontal-view images placed side-by-side. Examples can be observed in **Figure 1**. Facial external features were removed by an oval mask, which was sized 13.80 cm in width and 19.07 cm in height. We made further adjustments to make sure that all faces were of the same size and comparable in brightness and contrast. The faces were between 25 and 29 years of age. Overall, three pairs of Asian faces and three pairs of Caucasian faces were used. Eight Chinese adults who were blind to the purpose of the research were recruited to rate each face's attractiveness and gender distinctiveness using Likert scales. The results showed that these faces were comparable in attractiveness ($p = 0.874$) and gender distinctiveness ($p = 0.723$).

Procedure

All infants were tested in a quiet room while sitting on the lap of a parent. A 42-inch monitor (width = 90.00 cm, height = 57.00 cm) was placed in front of participants for a viewing distance of approximately 75 cm. Before the test started, an experimenter examined the position of the infants to make sure they were aligned to the center of the monitor. The parent who held the infant was asked to keep the infant stable during the test. The parent was also asked to keep their eyes closed during testing.

Once a participant was in place, the test started. For the own-race condition, participants would see one own-race female and one own-race male face displayed on each side of the monitor.
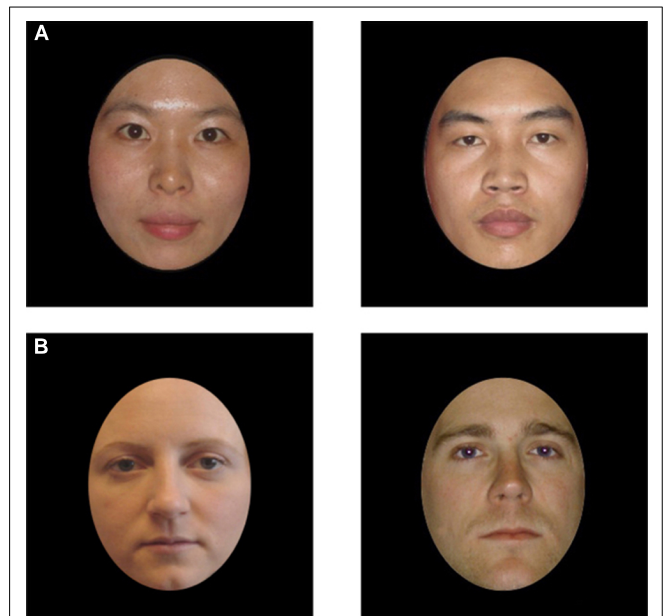
FIGURE 1 | Examples of a female–male own-race Asian face pair (A) and a female–male other-race Caucasian face pair (B).

The two faces were displayed at the eye level of the participants, and were separated by a 33.04 cm (24.85° in visual angle) gap. The face pair was presented cumulatively for 10 s. Then, the left–right positions of the two faces were switched and the faces were presented for another 10 s. Stimulus presentation was controlled by a computer program. The other-race condition was the same as the own-race condition, except that two Caucasian faces were presented. Looking time of the participants was recorded by a camera placed above the monitor. These recordings were used for offline coding to indicate each infant's proportional looking time for female versus male faces. One of the three female–male face pairs was randomly selected for each participant in each condition. The order of presentation of the races was counterbalanced across infants. The initial position (i.e., the left or right side) of the female face was randomly selected by the computer program.

Each infant's looking time for faces on the left and right sides was measured by examining each frame of the camera recordings. The camera recordings contained only each participant's face, without any cues of the screen content that participants were watching. Each video was trimmed according to face pair onset and offset time. We extracted each frame of the video (25 frames/second) to save as an image file. We then randomly presented these images to raters, who determined which side of the screen participants looked at. We finally calculated the numbers of left looking frames and right looking frames for each participant in the own- and other-race conditions, and these were used for calculating percentage looking time. Two independent raters participated in the coding process, and their inter-rater agreement was high (*Pearson r* = 0.90). Therefore, we averaged the looking time measurement from the two raters for the following analyses. A preliminary analysis revealed no significant gender difference for participants; thus, data were collapsed across participant gender in subsequent analyses.
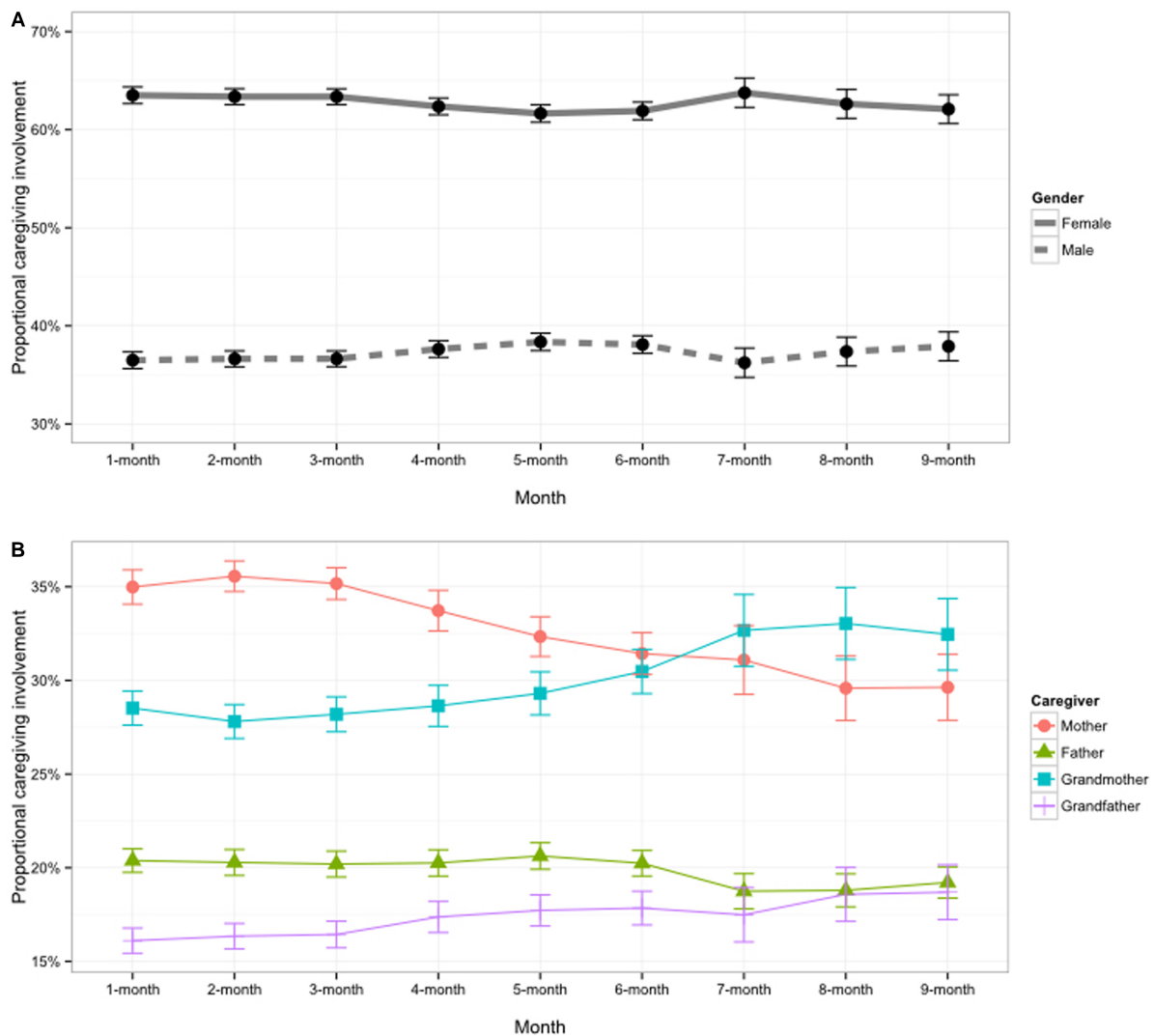
FIGURE 2 | The proportional caregiving involvement of female and male caregivers (A) and the proportional caregiving involvement of each type of caregiver (i.e., mother, father, grandmother, and grandfather) (B) within each month. Error bars represents one standard error.

In addition to the face gender preference test, parents completed a questionnaire to report the monthly involvement of each family member in caregiving. The family members included the mother, father, grandmother on the mother's side, grandmother on the father's side, grandfather on the mother's side, and grandfather on the father's side. For each family member, one of four involvement levels was requested: (1) never, (2) occasionally, (3) often, and (4) always. Parents were asked to report the involvement in every month after the infant's birth. For example, a 3-month-old's parents should provide the involvement in the 1st, 2nd, and 3rd month, while a 6-month-old's parents should provide the involvement from the 1st to the 6th month.

Chinese Caregiving Arrangements

To derive the amount of female and male caregiver involvement for each participant, we translated the involvement responses into proportional involvement scores. To do so, we first converted

parent responses into numbers (never = 0, occasionally = 1, often = 2, and always = 3). Based on the caregiving data provided by the parents and in accord with Chinese culture, caregiving provided by grandparents on the mother and father sides was coordinated and offered in alternate months. Thus, for example if the grandparents on the mother side provided caregiving in the first month, then the grandparents on the father side provided caregiving in the second month, and so on. We therefore used a single grandmother score and a single grandfather score in any given month to indicate the involvements of grandparents on the mother and father side. Then, a family involvement score was derived from the sum of each caregiver's involvement ($I_{family} = I_{mother} + I_{father} + I_{grandmother} + I_{grandfather}$). We calculated female involvement scores by summing scores of mother and grandmother ($I_{female} = I_{mother} + I_{grandmother}$) and male involvement scores by adding up scores of father and grandfather ($I_{male} = I_{father} + I_{grandfather}$). To derive proportional involvement

scores for females, we further divided female and male scores by the total family involvement score ($ProI_{female} = I_{female}/I_{family}$). On average, females played the majority part in caregiving (63.86%), which did not vary across months [ANOVA, $F(8, 626) = 0.66$, $p = 0.729$]. In addition, we further calculated the proportional involvement scores for each category of caregiver (i.e., mother, father, grandmother, and grandfather). The monthly proportional caregiving involvements are plotted in **Figure 2**, broken down by female versus male in the top panel, and the more specific caregiving categories in the bottom panel. Overall, **Figure 2** illustrates the consistent advantage in female caregiving across the first 9 months of infancy, and provides evidence that the female advantage holds whether one is comparing mother with father or grandmother with grandfather.

Results

Face Gender Preference for Own- and Other-Race Faces

We first calculated a female face preference for each participant. This percentage preference score was derived by dividing the female face looking time by the total looking time on the female and male faces and then multiplying that fraction by 100.

To examine the effect of face race and participant age on the female face preference, a 2 (face race: own and other) × 3 (age group: 3, 6, and 9 months) mixed repeated measures analysis of variance (ANOVA) was conducted. We found a significant main effect of race, $F(1, 81) = 9.25$, $p = 0.003$, $\eta_p^2 = 0.85$. On average, infants showed a significantly larger female face preference for own-race faces ($M = 57.27\%$, SD $= 17.63\%$) than for other-race faces ($M = 46.45\%$, SD $= 21.62\%$). There was no significant age group effect [$F(2, 81) = 0.24$, $p = 0.789$, $\eta_p^2 = 0.09$], or interaction between race and age [$F(2, 81) = 1.59$, $p = 0.210$, $\eta_p^2 = 0.33$].

To further examine whether the female face preference at each age was above chance, we performed one-sample t-tests for each age group, in which the female face preference was compared to chance responding (i.e., 50%). For own-race faces, as shown in **Figure 3**, we found that 3- and 6-month-olds showed a significant female face preference [3-month-olds: $M = 65.18\%$, SD $= 15.94\%$, $t(10) = 3.16$, Cohen's $d = 2.00$, $p = 0.010$; 6-month-olds: $M = 56.79\%$, SD $= 19.07\%$, $t(39) = 2.25$, Cohen's $d = 0.72$, $p = 0.030$]. However, 9-month-olds did not show a preference for female or male faces [$M = 52.67\%$, SD $= 15.57\%$, $t(32) = 0.98$, Cohen's $d = 0.35$, $p = 0.332$]. Moreover, a Pearson correlation analysis revealed that the female face preference decreased with age [$r(82) = -0.208$, $p = 0.029$]. As for the other-race Caucasian faces, we did not find a gender preference at any age: 3-month-olds: $M = 41.67\%$, SD $= 35.20\%$, $t(10) = -0.77$, Cohen's $d = -0.50$, $p = 0.451$; 6-month-olds: $M = 45.56\%$, SD $= 21.90\%$, $t(39) = -1.28$, Cohen's $d = -0.41$, $p = 0.207$; 9-month-olds: $M = 49.13\%$, SD $= 14.99\%$, $t(32) = -0.33$, Cohen's $d = -0.11$, $p = 0.742$.

Taken together, the present results regarding the female face preference replicated previous findings that infants at 3 months of age show a reliable preference for own-race female faces over own-race male faces (Quinn et al., 2002, 2008; Hillairet de Boisferon

FIGURE 3 | Mean proportional female face looking time for own- and other-race faces. Error bars represent one standard error. The asterisks suggest a significant preference for female over male faces (*$p < 0.05$).
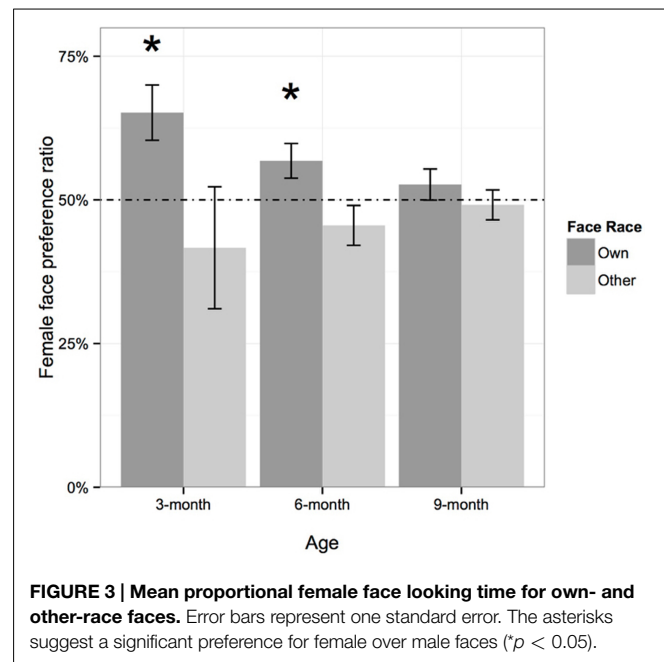
et al., 2014). Also, the current findings indicate that the female face preference is present only in 3- and 6-month-olds. The 9-month-olds spent equivalent amounts of time looking at female versus male faces. When we consider the downturn in the female face preference with age coupled with the relative constancy of the female advantage in caregiving, the two outcomes taken together suggest that infants might gradually gain sufficient male face experience to form comparable representations of female and male faces, thereby significantly reducing their visual preference for female faces, a point that we explore further in the next section of the Results.

In addition, in accord with Quinn et al. (2008), the female face preference was only observed in own-race faces, but not in other-race faces. This finding indicates that the findings of Quinn et al. (2008) are not limited to Caucasian participants or Caucasian faces. Rather, the results seem to reflect a general phenomenon regarding difference in access to gender diagnostic information in own- versus other-race faces.

The Role of Cumulative Male Caregiving Involvement in Face Gender Preference

To examine the contribution of accumulated male face experience to the downturn in the female face preference, we first calculated the cumulative male face experience by taking the data in **Figure 2** and summing each participant's male face experience across months. Not surprisingly, cumulative male face experience increased with age [$r(82) = 0.52$, $p < 0.001$]. We then performed a partial Pearson correlation between the accumulated male face experiences gained in caregiving and the proportional female face preference after controlling for the effect of the proportional female face experience. Importantly, the results indicated that the amount of male face experience was negatively correlated with the female face preference [partial $r(81) = -0.22$, $p = 0.046$, **Figure 4**],

FIGURE 4 | Linear correlation between the accumulated male face experience and the proportional own-race female face preference after controlling for the effect of the female–male face experience ratio.
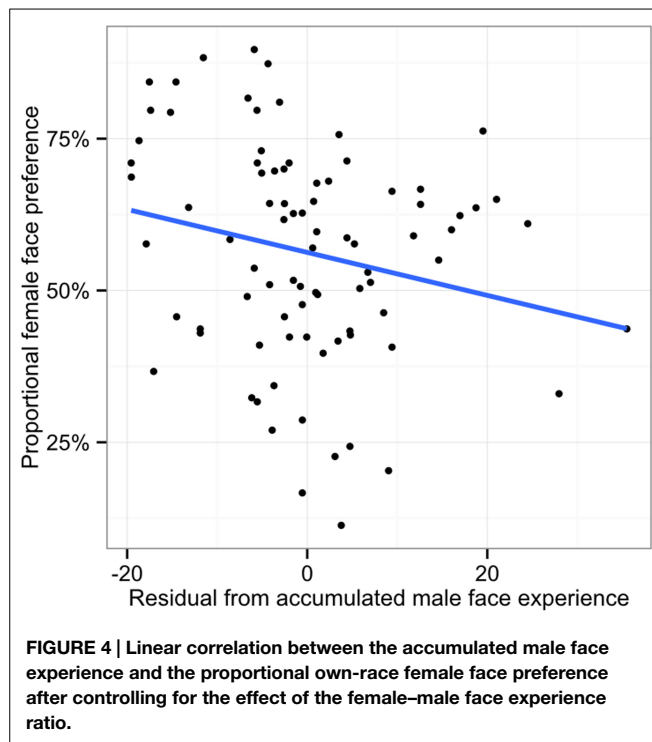
when the influence of female–male experience ratio was controlled. The pattern of correlations suggests that with increased age, infants may obtain sufficient male experience to account for the downturn in the female face preference. Additionally, we failed to find such partial correlation between the male caregiver experience and the other-race Caucasian female face preference, partial $r(81) = 0.16$, $p = 0.138$. The results indicate that male experience in caregiving only influenced infant's own-race face gender preference, but not other-race face gender preference.

Discussion

The present study investigated face gender preference in 3-, 6-, and 9-month-old Chinese infants who were primarily reared by female caregivers. Three major findings were obtained: (1) Three- and 6-month-olds exhibited a significant preference for female own-race faces over male own-race faces. No such gender preference was revealed in 9-month-olds. (2) The female face preference decreased with the increase in cumulative male face experience. (3) Infants failed to show any gender preference for other-race Caucasian faces at any age. These results together indicate the important role of face gender experience within caregiving practice in how infants respond to face gender categories.

The finding of a female face preference in 3-month-olds is consistent with previous findings that 3- to 4-month-old Caucasian infants preferred own-race Caucasian female faces over male faces (Quinn et al., 2002, 2008, 2010; Hillairet de Boisferon et al., 2014). Moreover, by examining female and male caregiving involvement, we found a caregiving bias in favor of females in Asian culture similar to that previously reported in Western culture, in which females form the category of greater experience (Rennels and

Davis, 2008; Sugden et al., 2014). The data from the three studies taken together suggest female face dominance in infant face experience. Moreover, the fact that the current study and that of Rennels and Davis (2008) relied on parental report data, whereas Sugden et al. (2014) reported data obtained with head mounted cameras, suggests that an approximate 2 to 1 female to male ratio is independent of the particular measure of experience. In addition, the fact that the prior studies were conducted in North America, one in the US (Rennels and Davis, 2008) and one in Canada (Sugden et al., 2014), whereas the current study was conducted in China, suggests that the dominant female to male face experience transcends culturally-based differences in Western and Eastern child rearing practices. The relatively rich interaction with female caregivers contributes to forming a more robust representation of female faces relative to that of male faces at an early stage of life, and has even been shown to have specific neural correlates (Righi et al., 2014). The discrepancy in the robustness of the representations based on differential experience presumably leads to a preference for female faces, which has been shown to generalize even to a girl prototype face (Quinn et al., 2010).

In addition to the female face preference found in the 3-month-olds, the present study also examined the gender preference in older infants. Despite the fact that female proportional involvement in caregiving remains stable across the first 9 months of age, we only observed the female face preference in the 3- and 6-month-olds. No significant gender preference was found in the 9-month-olds. Moreover, as indicated by the correlation analysis, the female face preference decreased significantly with the increase in accumulated male face experience. This overall pattern of results is consistent with an account in which a representation of female faces is formed relatively early because of the greater female caregiving experience. Representation of male faces, due to the lesser male face experience, may take longer to develop. The quality of the representation of male faces may eventually reach the same level as the quality of the representation for female faces, and then eventuate in a null preference between female and male faces in older infants. The finding of a developmental downturn in face gender preference supports the hypothesis that a face gender representation is developing with face experience until the amount of experience reaches a threshold amount. Such a mechanism would account for the age-related female preference decrease.

The age-related change in the female face preference is broadly consistent with other changes in the development of face representation more generally. The emergence of face prototypes was reported around 3 months of age (de Haan et al., 2001). Evidence for categorical differentiation of faces can be found in the preference for the more familiar gender (Quinn et al., 2002) and race (Kelly et al., 2005) at 3–4 months of age. This preliminary representation of social categories may undergo further development as the infant gains more experience, which may occur around 7–10 months of age. Infants at this age start to exhibit more reliable abilities to categorize face gender and race (Leinbach and Fagot, 1993; Younger and Fearing, 1999; Anzures et al., 2010). At these older ages, the preference found in young infants may be likely to disappear, which reflects comparable efficiency in processing faces of greater or lesser experience

(de Haan and Nelson, 1999) or a transition phase for further development (Liu et al., 2015). Moreover, representation of social categories in the changeover from infancy to early childhood might follow a perceptual-to-conceptual transition (Madole and Oakes, 1999; Quinn et al., 2001, 2013). For example, 3-year-olds tend to choose a novel object from a person of their own gender (Shutts et al., 2010). Five-year-olds show a similar race-based social preference for own-race individuals over other-race ones (Kinzler and Spelke, 2011). It would be worthwhile in future studies to bridge the gap between the face biases of infants and social biases of children so as to reveal a fuller trajectory in the development of responsiveness to the social world.

As for the other-race Caucasian faces, we did not find any preference for female or male faces at any age. This finding supports the argument that it is face race (i.e., own-race vs. other-race) rather than a specific type of face (e.g., Asian faces or Caucasian faces) that leads to the female face preference being observed only for own-race faces. This race-related difference may indicate that the extensive experience with own-race faces influences the perceptual cues used for categorizing face gender. With extensive experience with own-race faces, infants might develop a race specific strategy for processing facial gender information by relying on the gender diagnostic information specific to own-race faces. However, the cues used to process gender in own-race faces might not apply to the processing of gender in other-race faces, which results in the disappearance of the gender preference for other-race faces. This perceptual argument is supported by a recent study with Caucasian infants, which reported that hairline is a diagnostic cue for processing the gender of Caucasian faces (Hillairet de Boisferon et al., 2014). However, this cue could not apply for processing the gender of the Asian faces in the present study, given that the hairline and other external facial features were removed.

The race specific female face preference suggests a hierarchical structure in representing own-race faces, in which race is superordinate to gender information. The present findings, along with those of Quinn et al. (2008), are additionally consistent with several adult findings indicating that face gender information in own-race faces is processed more readily relative to that in other-race faces (e.g., O'Toole et al., 1996; Rossion, 2002). Together with the findings that the processing of face race information by infants was not affected by face gender (Kelly et al., 2005, 2007a,b, 2009), the pattern of outcomes taken together supports a model of face processing, in which face race information may be processed before gender information. This model, which is further supported by event related potential (ERP) evidence in which the face race processing related ERP components can be observed earlier than components associated with gender processing (Ito and Urland, 2003), challenges the classic face processing model that race and gender are processed independently (Bruce and Young, 1986).

We acknowledge a few limitations in the current study. It only investigated face experience within the context of family caregiving practices. However, infants, especially older ones, also acquire face experience outside the family environment, which could potentially affect the development of gender categories. For example, in Rennels and Davis (2008), the ratio of non-caregiver female faces to non-caregiver male faces was more than 2 to 1, which is consistent with the possibility that female caregivers may have more same-sex friends who come to interact with the infant. To overcome this limitation, future studies could consider using other investigative techniques, such as head-mounted cameras to record the infant's face experience both inside and outside the family caregiving environment (Sugden et al., 2014). In addition, it is not completely clear whether disproportional gender experience within caregiving contributes to the development of other aspects of face processing. Although Quinn et al. (2002) showed that infants are more likely to represent a set of female faces as individual exemplars and a set of male faces at the category level of male, we do not know if perceptual narrowing occurs in the processing of gender. For example, it is possible that infants might develop more sophisticated capabilities to recognize faces (Kelly et al., 2007b, 2009) or to process facial expressions (Kahana-Kalman and Walker-Andrews, 2001) specific to their more familiar gender category. Future studies might consider using multiple tasks to further investigate the role of face experience for a variety of aspects of face processing.

Acknowledgments

References

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., and Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Dev. Sci.* 13, 553–564. doi: 10.1111/j.1467-7687.2009.00900.x

Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77, 305–327. doi: 10.1111/j.2044-8295.1986.tb02199.x

Cassia, V. M., Kuefner, D., Picozzi, M., and Vescovo, E. (2009). Early experience predicts later plasticity for face processing: evidence for the reactivation of dormant effects. *Psychol. Sci.* 20, 853–859. doi: 10.1111/j.1467-9280.2009.02376.x

de Haan, M., Johnson, M. H., Maurer, D., and Perrett, D. I. (2001). Recognition of individual faces and average face prototypes by 1- and 3-month-old infants. *Cogn. Dev.* 16, 659–678. doi: 10.1016/S0885-2014(01)00051-X

de Haan, M., and Nelson, C. A. (1999). Brain activity differentiates face and object processing in 6-month-old infants. *Dev. Psychol.* 35, 1113–1121. doi: 10.1037/0012-1649.35.4.1113

Hillairet de Boisferon, A., Uttley, L., Quinn, P. C., Lee, K., and Pascalis, O. (2014). Female face preference in 4-month-olds: the importance of hairline. *Infant Behav. Dev.* 37, 676–681. doi: 10.1016/j.infbeh.2014.08.009

Ito, T. A., and Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *J. Pers. Soc. Psychol.* 85, 616–626. doi: 10.1037/0022-3514.85.4.616

Kahana-Kalman, R., and Walker-Andrews, A. S. (2001). The role of person familiarity in young infants' perception of emotional expressions. *Child Dev.* 72, 352–369. doi: 10.1111/1467-8624.00283

Kelly, D. J., Liu, S., Ge, L., Quinn, P. C., Slater, A. M., Lee, K., et al. (2007a). Cross-race preferences for same-race faces extend beyond the African versus Caucasian contrast in 3-month-old infants. *Infancy* 11, 87–95. doi: 10.1207/s15327078in1101_4

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., and Pascalis, O. (2007b). The other-race effect develops during infancy: evidence of perceptual narrowing. *Psychol. Sci.* 18, 1084–1089. doi: 10.1111/j.1467-9280.2007.02029.x

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., et al. (2009). Development of the other-race effect during infancy: evidence toward universality? *J. Exp. Child Psychol.* 104, 105–114. doi: 10.1016/j.jecp.2009.01.006

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., et al. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Dev. Sci.* 8, F31–F36. doi: 10.1111/j.1467-7687.2005.0434a.x

Kinzler, K. D., and Spelke, E. S. (2011). Do infants show social preferences for people differing in race? *Cognition* 119, 1–9. doi: 10.1016/j.cognition.2010.10.019

Lee, K., Quinn, P. C., Pascalis, O., and Slater, A. (2013). "Development of face processing ability in childhood," in *Oxford Handbook of Developmental Psychology*: Vol. 1, *Body and Mind*, ed. P. D. Zelazo (Oxford: Oxford University Press), 338–370.

Le Grand, R., Mondloch, C. J., Maurer, D., and Brent, H. P. (2001). Early visual experience and face processing. *Nature* 410, 890–890.

Leinbach, M. D., and Fagot, B. I. (1993). Categorical habituation to male and female faces: Gender schematic processing in infancy. *Infant Behav. Dev.* 16, 317–332. doi: 10.1016/0163-6383(93)80038-A

Liu, S., Xiao, W. S., Xiao, N. G., Quinn, P. C., Zhang, Y., Chen, H., et al. (2015). Development of visual preference for own- versus other-race faces in infancy. *Dev. Psychol.* 51, 500–511. doi: 10.1037/a0038835

Madole, K. L., and Oakes, L. M. (1999). Making sense of infant categorization: stable processes and changing representations. *Dev. Rev.* 19, 263–296. doi: 10.1006/drev.1998.0481

O'Toole, A. J., Peterson, J., and Deffenbacher, K. A. (1996). An "other-race effect" for categorizing faces by sex. *Perception* 25, 669–676. doi: 10.1068/p250669

Pascalis, O., de Haan, M., and Nelson, C. A. (2002). Is face processing species-specific during the first year of life? *Science* 296, 1321–1323. doi: 10.1126/science.1070223

Pascalis, O., Scott, L. S., Kelly, D. J., Shannon, R. W., Nicholson, E., Coleman, M., et al. (2005). Plasticity of face processing in infancy. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5297–5300. doi: 10.1073/pnas.0406627102

Quinn, P. C. (2002). "Beyond prototypes: asymmetries in infant categorization and what they teach us about the mechanisms guiding early knowledge acquisition," in *Advances in Child Development and Behavior*, Vol. 29, eds R. Kail and H. Reese (San Diego, CA: Academic Press), 161–193.

Quinn, P. C., Anzures, G., Lee, K., Pascalis, O., Slater, A., and Tanaka, J. W. (2013). "On the developmental origins of differential responding to social category information," in *Navigating the Social World: What infants, Children, and Other Species can Teach us*, eds M. R. Banaji and S. A. Gelman (New York: Oxford University Press), 286–291.

Quinn, P. C., Conforto, A., Lee, K., O'Toole, A. J., Pascalis, O., and Slater, A. M. (2010). Infant preference for individual women's faces extends to girl prototype faces. *Infant Behav. Dev.* 33, 357–360. doi: 10.1016/j.infbeh.2010.03.001

Quinn, P. C., Eimas, P. D., and Tarr, M. J. (2001). Perceptual categorization of cat and dog silhouettes by 3- to 4-month-old infants. *J. Exp. Child Psychol.* 79, 78–94. doi: 10.1006/jecp.2000.2609

Quinn, P. C., Uttley, L., Lee, K., Gibson, A., Smith, M., Slater, A. M., et al. (2008). Infant preference for female faces occurs for same- but not other-race faces. *J. Neuropsychol.* 2, 15–26. doi: 10.1348/174866407X231029

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., and Pascalis, O. (2002). Representation of the gender of human faces by infants: a preference for female. *Perception* 31, 1109–1121. doi: 10.1068/p3331

Rennels, J. L., and Davis, R. E. (2008). Facial experience during the first year. *Infant Behav. Dev.* 31, 665–678. doi: 10.1016/j.infbeh.2008.04.009

Righi, G., Westerlund, A., Congdon, E. L., Troller-Renfree, S., and Nelson, C. A. (2014). Infants' experience-dependent processing of male and female faces: insights from eye tracking and event-related potentials. *Dev. Cogn. Neurosci.* 8, 144–152. doi: 10.1016/j.dcn.2013.09.005

Rossion, B. (2002). Is sex categorization from faces really parallel to face recognition? *Vis. Cogn.* 9, 1003–1020. doi: 10.1080/13506280143000485

Shutts, K., Banaji, M. R., and Spelke, E. S. (2010). Social categories guide young children's preferences for novel objects. *Dev. Sci.* 13, 599–610. doi: 10.1111/j.1467-7687.2009.00913.x

Sugden, N. A., Mohamed-Ali, M. I., and Moulson, M. C. (2014). I spy with my little eye: typical, daily exposure to faces documented from a first-person infant perspective. *Dev. Psychobiol.* 56, 249–261. doi: 10.1002/dev.21183

Younger, B. A., and Fearing, D. D. (1999). Parsing items into separate categories: developmental change in infant categorization. *Child Dev.* 70, 291–303. doi: 10.1111/1467-8624.00022

Infants' neural responses to facial emotion in the prefrontal cortex are correlated with temperament: a functional near-infrared spectroscopy study

Miranda M. Ravicz[1], Katherine L. Perdue[1,2], Alissa Westerlund[1], Ross E. Vanderwert[1,2] and Charles A. Nelson[1,2,3]*

[1] *Laboratories of Cognitive Neuroscience, Division of Developmental Medicine, Boston Children's Hospital, Boston, MA, USA,* [2] *Department of Pediatrics, Harvard Medical School, Boston, MA, USA,* [3] *Harvard Graduate School of Education, Cambridge, MA, USA*

Accurate decoding of facial expressions is critical for human communication, particularly during infancy, before formal language has developed. Different facial emotions elicit distinct neural responses within the first months of life. However, there are broad individual differences in such responses, so that the same emotional expression can elicit different brain responses in different infants. In this study, we sought to investigate such differences in the processing of emotional faces by analyzing infants' cortical metabolic responses to face stimuli and examining whether individual differences in these responses might vary as a function of infant temperament. Seven-month-old infants ($N = 24$) were shown photographs of women portraying happy expressions, and neural activity was recorded using functional near-infrared spectroscopy (fNIRS). Temperament data were collected using the Revised Infant Behavior Questionnaire Short Form, which assesses the broad temperament factors of Surgency/Extraversion (S/E), Negative Emotionality (NE), and Orienting/Regulation (O/R). We observed that oxyhemoglobin (oxyHb) responses to happy face stimuli were negatively correlated with infant temperament factors in channels over the left prefrontal cortex (uncorrected for multiple comparisons). To investigate the brain activity underlying this association, and to explore the use of fNIRS in measuring cortical asymmetry, we analyzed hemispheric asymmetry with respect to temperament groups. Results showed preferential activation of the left hemisphere in low-NE infants in response to smiling faces. These results suggest that individual differences in temperament are associated with differential prefrontal oxyHb responses to faces. Overall, these analyses contribute to our current understanding of

---

Abbreviations: deoxyHb, deoxyhemoglobin; EEG, electroencephalography; ERP, event-related potential; FFA, fusiform face area; fMRI, functional magnetic resonance imaging; fNIRS, functional near-infrared spectroscopy; lPFC, lateral prefrontal cortex; mPFC, medial prefrontal cortex; NE, negative emotionality; OFA, occipital face area; OFC, orbitofrontal cortex; oxyHb, oxyhemoglobin; PFC, prefrontal cortex; R-IBQsf, Revised Infant Behavior Questionnaire Short Form; RM-ANOVA, repeated measures analysis of variance; S/E, Surgency/Extraversion; STS, superior temporal sulcus; totalHb, total hemoglobin.

face processing during infancy, demonstrate the use of fNIRS in measuring prefrontal asymmetry, and illuminate the neural correlates of face processing as modulated by temperament.

Keywords: functional near-infrared spectroscopy, infancy, temperament, negative emotionality, emotion, face processing, prefrontal cortex

Introduction

Much of human communication is unspoken. When we are angry or fearful or happy, those emotions are often reflected in our faces, and we observe others' facial expressions in order to gather information about our social environment (Adolphs, 2002; Blair, 2003). Emotion processing is studied in infants in order to better understand where, when, and how this specialized ability develops. At the level of individuals, there are differences in the strength and sensitivity of neural responses to emotional faces (Etkin et al., 2004; Jessen and Grossmann, 2015). The present study examines the relation between infants' temperaments and their neural responses to happy face stimuli.

Research on face processing in infants has frequently focused on responses to fearful faces, and there are several compelling reasons why that is the case. The fear circuit is one of the clearest and best-documented brain circuits, and it develops at an early age in humans (LeDoux, 2000; Leppänen and Nelson, 2012). Between 5 and 7 months, infants develop a proclivity to respond preferentially to fearful faces; for example, 7-month-old infants spend more time scanning fearful faces than neutral or happy faces, and they show greater brain activity in certain face processing areas when looking at fearful faces (Nelson and Dolgin, 1985; Leppänen et al., 2007; Hoehl et al., 2008; Vanderwert and Nelson, 2014). As early as 7 months of age, infants' brain responses to fearful faces versus happy faces are distinct, with greater attention allocated to fearful faces, even when the infants do not consciously perceive the faces (Jessen and Grossmann, 2015). Research suggests that further developmental changes in emotional face processing occur between 7 and 12 months. ERP analysis shows that 7-month-old infants allocate greater attention to happy faces versus angry faces, whereas 12-month-old infants show the opposite preference (Grossmann et al., 2007). Further study is required to understand how the neural architecture underlying emotional face processing develops over the first year of life.

In the present study, we analyze infants' neural responses to happy faces. A happy face is likely the first expression that an infant sees in the world, and the facial expression most commonly experienced from a very early age. Despite this relevance to early life experience, little is known about the development of the neural architecture involved in processing happy faces.

Over the past several decades, researchers have developed a number of tools to study the developing brain and draw inferences about the neural bases of perceptual and cognitive functions. An emergent technique, fNIRS, is a non-invasive and infant-friendly methodology that measures changes in

hemoglobin concentrations as an indicator of localized brain activity. As with fMRI, the fNIRS methodology assumes that increased oxyHb concentration and decreased deoxyHb concentration correspond to increased local brain activity (Lloyd-Fox et al., 2010). The fNIRS hardware is relatively inexpensive and portable, and the optodes (i.e., emitters and detectors) are arranged in a wearable cap, which is much more tolerable for infants and relatively more robust to movement artifact than is the fMRI scanner. Moreover, because fNIRS data can be collected in awake, behaving infants, this method allows for more ecologically valid experimental tasks (**Figure 1A**). fNIRS has better spatial resolution than EEG techniques, and better temporal resolution than fMRI. It is important to note, however, that because the brain topography is less well mapped for infants than for adults, it is difficult to know which brain regions underlie specific channels in the fNIRS probe, although recent modeling work has begun to address this issue (Lloyd-Fox et al., 2014). One drawback to fNIRS, in comparison with fMRI, is that it is limited to interrogating the cortical surface. However, as fNIRS methodology continues to evolve, increasingly complex

FIGURE 1 | Functional near-infrared spectroscopy probe design.
(A) Proper placement of fNIRS cap on infant. **(B)** fNIRS probe, top view.
(C) Prefrontal panel of fNIRS probe, consisting of 22 channels (emitter-detector pairs), and the approximate locations of channels on the infant forehead.
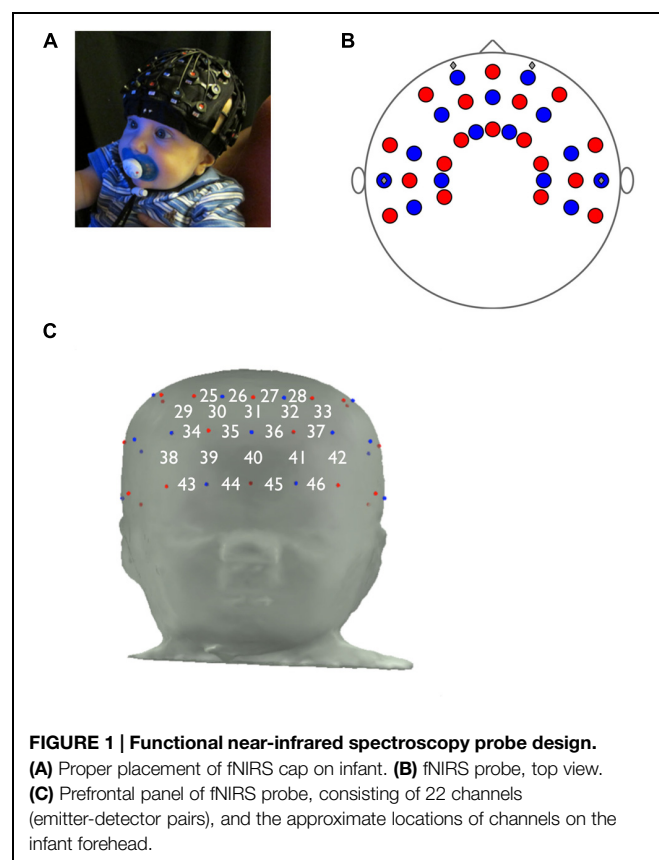
experimental paradigms will provide critical insights into the functions of the developing brain.

Recently, fNIRS studies have examined the hemodynamic response to face stimuli, confirming that a distinct brain response to faces can be measured with this methodology (Blasi et al., 2007; Lloyd-Fox et al., 2009; Vanderwert and Nelson, 2014). Nakato et al. (2011) recorded fNIRS responses over the temporal regions, demonstrating that infant hemodynamic responses to distinct types of facial stimuli (happy and angry) are significantly different. They found that the left temporal area overlying the STS was significantly activated when infants viewed happy faces, and the right temporal area was significantly activated for angry faces. These results suggest that emotions are processed differentially over regions associated with face processing in the temporal cortices; however, this study did not examine evaluative processing of emotional faces in the PFC.

Another infant fNIRS study provides evidence that the OFC is involved with face processing and emotional arousal, and that these brain functions are measurable with fNIRS methodology by 12 months of age (Minagawa-Kawai et al., 2009). In this study, infants were shown videos of their mothers and of an unfamiliar woman, posing with neutral and smiling expressions. The results found that, in a channel in the medial PFC, there was a significant difference in oxyHb activation in the smiling condition versus the neutral condition when the infant watched the video of his/her own mother. The same effect was found for the smiling versus neutral unfamiliar woman stimuli, with marginal significance. These data indicate that there is a detectable response to smiling faces in the medial OFC.

More recently, an fNIRS study of 6- and 7-month-old infants at high and low risk for autism provides further insight into how infants process smiling versus neutral faces (Fox et al., 2013). The study examined oxy- and deoxyHb responses to videos of female faces changing from a neutral expression to a smiling one. In three channels over the right frontal cortex, there was a main effect of emotion, in which the oxyHb response to smiling faces was greater than the response to the neutral expression. In three channels over the left frontal cortex, infants showed greater (that is, more negative) deoxyHb responses to smiling versus neutral faces. Taken together, these studies demonstrate that fNIRS methodology can detect stimulus-driven responses to happy faces in distinct channels over the frontal cortex.

In the research conducted to date, infants' data were pooled and examined at the group level, with little attention paid to individual differences. Thus, in the current study we sought to add an individual difference dimension — specifically, to examine whether differences in infants' temperament might be associated with differences in infants' hemodynamic response. Temperament refers to a biologically determined disposition toward certain behaviors or feelings. A person's temperament, observable during infancy and relatively stable over the lifetime (Fox et al., 2001), affects susceptibility to certain emotional states, intensity of emotion, and the ability to regulate emotional responses. Evidence suggests that temperamental biases are determined by genes influencing neurochemistry and neuroanatomy, and by the prenatal environment (Kagan and Snidman, 2004).

Childhood environment and early experiences, as well as genetic expression that is modified throughout development, will determine how a child's temperament manifests as personality traits, and whether those traits will change over time as the individual develops from infancy to childhood to adolescence and eventually into adulthood (Kandler et al., 2013). Infant temperament has been shown to be moderately predictive of temperament in toddlerhood and early childhood, with strong longitudinal correlations for the factor levels of S/E and Negative Affect (Fox et al., 2001; Putnam et al., 2008). Furthermore, a large longitudinal study showed that temperament groups at age 3 — specifically the dimensions of undercontrolled, inhibited, and well-adjusted — predict personality style at age 18, and that temperament at age 3 predicts the quality of interpersonal relationships and social support, as well as the incidence of unemployment, psychiatric disorders, and criminal behavior, at age 21 (Caspi, 2000). Though the effect size for each of these connections was only small to medium, Caspi (2000) argues that the association between temperament during toddlerhood and multiple independent measures of psychosocial functioning in young adulthood provides important evidence for the developmental continuity of temperament.

Previous investigations have found that infant temperament is associated with individual differences in emotional face processing, specifically in the amplitude and latency of the Nc, an ERP component associated with allocation of attention. Martinos et al. (2012) found that, from 3 to 13 months, infants with higher NE allocated greater attention (as indexed by the Nc) to happy faces than to fearful faces. This pattern of findings has not always been consistent, however; de Haan et al. (2004) found that more fearful infants showed a larger Nc in response to fearful faces than to happy faces. Martinos et al. (2012) attributed this discrepancy to methodological differences in temperament assessment or ERP measurement, or to the different age ranges of the subjects.

Temperament has been robustly associated with differences in EEG activity observed over the left and right frontal scalp (known as 'frontal EEG asymmetry'). According to Davidson's (1993) 'motivational model' of EEG asymmetry, relatively greater activity over the left (versus right) frontal lobe is associated with 'approach' orientation or behavior. Relatively greater right (versus left) frontal activation is associated with 'withdrawal' orientation or behavior. Infants who demonstrate withdrawal behavior are reticent and distressed when presented with unknown people or novel objects; infants demonstrating approach behavior readily approach new people and toys, and remain unperturbed in stressful situations. Studies have examined both state-dependent effects (for example, smiling can increase relative left frontal activation) and trait-dependent effects (Ekman and Davidson, 1993; Coan and Allen, 2003). The trait-dependent effects are relatively stable measures of individual temperament, with high internal consistency and acceptable test–retest stability (Tomarken et al., 1992; Fox et al., 2001). The measurement of frontal asymmetry can provide insight into the biological basis of temperament.

Many studies of frontal asymmetry have investigated the cortical response to emotional faces (Davidson and Fox, 1982; reviewed in Davidson, 1993). A meta-analysis of fMRI research

analyzing adult brain responses to photos of emotional faces found that there were significant hemispheric interactions between the category of facial expression (approach- versus avoidance-inducing) and activity in the PFC (Fusar-Poli et al., 2009). Frontal asymmetry in response to emotional faces provides information about how the PFC is involved in emotion perception and how the response is modulated by individual temperament.

Preliminary work suggests that state- and trait-related asymmetry can also be studied using fNIRS. In a study by Tuscan et al. (2013), young adult subjects were asked to complete three tasks while fNIRS activity over the PFC was recorded during the different task manipulations: conversing with strangers, planning a 5-minute speech, and delivering this speech. Subjects reported their subjective anxiety levels after each of the three tasks, and each subject's trait anxiety was evaluated using the Social Phobia and Anxiety Inventory (SPAI). During the anticipation and speech phases, subjects experienced relatively greater blood volume and oxyHb concentrations in the right relative to the left hemisphere. Further, the participants who were identified as higher in anxiety showed a trend toward greater right frontal activation relative to low-anxious subjects (Tuscan et al., 2013). This observed effect is analogous to right frontal EEG asymmetry, which is associated with withdrawal behavior and would be predicted in tasks designed to induce social stress.

We selected channels overlying the PFC as the region of interest. The PFC is involved in distinguishing facial emotions (Leppänen and Nelson, 2009) and in emotional regulation (for example, Tranel et al., 2002), roles that are likely modulated by reciprocal connections with subcortical structures involved in emotion perception, particularly the amygdala and superior colliculus (Leppänen and Nelson, 2009). We were not able to collect data from the OFA or FFA because these primary face processing areas lie too deep below the cortical surface to detect with fNIRS (Otsuka et al., 2007), and because the fNIRS probe used in this study lies over the infant's temporal and prefrontal cortices. Some fNIRS studies of infant face processing have examined the area overlying the STS, another face processing area (Lloyd-Fox et al., 2009; Nakato et al., 2011). However, because we were interested in measuring frontal asymmetry in response to facial stimuli, we chose to focus on the PFC. Furthermore, Minagawa-Kawai et al. (2009) showed significant activity in a region of the PFC for infants' response to smiling faces.

We hypothesized that there would be channels in the prefrontal panel in which hemoglobin activity would be correlated with temperament. Specifically, we hypothesized that the temperament factors of S/E and O/R would be positively correlated with brain activation, that NE would be negatively correlated with brain activation. Furthermore, to explore the capacity of fNIRS methodology to record relative asymmetry in oxyHb activity between left and right hemispheres, we analyzed the interactions of temperament and hemisphere. We hypothesized that infants with higher S/E scores would show relatively greater left frontal activity, and infants with higher NE scores would show relatively greater right frontal activity. These analyses contribute to our current understanding of

face processing during infancy, investigate the use of fNIRS in measuring prefrontal asymmetry, and examine the neural correlates of face processing as modulated by temperament.

Materials and Methods

Participants

Twenty-four 7-month-old infants were included in the study (mean age 212 ± 1.0 days, range 205–221 days; 11 females). Twenty additional infants were tested but were excluded from the study for incorrect optode placement ($n = 7$), for more than 25% of channels in the prefrontal panel rejected for artifact ($n = 6$), for equipment failure ($n = 5$), or for movement artifact ($n = 2$). The 45% attrition rate is comparable to the rate in other infant fNIRS studies (Lloyd-Fox et al., 2010). Infants who were included or excluded from the study did not differ in measures of S/E, $t(42) = 1.16$, $p = 0.252$, NE, $t(42) = -1.02$, $p = 0.314$, or O/R, $t(42) = 0.196$, $p = 0.846$. Infants were recruited from a registry of local births set up by the Laboratories of Cognitive Neuroscience. Infants were excluded from recruitment if they were born more than 3 weeks before their due date, or if they had any neurological disorders, including neurological trauma, developmental delay, uncorrected vision difficulty, or birth-related complications. Written informed consent was obtained from each infant's parent or primary caregiver prior to the start of the experiment, and the experimental protocol was approved by the Boston Children's Hospital Institutional Review Board. Written informed consent was also obtained from the parent for use of the photo in **Figure 1A**.

Infant Behavior Questionnaire

We used the R-IBQsf (Putnam et al., 2013), a parent-report measure of temperament that was completed by the mother or primary caregiver of each subject prior to the visit. The tool is validated for 3- to 12-month-old infants and has showed adequate internal consistency, inter-rater reliability between mothers and fathers, and convergence with laboratory observational assessments (Gartstein and Rothbart, 2003; Parade and Leerkes, 2008; Putnam et al., 2013).

The R-IBQsf consists of 14 subscales that load onto three broad temperament factors, as derived from principle factor analysis (Gartstein and Rothbart, 2003). The S/E factor includes temperament subscales of activity level, vocal reactivity, smiling and laughing, high intensity pleasure, perceptual sensitivity, and approach. The NE factor includes questions about fear, distress to limitations, sadness, and (loading negatively) falling reactivity. The third factor, O/R, includes the subscales of soothability, duration of orienting, cuddliness, and low intensity pleasure. Gartstein and Rothbart (2003) found low bivariate correlations between the factors.

NIRS Recording

Hemodynamic responses were recorded using a multichannel optical topography NIRS instrument (ETG-4000, Hitachi Medical Corporation, Tokyo, Japan). Near-infrared light at 695 and 830 nm was conveyed to the emitting optodes via optical

fibers and shined onto the scalp. The light that passed back through the scalp was then conveyed from the detecting optodes via optical fibers to photodetectors that measured the intensity of the attenuated light. The inter-optode distance was fixed at 3.0 cm. Data were sampled every 100 ms (10 Hz). The fNIRS probe was mounted inside a flexible cap, which was placed on the infant's head and worn for the duration of the experiment (**Figure 1A**). The probe was customized for this experiment and included 46 channels, each consisting of an emitter and detector combination, that were positioned over the frontal, temporal, and parietal cortices (**Figure 1B**). In this study the area of interest was the PFC, underlying channels 25 through 46 (**Figure 1C**).

To ensure precise and consistent spatial resolution in the fNIRS data, we adhered to stringent criteria for hat placement. During each session, photos were taken of the cap placement (frontal and lateral views), and the photos were reviewed by multiple experimenters. Subjects were excluded for incorrect hat placement if the cap was shifted by more than 1 cm in any direction (left, right, up, or down).

Task and Stimuli

Infants completed the task while sitting on their parent's lap. Parents were asked not to speak to the infant during the experiment, and they wore a visor to prevent any parental response to the visual stimuli from influencing the infant's reaction. The infant was seated approximately 60 cm from a 17-inch computer monitor. The stimuli were 16.5 cm high (visual angle: 14.3°) and 14 cm wide (visual angle: 12.2°). The testing room was soundproof and the lights were dimmed during the experiment to a standardized brightness. An experimenter sat next to the infant and parent and redirected the infant's attention to the screen before the start of each trial. In order to minimize data attrition, parents were asked to select a time for the visit when the infant was typically alert and content, and the infant was allowed to take breaks during the session as needed.

The stimuli were images of female models displaying happy, fearful, and angry facial expressions (**Figure 2A**), selected from the NimStim Face Stimulus Set (Tottenham et al., 2009). The stimuli were presented using the E-Prime Application Suite for Psychology (E-Prime 2.0, Psychology Software Tools, Sharpsburg, PA, USA).

The experiment consisted of a maximum of 30 trials, each of which included five images. The five images in a single trial were of the same emotional category (happy, fearful, or angry), portrayed by different models. Each image was shown for 1 s, with a randomly generated 200–400 ms inter-stimulus time. After the set of five images, a video of non-face shapes was shown for 10 s, resulting in a total trial length of 16 s (**Figure 2B**). There were 10 trials of each of the three emotional categories, for a total of 30 trials. The session ended when a participant viewed all 30 trials, or if the participant grew restless or upset. The order of stimulus presentation was counterbalanced across subjects.

Data Processing

Oxy- and deoxyHb data from 'happy' trials were included in this analysis. Subjects viewed a maximum number of 10 'happy' trials; they viewed fewer than 10 trials if they refused to complete the entire task, or if they looked away from the stimulus during a given trial. For each subject, a video recording of the experimental proceedings was coded offline using SuperCoder software (SuperCoder 1.7.1, Purdue University, West Lafayette, IN, USA) by observers who were blind to the emotional category. Inter-rater reliability was maintained at 0.90 with 15% coding overlap. Trials in which the infant was not looking at the stimulus for at least 50% of the time the stimulus was on the screen were excluded. Trials were not excluded for failure to look during the inter-trial video. Infants completed an average of 8 ± 0.3 'happy' trials ($N = 24$), and the range of valid trials was 5–10. We used an *a priori* threshold of three 'happy' trials for inclusion in the final sample, based on a previous study of face processing in infants (Lloyd-Fox et al., 2013).

Functional near-infrared spectroscopy data were processed using HOMER2 (MGH-Martinos Center for Biomedical Imaging, Boston, MA, USA), a MATLAB (The MathWorks, Inc., Natick, MA, USA) software package. The attenuated light intensities measured by the detecting optode at each channel were converted to optical density units, and then filtered using a band pass filter with a passband from 0.050–0.80 Hz. They were also processed using wavelet motion correction as implemented in HOMER2 with an interquartile range of 0.5 (Cooper et al., 2012; Molavi and Dumont, 2012). The filtered data were used to calculate the change in concentration of each hemoglobin chromophore according to the modified Beer–Lambert Law (Delpy et al., 1988), assuming a pathlength factor of 5 (Duncan et al., 1995). Chromophore concentrations were baseline
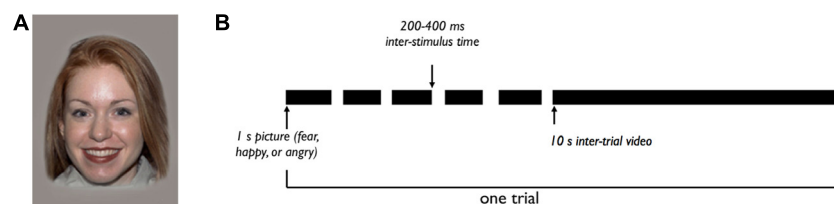
FIGURE 2 | Experimental design. (A) Example 'happy' stimulus. Each trial consists of five images like this one, with different female models portraying a happy expression. **(B)** Each trial includes five images, each presented for 1 s with a 200–400 ms inter-stimulus time. The five images in a trial represent a single emotional category (fearful, happy, or angry), with five different female models portraying the emotional expression. After the fifth image, a 10 s inter-trial video is shown, for a total trial time of about 16 s. The experiment consists of 30 total trials, with 10 trials of each emotional category.

corrected using the 2 s prior to stimulus presentation, as in previous fNIRS studies (for example, Watanabe et al., 2008).

In the data processing stream, channels in the fNIRS probe were excluded for artifact if the magnitude of the signal was greater than 98% or less than 2% of the total range for longer than 5 s during the recording. Subjects with more than 25% of channels in the region of interest marked unusable were excluded from further analysis. For this experimental probe design, there were 22 channels in the prefrontal panel, and subjects with more than five channels marked unusable were excluded.

Statistical Analyses

Statistical tests were conducted using IBM SPSS Statistics 21.0 (IBM Corporation, Armonk, NY, USA). One-sample t-tests were conducted to determine if the maximum changes in oxyHb or deoxyHb concentration in the channels of interest were significantly different from baseline levels. A time window of interest was selected between 0 and 10 s, with $t = 0$ s corresponding to the time of stimulus onset. Baseline values were measured between -2 and 0 s, as in previous fNIRS studies (for example, Watanabe et al., 2008).

Bivariate Pearson correlations were conducted between the three temperament factors (S/E, NE, and O/R). Pearson correlations were then calculated for channels in the region of interest to test for a relation between temperament and oxyHb and deoxyHb activity (maximum amplitude). Due to a significant correlation between S/E and O/R (see below), partial correlations were also conducted to test the independent relations between temperament and Hb activity. In order to fully explore the connection between happy faces and temperament, only the responses to happy face stimuli were included in these analyses.

Repeated measures analysis of variance was used to test for a hemispheric effect of temperament on oxyHb activity, with hemisphere (left versus right) as the within-subjects factor and the temperament group (low versus high) as the between-subjects factor. Activity in the left and right hemispheres was calculated as the mean value of the maximum change in oxyHb amplitude for two channels in the left hemisphere (36 and 41) and for two channels in the right hemisphere (35 and 39). A median split was used to divide subjects into low and high temperament groups for

each factor (S/E, NE), as in previous studies (for example, Baehr et al., 1998; Hagemann et al., 1999; Hagemann, 2004).

Results

Prefrontal Activation in Response to Happy Face Stimuli

We plotted the grand averaged time courses of the changes in concentration of oxy-, deoxy-, and totalHb for the channels in the prefrontal panel. The responses for channels 25 and 46 are shown in **Figure 3**. Based on these responses, we selected a time range of 0–10 s for analysis.

To test whether channels overlying the PFC were significantly activated by happy face stimuli, we conducted one-sample t-tests for the maximum change in oxyHb and deoxyHb concentrations in these channels. The changes in Hb concentration were calculated relative to the baseline value at $t = -2$ to 0 s. Across subjects, there was a significant decrease in oxyHb concentration in channel 25, $t(20) = -2.767$, $p = 0.012$, and in channel 46, $t(23) = -2.387$, $p = 0.026$. There was a significant increase in deoxyHb concentration in channel 34, $t(23) = 2.959$, $p = 0.007$.

Differential Brain Responses According to Temperament

In our sample, we found no significant correlation between NE and S/E or O/R, which is consistent with previous findings (Gartstein and Rothbart, 2003). We did find a significant correlation between S/E and O/R, $r(22) = 0.674$, $p < 0.001$.

We conducted Pearson correlations to test whether oxyHb and deoxyHb activity (calculated as maximum change in concentration from baseline) were correlated with temperament in the prefrontal channels. The results are shown in **Table 1**. A total of $N = 24$ subjects were tested, but because some channels did not have reliable data from all subjects, the number of subjects (n) tested for each channel is shown in the table. The temperament factor S/E was negatively correlated with oxyHb activity in channel 26, $r(21) = -0.521$, $p = 0.011$, and channel 32, $r(21) = -0.457$, $p = 0.028$; in these channels, infants with lower S/E scores showed greater activation in response to happy faces.

FIGURE 3 | Functional near-infrared spectroscopy activation in response to happy face stimuli in channels 25 and 46. Grand averaged time courses of the changes in concentration of oxy- (red), deoxy- (blue), and totalHb (green) in **(A)** channel 25 and **(B)** channel 46.
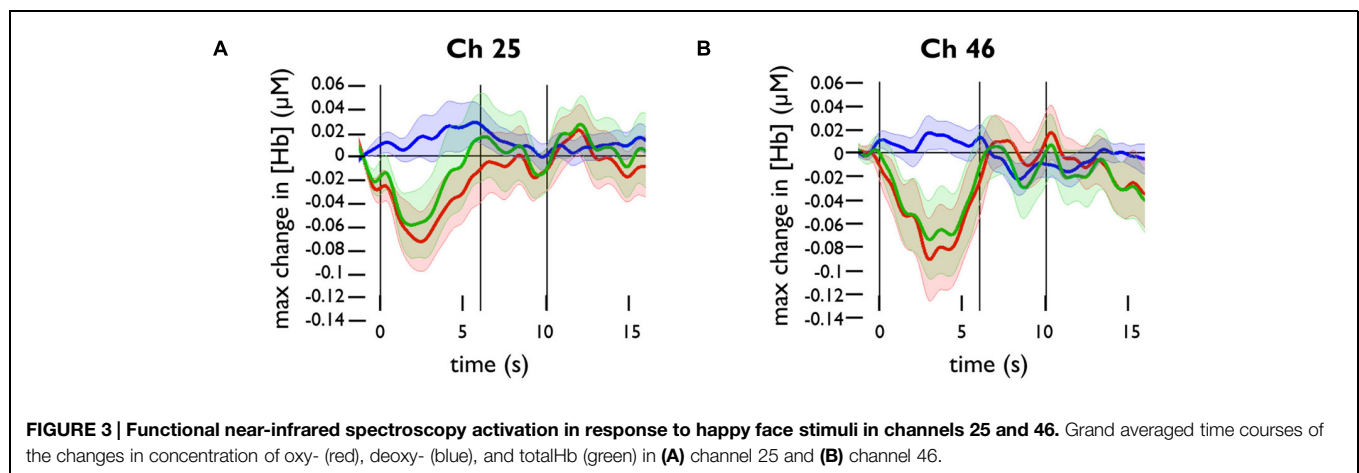
TABLE 1 | Summary of channels in which maximum change in concentrations of oxyHb or deoxyHb are significantly correlated with the temperament factors.

Channel	n	oxyHb			deoxyHb		
		S/E	NE	O/R	S/E	NE	O/R
26	23	−0.521*					
27	21			−0.525*	−0.500*		
28	21			−0.685**			−0.525*
32	23	−0.457*		−0.423*			
33	23			−0.585**			
36	24		−0.414*				
41	24		−0.476*				
42	23		−0.423*				
43	24					0.411*	
45	24					0.430*	

*Values are Pearson correlations. S/E, Surgency/Extraversion; NE, Negative Emotionality; O/R, Orienting/Regulation. *p < 0.05; **p < 0.01.*

FIGURE 4 | Placement of channels in which temperament is correlated with hemoglobin activity. (A) Channels in which the maximum change in amplitude from baseline of [oxyHb] is significantly correlated with S/E, NE, and O/R are indicated in light gray, patterned, and dark gray, respectively. Positive correlations are indicated with a square marker, and negative correlations are indicated with a circular marker. **(B)** Channels in which maximum change in amplitude from baseline of [deoxyHb] is significantly correlated with temperament factors.

Similarly, NE was negatively correlated with oxyHb in channel 36, $r(22) = −0.414$, $p = 0.044$, channel 41, $r(22) = −0.476$, $p = 0.019$, and channel 42, $r(21) = −0.423$, $p = 0.044$. O/R was negatively correlated with oxyHb activity in channel 27, $r(19) = −0.525$, $p = 0.015$, channel 28, $r(19) = −0.685$, $p = 0.001$, channel 32, $r(21) = −0.423$, $p = 0.044$, and channel 33, $r(21) = −0.585$, $p = 0.003$. Infants with lower NE scores and lower O/R scores, respectively, showed greater activation in response to happy faces than did infants with higher NE and O/R scores. In one channel (32), oxyHb activity was correlated with both S/E and O/R. DeoxyHb activity was correlated with S/E in channel 27, $r(19) = −0.500$, $p = 0.021$; with NE in channel 43, $r(22) = 0.411$, $p = 0.046$, and channel 45, $r(22) = −0.430$, $p = 0.036$; and with O/R in channel 28, $r(19) = −0.525$, $p = 0.014$.

The approximate locations of the correlated channels are shown in **Figure 4**. None of the correlations between temperament and concentration was significant at the level required to correct for multiple comparisons ($p < 0.0008$). However, the correlated channels are clustered by temperament group, suggesting a consistent pattern of activation. The four channels in which oxyHb activity is correlated with O/R are adjacent to one another, as are the three channels in which oxyHb is correlated with NE, as shown in **Figure 4A**. The two channels with significant correlations between S/E and oxyHb are nearly adjacent.

The channels in which deoxyHb concentration is correlated with temperament are shown in **Figure 4B**. The channels where deoxyHb activity is correlated with S/E and O/R overlap with regions where oxyHb activity is correlated with temperament, as would be expected. The channels in which NE is positively correlated with deoxyHb activity are located along the brow.

The three higher-order temperament factors (S/E, NE, and O/R) are, in theory, orthogonal and uncorrelated (Gartstein and Rothbart, 2003). However, because S/E and O/R were correlated for our data, we conducted partial correlations to examine the independent contributions of each factor. We found that, when controlling for S/E, oxyHb was significantly correlated with O/R in channel 28, $r(18) = −0.607$, $p = 0.005$, and channel 33, $r(20) = −0.532$, $p = 0.011$. The correlation approached significance in channel 44, $r(21) = −0.403$, $p = 0.056$. DeoxyHb was significantly correlated with O/R in channel 28, $r(18) = −0.506$, $p = 0.023$. When controlling for O/R, neither oxyHb nor deoxyHb was significantly correlated with S/E.
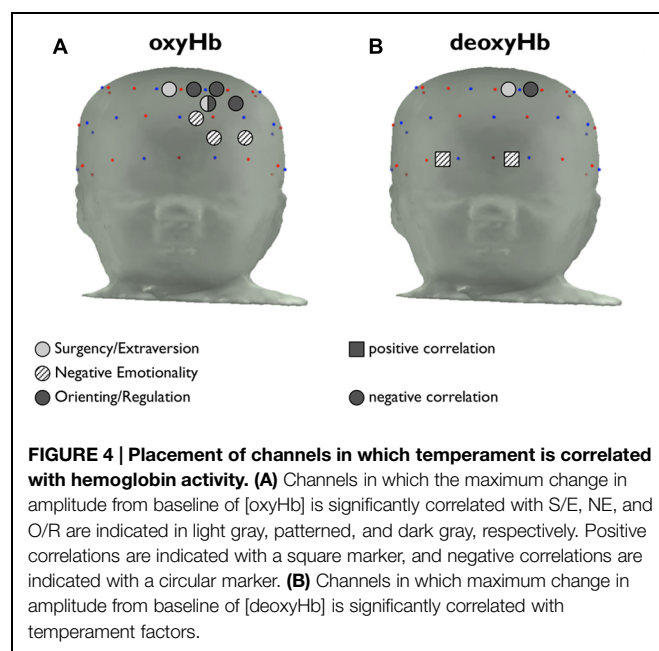
Frontal Asymmetry Measured by fNIRS

We hypothesized that fNIRS imaging could detect an asymmetry effect in the prefrontal response to happy face stimuli. Individuals with a greater proclivity for 'approach' behaviors show relatively greater left frontal activation, while individuals who tend to display 'withdrawal' behaviors show relatively greater right frontal activation. Based on these established findings in both infants and adults (Davidson and Fox, 1982; Davidson, 1993; Fox et al., 2001; Coan and Allen, 2003), we hypothesized that subjects with higher S/E temperament scores would show greater relative left frontal activation. Subjects with higher NE scores would show greater relative right frontal activation.

Negative Emotionality

We first tested for a hemispheric effect between the high and low NE groups. We conducted an RM-ANOVA with the two hemispheres (left and right) as the within-subjects factor, and the temperament group (high or low NE) as the between-subjects factor. To calculate the oxyHb activation in each hemisphere, we averaged the activity in two channels on the left side (36 and 41) and two channels on the right side (35 and 39) of the prefrontal panel of the fNIRS probe. These channels were selected because they were located in symmetrical positions on the medial region of the prefrontal probe, an area commonly analyzed for

hemispheric comparisons using fNIRS and EEG (Tuscan et al., 2013). Activity in channels 36 and 41 was significantly correlated with NE, so we expected to see an effect of temperament in this analysis.

There was a significant main effect of NE group, $F(1,22) = 4.80$, $p = 0.039$. The low-NE group had greater overall activation to happy faces ($M = 0.070$, SD = 0.137) compared to the high-NE group ($M = -0.061$, SD = 0.154) in the four channels of interest. This suggests that infants with less negative temperament are more responsive to images of happy faces than are infants with more negative temperament. There was no significant main effect of hemisphere, $F(1,22) = 1.11$, $p > 0.05$; the overall activation was not significantly different between the left hemisphere and the right hemisphere.

The main effect of group was modified by a Group × Hemisphere interaction, $F(1,22) = 4.75$, $p = 0.040$.

A priori pairwise comparisons testing hemispheric differences between the two temperament groups found that the low NE group and high NE group did not differ in right hemisphere activation, $t(22) = 0.487$, $p = 0.631$, but the two groups did differ significantly in left hemisphere activation, $t(22) = 3.614$, $p = 0.002$. Low-NE infants showed preferential activation in response to happy faces in the left hemisphere, while high-NE infants showed less overall activation in both hemispheres. The asymmetry effect is shown in **Figure 5**.

Surgency/Extraversion

We also conducted an RM-ANOVA with S/E group as the between-subjects factor. There were no significant main effects of S/E group, $F(1,22) = 0.354$, $p = 0.56$, or of hemisphere, $F(1,22) = 1.03$, $p = 0.32$, and no significant Group × Hemisphere interaction, $F(1,22) = 2.69$, $p = 0.12$.
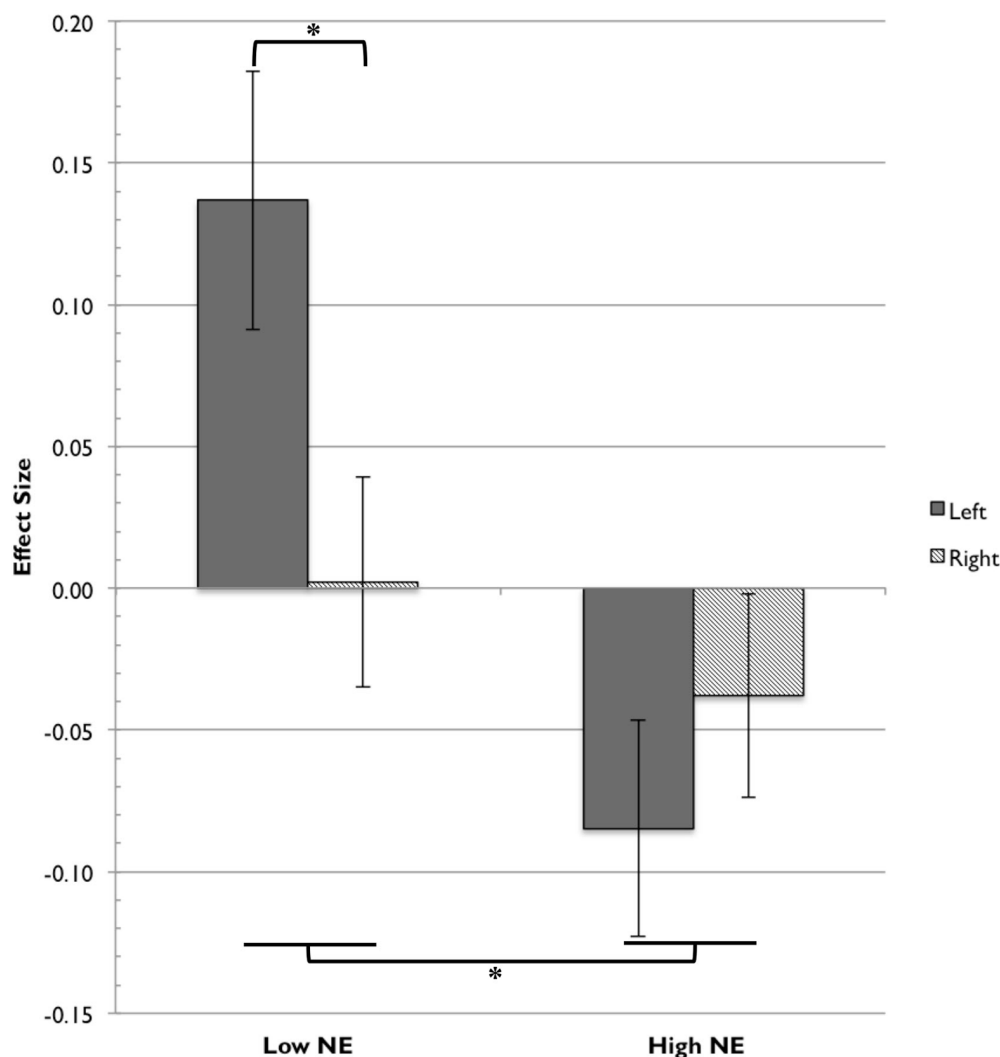

FIGURE 5 | Frontal fNIRS asymmetry in low NE and high NE groups. For the low NE group, the left hemisphere (channels 36, 41) showed significantly greater activation relative to the right hemisphere (channels 35, 39). The low NE group showed significantly greater activation overall than the high NE group. Asterisks indicate significance at the level of $p < 0.05$.

Discussion

Key Findings

The objectives of the present study were to characterize the prefrontal hemodynamic response of 7-month-old infants to happy face stimuli, to analyze the relation between infants' temperament and their brain responses to happy face stimuli, and to examine the capabilities of fNIRS methodology to provide information about frontal asymmetry. We showed that happy face stimuli elicited significant changes relative to baseline in oxyHb and deoxyHb concentrations in three channels (25, 34, and 46). Based on the time course of these responses, we selected a time range of interest of $t = 0$–10 s, with $t = 0$ s corresponding to the start of stimulus presentation. Through correlational analyses, we showed that the maximum change in both oxyHb and deoxyHb concentrations in response to happy faces was significantly correlated with S/E, NE, and O/R temperaments in channels overlying the left PFC. However, when controlling for the O/R factor, S/E is not correlated with oxy- or deoxyHb activity in any of the channels.

Further, we demonstrated that fNIRS can be used to study frontal asymmetry, a direction of analysis that is well-established in EEG literature but relatively unexplored in NIRS. We showed that there was a main effect of NE temperament group (low and high) modified by a Group × Hemisphere interaction. The low-NE infants preferentially activated the left hemisphere in response to happy faces. High-NE infants did not show this lateralization effect, and the overall activation for low-NE infants was higher than for high-NE infants. We had hypothesized that high-S/E infants would show relatively greater left frontal activation, but there were no notable effects of S/E group and hemisphere.

Hemodynamic Differences from Baseline

The typical NIRS response in adults shows an increase in oxyHb, and a corresponding decrease in deoxyHb that is relatively smaller in magnitude. This response at a given channel is thought to indicate an increase in brain activation in the cortical region underlying the channel. The activation in channels 25, 34, and 46 did not follow this pattern of typical activation; channels 25 and 46 showed a significant decrease in oxyHb concentration, and channel 34 showed a significant increase in deoxyHb concentration. Atypical hemodynamic patterns in infant brains (such as simultaneous increases in all three chromophores) have been described previously, but they are believed to be the result of immature neurovascular coupling (Lloyd-Fox et al., 2010). It has also been proposed that a decrease in oxyHb and corresponding increase in deoxyHb would indicate local decreased neural activity as compared to baseline, just as a decrease in fMRI BOLD signal is considered to represent brain deactivation (Fransson et al., 1999; Sakatani et al., 2006), or that this inverse pattern indicates activation in an adjacent brain region ("focal activation/surround deactivation"; Pfurtscheller et al., 2010).

Because channels 25, 34, and 46 are not adjacent to one another, it is difficult to draw conclusions about the implications of these activations. It is noteworthy, however, that channels 25 and 46 are located on the edges of the region in which oxyHb activity is correlated with temperament factors (channels 26, 27, 28, 32, 33, 36, 41, and 42). This suggests that the broad region could be involved with emotional face processing, with some areas activated across all subjects and other areas differentially activated as a result of individual differences — in this case, infant temperament.

It is unclear whether the observed responses were specific to happy face stimuli. Future research will compare the hemodynamic responses to happy face stimuli with the responses to other emotional expressions (fearful, angry, and neutral).

Previous studies in fNIRS have found a significant neural response to facial stimuli in the medial prefrontal region. In one fNIRS study that analyzed the oxyHb response to happy face stimuli, there was significant activation relative to neutral stimuli in a single channel overlying the medial PFC, demonstrating that there is a measurable response to happy face stimuli in this region (Minagawa-Kawai et al., 2009). Similarly, a second study showed a greater response to smiling faces than to neutral faces in six frontal channels, with three channels over the right frontal cortex showing a greater oxyHb response to happy faces, and three channels over the left frontal cortex showing a greater (more negative) deoxyHb response to happy faces (Fox et al., 2013). Based on these previous findings, we were surprised that we did not observe greater prefrontal activation to happy face stimuli. However, because we chose to analyze the response to facial stimuli relative to a non-face baseline, rather than to a neutral face baseline, the results are not directly comparable.

Temperament Correlated with OxyHb and DeoxyHb Activity

As hypothesized, temperament was correlated with hemodynamic activity in regions of the PFC. We chose not to correct these correlational results for multiple comparisons, given the exploratory nature of this analysis. However, the finding that local hemoglobin activation in these channels correlates with temperament is particularly compelling due to the clustering of channels by temperament group. The correlated activity is found in contiguous regions, rather than in scattered or isolated channels.

We would expect that activation to happy faces would be negatively correlated with NE. This study shows that, in these channels, infants with greater negative affect show less activation to happy face stimuli, while infants with less negative affect show greater neural activation. These results suggest that low-NE infants tend to be more responsive to happy faces.

It is more surprising that S/E was also negatively correlated with the oxyHb response to happy faces. We hypothesized that infants with higher S/E scores — infants who are more approach oriented — would show greater activation to happy faces. When we assumed that the three temperament factors were independent (Gartstein and Rothbart, 2003), this hypothesis was unconfirmed. However, when we controlled for O/R, there were no channels where S/E was correlated with oxyHb or deoxyHb activity. This suggests that the observed negative correlation between S/E and temperament was primarily driven by the O/R factor.

We noted that the channels in which NE was correlated with oxyHb activity were medial relative to the channels in which O/R was correlated with oxyHb activity. Previous work using fNIRS, fMRI, and other brain imaging technologies has shown a functional division in the infant PFC between the mPFC and lPFC. The mPFC has reciprocal connections with the amygdala, hippocampus, and temporal cortex — regions implicated in emotion, memory, and sensory processing — while the lPFC has reciprocal connections with motor regions, the cingulate cortex, and the parietal cortex. Broadly, the mPFC is involved with emotional processes, and the lPFC is involved with cognitive processes (Grossmann, 2013). The present findings are consistent with this distinction. The temperament factor of NE measures infants' emotional proclivities (including fear and sadness), and the channels where NE is correlated with brain activity are relatively medial. The O/R factor measures attentional and regulatory tendencies, which would be consistent with cognitive processing in the left lateral PFC, and the channels where O/R is correlated with oxyHb activity are relatively lateral.

All of the channels except one (channel 26, correlated with S/E) are located in the left hemisphere. As described in the introduction, activation of the left hemisphere has been associated with motivation to approach (versus withdraw). We would expect that non-threatening face stimuli, such as the images used in this study, would elicit an 'approach' response in some individuals. It makes sense, then, that the subjects' temperamental biases in their responses to happy face stimuli appeared to be driven by activation in the left hemisphere.

The correlations between temperament factors and deoxyHb activity are difficult to interpret. Most fNIRS studies of infant hemodynamic activity report the oxyHb responses because these data have higher signal-to-noise ratio than either deoxyHb or totalHb responses, and studies that do report deoxyHb activity have found inconsistent trends (Lloyd-Fox et al., 2010). However, a more complete understanding of infant metabolic activity requires that both oxyHb and deoxyHb activity be reported. In this study, the deoxyHb activation is reasonably consistent with the oxyHb activation. The fact that fewer channels show significant correlations with temperament could be due to the smaller amplitude of the deoxyHb response; there is less variation in the maximum amplitude of deoxyHb activity and thus fewer meaningful correlations.

Overall, these results provide additional evidence that infant temperament, as measured by the Infant Behavior Questionnaire, is associated with individual differences in the neural response to emotional faces. A previous study showed that NE correlated positively with a greater Nc to happy faces (Martinos et al., 2012), whereas our results demonstrate that lower NE scores are associated with greater activation to the happy face stimuli. Because the present study did not assess infants' allocation of attention to the face stimuli, these results cannot be directly compared to previous studies (de Haan et al., 2004; Martinos et al., 2012). However, the accumulating evidence suggests that the association between infant temperament and individual differences in emotional face processing is a fruitful topic for further investigation.

Frontal Asymmetry in fNIRS Response to Happy Face Stimuli

In analyzing the effects of hemisphere and temperament on oxyHb activity in these data, we found that there was a frontal asymmetry effect detectable with fNIRS. There was no main effect of hemisphere on oxyHb activation, but there was an interaction with NE group in the expected direction. The low-NE infants preferentially activated the left hemisphere, which is consistent with previous findings that the left hemisphere is preferentially activated during approach behaviors (Davidson, 1993), but there was no difference in hemispheric activation for high-NE infants. In response to happy faces, low-NE infants seem to demonstrate an asymmetrical approach response, but this lateralization effect is absent in high-NE infants.

The asymmetry effect measured in the left and right OFC is likely to be a response elicited by the happy face stimuli (state-dependent) that is stronger in infants with low-NE temperament than those with high-NE temperament (trait-dependent). An investigation of regional specificity in EEG asymmetry found that the effects in frontopolar recordings were more transient and were assumed to reflect OFC activity. In contrast, asymmetries calculated for dorsolateral, temporal, and parietal areas were more stable over time (Papousek and Schulter, 1998). Previous research confirms that emotional face stimuli elicit state-dependent frontal EEG asymmetry, even in infants (Davidson and Fox, 1982; Ekman et al., 1990). It is thought that asymmetrical cortical activation might be caused by differential inputs to the two hemispheres of the cortex from subcortical structures, particularly the amygdala (Kagan and Snidman, 2004). The OFC is strongly innervated by the amygdala, and it is reasonable that it would receive transient state-dependent inputs to each hemisphere.

What neurobiological activity might be driving the asymmetrical left-hemisphere activation in low-NE infants? Traditionally, frontal asymmetry is measured using EEG. In this measure, increased cortical activation is associated with desynchronization of the neural activity that produces the alpha wave, resulting in reduced power in the alpha frequency band (Davidson, 2004; Kagan and Snidman, 2004). Studies that simultaneously record electrical activity (using EEG) and metabolic activity (using fMRI or PET) have shown that alpha power is inversely correlated with high metabolism in several cortical brain regions, providing justification for using alpha power and hemodynamic activity as alternative measures of brain activation (Oakes et al., 2004). In the present study, the increased oxyHb concentrations in the left prefrontal area, relative to the right, suggest that left-hemispheric neural activation is relatively greater. As the neurons fire at a greater rate and consume greater amounts of oxygen, greater amounts of oxyHb are delivered to the local cortical area, and are detected by the fNIRS recording.

This finding provides insight into the development of hemispheric asymmetry in emotional face processing. Previous work has shown that adults' responses to emotional faces

are lateralized (Davidson, 1993; Fusar-Poli et al., 2009), and the present experiment provides evidence that this laterality is present as early as 7 months of age, and that it can be measured with fNIRS. Because of the spatial resolution of fNIRS, this technique could prove useful in parsing out the functional divisions of the PFC, and the development of these functions over the first years of life. Furthermore, studying the typical development of hemispheric asymmetry in emotional face processing will reveal information about the atypical development of these processes (Davidson, 1993).

Limitations and Future Directions

In the current study, we used a parent report measure of infant temperament. Gartstein and Rothbart (2003) have discussed at length the limitations of various temperament assessments: parent report measures are less controlled than laboratory observation, but on the other hand, the novel setting of the laboratory could influence infants' behavior (Putnam et al., 2008). Future studies should employ both parent-report survey data and laboratory observational assessment of temperament. Furthermore, we examined only the hemodynamic responses to happy face stimuli. Future analyses of the neural responses to neutral expressions and to other emotional expressions would provide additional insight into the interaction between temperament and prefrontal brain activity. Finally, our analysis included only one age of participant. The investigation of

whether, when, and how individual temperaments change over time, from infancy into adulthood, is an important direction of future study, especially as it informs our understanding of how anxiety, depression, and social pathologies develop. Measuring the prefrontal responses to emotional faces in infants of other ages would provide information about how neural activity develops over the first year of life and would provide useful context for the interpretation of our results.

Author Contributions

Acknowledgments

References

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* 1, 21–62. doi: 10.1177/1534582302001001003

Baehr, E., Peter Rosenfeld, J., Baehr, R., and Earnest, C. (1998). Comparison of two EEG asymmetry indices in depressed patients vs. normal controls. *Int. J. Psychophysiol.* 31, 89–92. doi: 10.1016/S0167-8760(98)00041-5

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 561–572. doi: 10.1098/rstb.2002.1220

Blasi, A., Fox, S., Everdell, N., Volein, A., Tucker, L., Csibra, G., et al. (2007). Investigation of depth dependent changes in cerebral haemodynamics during face perception in infants. *Phys. Med. Biol.* 52, 6849–6864. doi: 10.1088/0031-9155/52/23/005

Caspi, A. (2000). The child is father of the man: personality continuities from childhood to adulthood. *J. Pers. Soc. Psychol.* 78, 158–172. doi: 10.1037//0022-3514.78.1.158

Coan, J. A., and Allen, J. J. B. (2003). "The state and trait nature of frontal EEG asymmetry in emotion," in *The Asymmetrical Brain*, eds K. Hugdahl and R. J. Davidson (Cambridge, MA: MIT Press).

Cooper, R. J., Selb, J., Gagnon, L., Phillip, D., Schytz, H. W., Iversen, H. K., et al. (2012). A systematic comparison of motion artifact correction techniques for functional near-infrared spectroscopy. *Front. Neurosci.* 6:147. doi: 10.3389/fnins.2012.00147

Davidson, R. J. (1993). Cerebral asymmetry and emotion: conceptual and methodological conundrums. *Cogn. Emot.* 7, 115–138. doi: 10.1080/02699939308409180

Davidson, R. J. (2004). What does the prefrontal cortex "do" in affect: perspectives on frontal EEG asymmetry research. *Biol. Psychol.* 67, 219–234. doi: 10.1016/j.biopsycho.2004.03.008

Davidson, R. J., and Fox, N. A. (1982). Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants. *Science* 218, 1235–1237. doi: 10.1126/science.7146906

de Haan, M., Belsky, J., Reid, V., Volein, A., and Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *J. Child Psychol. Psychiatry* 45, 1209–1218. doi: 10.1111/j.1469-7610.2004.00320.x

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S., Wray, S., and Wyatt, J. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Phys. Med. Biol.* 33, 1433–1442. doi: 10.1088/0031-9155/33/12/008

Duncan, A., Meek, J. H., Clemence, M., Elwell, C. E., Tyszczuk, L., Cope, M., et al. (1995). Optical pathlength measurements on adult head, calf and forearm and the head of the newborn infant using phase resolved optical spectroscopy. *Phys. Med. Biol.* 40, 295–304. doi: 10.1088/0031-9155/40/2/007

Ekman, P., and Davidson, R. J. (1993). Voluntary smiling changes regional brain activity. *Psychol. Sci.* 4, 342–345. doi: 10.1111/j.1467-9280.1993.tb00576.x

Ekman, P., Davidson, R. J., and Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology. II. *J. Pers. Soc. Psychol.* 58, 342–353. doi: 10.1111/j.1467-9280.1993.tb00576.x

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., et al. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron* 44, 1043–1055. doi: 10.1016/j.neuron.2004.12.006

Fox, N. A., Henderson, H. A., Rubin, K. H., Calkins, S. D., and Schmidt, L. A. (2001). Continuity and discontinuity of behavioral inhibition and exuberance: psychophysiological and behavioral influences across the first four years of life. *Child Dev.* 72, 1–21. doi: 10.1111/1467-8624.00262

Fox, S. E., Wagner, J. B., Shrock, C. L., Tager-Flusberg, H., and Nelson, C. A. (2013). Neural processing of facial identity and emotion in infants at high-risk for autism spectrum disorders. *Front. Hum. Neurosci.* 7:89. doi: 10.3389/fnhum.2013.00089

Fransson, P., Krüger, G., Merboldt, K. D., and Frahm, J. (1999). MRI of functional deactivation: temporal and spatial characteristics of oxygenation-sensitive responses in human visual cortex. *Neuroimage* 9, 611–618. doi: 10.1006/nimg.1999.0438

Fusar-Poli, P., Placentino, A., Carletti, F., Allen, P., Landi, P., Abbamonte, M., et al. (2009). Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neurosci. Lett.* 452, 262–267. doi: 10.1016/j.neulet.2009.01.065

Gartstein, M. A., and Rothbart, M. K. (2003). Studying infant temperament via the Revised Infant Behavior Questionnaire. *Infant Behav. Dev.* 26, 64–86. doi: 10.1016/S0163-6383(02)00169-8

Grossmann, T. (2013). Mapping prefrontal cortex functions in human infancy. *Infancy* 18, 303–324. doi: 10.1111/infa.12016

Grossmann, T., Striano, T., and Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: a neurobehavioral study. *Brain Cogn.* 64, 30–41. doi: 10.1016/j.bandc.2006.10.002

Hagemann, D. (2004). Individual differences in anterior EEG asymmetry: methodological problems and solutions. *Biol. Psychol.* 67, 157–182. doi: 10.1016/j.biopsycho.2004.03.006

Hagemann, D., Naumann, E., Lürken, A., Becker, G., Maier, S., and Bartussek, D. (1999). EEG asymmetry, dispositional mood and personality. *Pers. Indiv. Differ.* 27, 541–568. doi: 10.1016/S0191-8869(98)00263-3

Hoehl, S., Palumbo, L., Heinisch, C., and Striano, T. (2008). Infants' attention is biased by emotional expressions and eye gaze direction. *Neuroreport* 19, 579–582. doi: 10.1097/WNR.0b013e3282f97897

Jessen, S., and Grossmann, T. (2015). Neural signatures of conscious and unconscious emotional face processing in human infants. *Cortex* 64, 260–270. doi: 10.1016/j.cortex.2014.11.007

Kagan, J., and Snidman, N. C. (2004). *The Long Shadow of Temperament.* Cambridge, MA: Belknap Press of Harvard University Press.

Kandler, C., Riemann, R., and Angleitner, A. (2013). Patterns and sources of continuity and change of energetic and temporal aspects of temperament in adulthood: a longitudinal twin study of self- and peer reports. *Dev. Psychol.* 49, 1739–1753. doi: 10.1037/a0030744

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev.neuro.23.1.155

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., and Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child Dev.* 78, 232–245. doi: 10.1111/j.1467-8624.2007.00994.x

Leppänen, J. M., and Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nat. Rev. Neurosci.* 10, 37–47. doi: 10.1038/nrn2554

Leppänen, J. M., and Nelson, C. A. (2012). Early development of fear processing. *Curr. Dir. Psychol. Sci.* 21, 200–204. doi: 10.1177/0963721411435841

Lloyd-Fox, S., Blasi, A., and Elwell, C. E. (2010). Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neurosci. Biobehav. Rev.* 34, 269–284. doi: 10.1016/j.neubiorev.2009.07.008

Lloyd-Fox, S., Blasi, A., Elwell, C. E., Charman, T., Murphy, D., and Johnson, M. H. (2013). Reduced neural sensitivity to social stimuli in infants at risk for autism. *Proc. R. Soc. B Biol. Sci.* 280:20123026. doi: 10.1098/rspb.2012.3026

Lloyd-Fox, S., Blasi, A., Volein, A., Everdell, N., Elwell, C. E., and Johnson, M. H. (2009). Social perception in infancy: a near infrared spectroscopy study. *Child Dev.* 80, 986–999. doi: 10.1111/j.1467-8624.2009.01312.x

Lloyd-Fox, S., Richards, J. E., Blasi, A., Murphy, D. G. M., Elwell, C. E., and Johnson, M. H. (2014). Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants. *Neurophotonics* 1:025006. doi: 10.1117/1.NPh.1.2.025006

Martinos, M., Matheson, A., and de Haan, M. (2012). Links between infant temperament and neurophysiological measures of attention to happy and fearful faces. *J. Child Psychol. Psychiatry* 53, 1118–1127. doi: 10.1111/j.1469-7610.2012.02599.x

Minagawa-Kawai, Y., Matsuoka, S., Dan, I., Naoi, N., Nakamura, K., and Kojima, S. (2009). Prefrontal activation associated with social attachment: facial-emotion recognition in mothers and infants. *Cereb. Cortex* 19, 284–292. doi: 10.1093/cercor/bhn081

Molavi, B., and Dumont, G. A. (2012). Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiol. Meas.* 33, 259–270. doi: 10.1088/0967-3334/33/2/259

Nakato, E., Otsuka, Y., Kanazawa, S., Yamaguchi, M. K., and Kakigi, R. (2011). Distinct differences in the pattern of hemodynamic response to happy and angry facial expressions in infants — A near-infrared spectroscopic study. *Neuroimage* 54, 1600–1606. doi: 10.1016/j.neuroimage.2010.09.021

Nelson, C. A., and Dolgin, K. G. (1985). The generalized discrimination of facial expressions by seven-month-old infants. *Child Dev.* 56, 58–61. doi: 10.2307/1130173

Oakes, T. R., Pizzagalli, D. A., Hendrick, A. M., Horras, K. A., Larson, C. L., Abercrombie, H. C., et al. (2004). Functional coupling of simultaneous electrical and metabolic activity in the human brain. *Hum. Brain Mapp.* 21, 257–270. doi: 10.1002/hbm.20004

Otsuka, Y., Nakato, E., Kanazawa, S., Yamaguchi, M. K., Watanabe, S., and Kakigi, R. (2007). Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *Neuroimage* 34, 399–406. doi: 10.1016/j.neuroimage.2006.08.013

Papousek, I., and Schulter, G. (1998). Different temporal stability and partial independence of EEG asymmetries from different locations: implications for laterality research. *Int. J. Neurosci.* 93, 87–100. doi: 10.3109/00207459808986415

Parade, S. H., and Leerkes, E. M. (2008). The reliability and validity of the Infant Behavior Questionnaire-Revised. *Infant Behav. Dev.* 31, 637–646. doi: 10.1016/j.infbeh.2008.07.009

Pfurtscheller, G., Bauernfeind, G., Wriessnegger, S. C., and Neuper, C. (2010). Focal frontal (de)oxyhemoglobin responses during simple arithmetic. *Int. J. Psychophysiol.* 76, 186–192. doi: 10.1016/j.ijpsycho.2010.03.013

Putnam, S. P., Helbig, A. L., Gartstein, M. A., Rothbart, M. K., and Leerkes, E. (2013). Development and assessment of short and very short forms of the Infant Behavior Questionnaire-Revised. *J. Pers. Assess.* 96, 445–458. doi: 10.1080/00223891.2013.841171

Putnam, S. P., Rothbart, M. K., and Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant Child Dev.* 17, 387–405. doi: 10.1002/icd.582

Sakatani, K., Yamashita, D., Yamanaka, T., Oda, M., Yamashita, Y., Hoshino, T., et al. (2006). Changes of cerebral blood oxygenation and optical pathlength during activation and deactivation in the prefrontal cortex measured by time-resolved near infrared spectroscopy. *Life Sci.* 78, 2734–2741. doi: 10.1016/j.lfs.2005.10.045

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., and Kinney, L. (1992). Psychometric properties of resting anterior EEG asymmetry: temporal stability and internal consistency. *Psychophysiology* 29, 576–592. doi: 10.1111/j.1469-8986.1992.tb02034.x

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 168, 242–249. doi: 10.1016/j.psychres.2008.05.006

Tranel, D., Bechara, A., and Denburg, N. L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision-making, and emotional processing. *Cortex* 38, 589–612. doi: 10.1016/S0010-9452(08)70024-8

Tuscan, L.-A., Herbert, J. D., Forman, E. M., Juarascio, A. S., Izzetoglu, M., and Schultheis, M. (2013). Exploring frontal asymmetry using functional near-infrared spectroscopy: a preliminary study of the effects of social anxiety during interaction and performance tasks. *Brain Imaging Behav.* 7, 140–153. doi: 10.1007/s11682-012-9206-z

Vanderwert, R. E., and Nelson, C. A. (2014). The use of near-infrared spectroscopy in the study of typical and atypical development. *Neuroimage* 85, 264–271. doi: 10.1016/j.neuroimage.2013.10.009

Watanabe, H., Homae, F., Nakano, T., and Taga, G. (2008). Functional activation in diverse regions of the developing brain of human infants. *Neuroimage* 43, 346–357. doi: 10.1016/j.neuroimage.2008.07.014

Dopamine D4 receptor polymorphism and sex interact to predict children's affective knowledge

Sharon Ben-Israel[1,2], Florina Uzefovsky[1,3], Richard P. Ebstein[4] and Ariel Knafo-Noam[1]*

[1] Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel, [2] Department of Psychology, Academic College of Tel Aviv-Yaffo, Tel Aviv, Israel, [3] Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, UK, [4] Department of Psychology, National University of Singapore, Singapore, Singapore

Affective knowledge, the ability to understand others' emotional states, is considered to be a fundamental part in efficient social interaction. Affective knowledge can be seen as related to cognitive empathy, and in the framework of theory of mind (ToM) as affective ToM. Previous studies found that cognitive empathy and ToM are heritable, yet little is known regarding the specific genes involved in individual variability in affective knowledge. Investigating the genetic basis of affective knowledge is important for understanding brain mechanisms underlying socio-cognitive abilities. The 7-repeat (7R) allele within the third exon of the dopamine D4 receptor gene (*DRD4-III*) has been a focus of interest, due to accumulated knowledge regarding its relevance to individual differences in social behavior. A recent study suggests that an interaction between the *DRD4-III* polymorphism and sex is associated with cognitive empathy among adults. We aimed to examine the same association in two childhood age groups. Children (*N* = 280, age 3.5 years, *N* = 283, age 5 years) participated as part of the Longitudinal Israel Study of Twins. Affective knowledge was assessed through children's responses to an illustrated story describing different emotional situations, told in a laboratory setting. The findings suggest a significant interaction between sex and the *DRD4-III* polymorphism, replicated in both age groups. Boy carriers of the 7R allele had higher affective knowledge scores than girls, whereas in the absence of the 7R there was no significant sex effect on affective knowledge. The results support the importance of *DRD4-III* polymorphism and sex differences to social development. Possible explanations for differences from adult findings are discussed, as are pathways for future studies.

Keywords: dopamine, *DRD4*, cognitive empathy, affective perspective taking, gender, affective knowledge

Introduction

Affective knowledge, the ability to understand others' emotional states (e.g., Knafo et al., 2009), is important for children's social functioning, and for the ability to communicate, cooperate, and cope with complex social interactions (Denham, 1986; Bauminger, 2002; Walker, 2005; Knafo et al., 2011b; Garner and Waajid, 2012).

Affective knowledge has been linked to *Empathy* – the tendency to share and understand the thoughts and feelings of others (Eisenberg and Strayer, 1990; Walter, 2012). Indeed, affective knowledge is often seen as an aspect of *Cognitive empathy*, the ability to recognize and understand what the other feels (Shamay-Tsoory et al., 2009). Cognitive empathy can also be defined in the framework of theory of mind (ToM) as, *Affective ToM*, the ability to represent and understand the affective mental states of others (Walter, 2012).

Affective knowledge, affective ToM, and cognitive empathy are all part of a network of interpersonal abilities that also includes *Affective empathy*, i.e., the ability to experience the emotion of the other while maintaining an emotional distinction between the self and the other (Decety and Lamm, 2006; de Vignemont and Singer, 2006). Studies in neuro-psychology and developmental psychology show that the two components are different, but not independent, aspects of the tendency to empathize (Singer, 2006; Volbrecht et al., 2007; Knafo et al., 2008a). In our literature review, therefore, we draw on evidence from the field of empathy research to understand the development of cognitive empathy and specifically affective knowledge.

Denham (1986) operationalized affective knowledge as composed of affective labeling (matching facial expressions to emotions) and affective perspective taking (matching a facial expression to someone based on their supposed emotional state). Garner and Waajid (2012) described affective knowledge as including awareness of emotion and knowledge of basic facial expressions (i.e., expression knowledge) and the situations that elicit emotions (situational knowledge). Similarly, many operationalizations of cognitive empathy (Matsumoto et al., 2000; Baron-Cohen et al., 2001), measured the ability to recognize an emotion from a facial expression, or predict an emotional response from a specific context. Based on these studies, our operationalization of affective knowledge (Knafo et al., 2009) combines aspects of these related approaches, by measuring attribution of emotional states, attribution of affective expressions, and matching facial expressions to attributed states.

Sex Differences in Empathy and in Affective Knowledge

Studies typically show significant sex differences in empathy, with women scoring higher than men (Hoffman, 1977; Baron-Cohen and Wheelwright, 2004). Disorders associated with deficits in empathy and particularly in ToM (e.g., autism, Asperger syndrome; Baron-Cohen et al., 1997) are more prevalent in men than in women (Baron-Cohen and Wheelwright, 2004). Disorders such as depression and anxiety, that tend to be more prevalent in women, are associated with higher empathy (Zahn-Waxler et al., 1991, 2008). However, the association between sex and empathy is unclear, and may be dependent on the method used to measure empathy (Eisenberg and Lennon, 1983). Sex differences favoring females usually emerge using questionnaire measures (e.g., Adams et al., 1979), while performance measures usually do not yield sex differences (Hoffman, 1977; Eisenberg and Lennon, 1983). With respect to ToM, most investigations

of individual differences in ToM did not specifically examine the issue of sex differences. Two exceptions are the studies of Charman et al. (2002) and of Walker (2005). Charman et al. (2002) investigated sex differences in false belief development and found a slight advantage for girls on false-belief task performance, an advantage that was only apparent in younger but not older children from a sample of 3–5 year-olds. In the second study, Walker (2005) showed that preschool girls are more competent than boys in ToM tasks, and that these sex differences were associated with peer-related social competence. Findings from studies that specifically focused on affective knowledge present a less than consistent picture (Hoffman, 1977; Eisenberg and Lennon, 1983; Gross and Ballif, 1991), with some studies showing a female advantage in affective knowledge tasks (Zahn-Waxler et al., 1984; Casey, 1993), and other studies finding no such effect (Cutting and Dunn, 1999).

These studies suggest that there are yet unanswered questions regarding the role of sex in explaining individual differences in empathy and affective knowledge, and that these differences may be important for understanding children's social functioning outcomes.

Heritability

There is much individual variation in empathic ability and specifically in affective knowledge – from extreme deficits, as in autism (Yirmiya et al., 1992) to differences in empathy and affective knowledge within the normal range seen in adults (Davis, 1980; Lawrence et al., 2004), children (Bryant, 1982; Denham, 1986; Knafo-Noam et al., 2015), and infants (Knafo et al., 2008a).

We are aware of a single study on affective ToM, in which 10-year-old's recognition of facial expressions showed substantial heritability (Lau et al., 2009). Additional studies on children (Zhu et al., 2010) and adult population (McKone and Palermo, 2010; Wilmer et al., 2010) have found significant heritability effects for face recognition. Although we are not aware of additional published studies concerning the heritability of affective knowledge among children, it is possible to learn about its development from research on empathy and ToM.

Empathy

To the best of our knowledge, up to the time of writing this paper, there have been only eight twin studies addressing the genetic and environmental influences on individual differences in empathy, and in all of these studies (except one study which had extremely high correlations for both MZ and DZ twins) a significant genetic influence on empathy was found to exist from early childhood onward (Knafo and Uzefovsky, 2013). Our meta-analysis (Knafo and Uzefovsky, 2013), found that heritability accounts for 35% of the variance in empathy (the influence of the shared environment was negligible, and the influence of the non-shared environment, which also includes measurement error, was estimated as explaining 63% of individual variability). Interestingly, when examined separately, cognitive and affective empathy were found to have different patterns of genetic and environmental effects. Heritability explained 30% of the variance in affective empathy and 26% of the variance in cognitive

empathy. Shared environment, estimated at 17%, explained individual variability in cognitive empathy only (the rest of the variance of both empathy facets was explained by non-shared environment and error, Knafo and Uzefovsky, 2013).

ToM

Similarly, there has been scant research estimating the genetic and environmental contributions to ToM. The available research shows moderate genetic as well as significant environmental contributions to individual differences in ToM (Ronald et al., 2005, 2006). A small study of 3.5 year-old twins showed a strong genetic effect on cognitive aspects of ToM (Hughes and Cutting, 1999). A larger study of 5-year-olds showed a genetic contribution to cognitive ToM, which overlapped with the genetic effects on language abilities (Hughes et al., 2005).

Taken together, the above-cited studies suggest that investigating the genetic basis of affective knowledge is a worthwhile endeavor.

Specific Genetic Effects

Studies have shown that the social hormones oxytocin (OT) and vasopressin (AVP) facilitate and promote social interactions by modulating dopaminergic activity in the brain reward system (Young and Wang, 2004; Skuse and Gallagher, 2009). Relatedly, many molecular genetic studies have so far focused on the oxytocin receptor *(OXTR)* gene, repeatedly showing that variations in the *OXTR* are significantly associated with measures of empathy (Chakrabarti et al., 2009; Rodrigues et al., 2009; Wu et al., 2012; Lucht et al., 2013; Uzefovsky et al., 2015), as well as with difficulties in empathy (Schneiderman et al., 2013) and with autism (Chakrabarti et al., 2009) and ToM (Wu and Su, 2014). For a recent review, see Israel et al. (2015).

An additional recent study has also suggested that the association between genetic variation in *OXTR* and prosocial behavior is mediated by perspective taking and empathic concern, and that this pathway is contingent on sex (Christ et al., 2015). This study has shown an interaction between *OXTR* polymorphisms and sex in predicting prosocial tendencies (empathic concern and perspective taking), which in turn predict prosocial behavior. The patterns of genotype effects on prosocial tendencies were different for males and females (Christ et al., 2015). For example, the interaction between sex and rs2254298 showed that males with at least one A allele have significantly lower perspective taking scores compared to males who are homozygous for G allele, while no significant genotype effect in this polymorphism was found for females.

As noted, OT and AVP, together with activity of dopaminergic receptors, comprise an integrated neuronal system of social cognition (Skuse and Gallagher, 2009). The dopaminergic system is a critical component of the "social brain" which comprises areas of the brain that are involved in social cognition and behavior (Skuse and Gallagher, 2009). Studies have suggested that dopamine is crucial to empathy-motivated prosocial behavior, which has evolutionary roots in offspring care (Preston, 2013). Thus, the dopaminergic system is another good candidate system for investigating the genetic underpinnings of empathy. Since most studies to date have focused on the role of *OXTR* in empathy, there is still a big gap in the understanding of the role of dopamine, even though it is a crucial part of the social brain. Indeed, one recent study conducted on Chinese college students has shown that genetic variation in the dopamine system made significant contributions to individual differences in facial expression recognition, and specifically in the recognition of disgust faces (Zhu et al., 2011).

DRD4 as a Candidate Gene for Affective Knowledge

DRD4 is a gene that encodes the D4 receptor of dopamine and is one of the most studied candidate genes in relation to social behavior (Kang et al., 2008; Zhong et al., 2010). The gene has a number of variations (polymorphisms). One of the most researched polymorphisms of the gene is found in exon 3, which is characterized by a repeat region of 48 bp (translated to 16 amino acids) that can be repeated 2 – 11 times. This polymorphism (*DRD4-III*) was associated with behaviors and traits that are related to empathy. An example of this is the association between *DRD4-III* polymorphism and altruistic behavior (Bachner-Melman et al., 2005), and the possible role of the polymorphism in ADHD (Faraone et al., 2001), a disorder that was shown to be related to ToM deficits and reduced empathy (Uekermann et al., 2010). Another interesting investigation focused on the function of *DRD4-III* polymorphism in representational ToM (RTM) – the ability to explicitly understand that other's mental states (beliefs, desires, knowledge) are person-specific representations of the world (Lackner et al., 2012). This study has suggested that variations in the *DRD4-III* may predict preschoolers' performance in RTM, showing that individuals with two shorter alleles (4 repeats or less) outperformed those with one or two longer alleles (6 repeats or more), (Lackner et al., 2010).

In addition to these investigations, the *DRD4-III* has recently become the focus of research into gene by environment interactions (GxE). Carriers of the 7 repeat allele (7R), the second most common repeat in Caucasian populations (Chang et al., 1996) are thought to be more sensitive to environmental influence (Belsky and Pluess, 2009). An example of this is the study by Knafo et al. (2011a), which focused, using a sample that partially overlaps with the current sample, on prosocial behavior. Prosocial behavior is relevant in this context because many (but not all) prosocial behaviors are associated with empathy. Positive parenting was positively related to prosocial behavior (as reported by the mother), and unexplained punishment related positively to experimentally-elicited self-initiated pro-social behavior, but only for children carriers of the 7R allele. Carriers of other alleles showed no association between parenting and behavior (Knafo et al., 2011a). In a subsample of that study, mother-reported negativity toward the child was negatively associated with observed empathic concern toward an examiner, again only among children carriers of the 7R allele (Knafo and Uzefovsky, 2013).

One recent study examined the association between *DRD4-III* polymorphism and empathy among adults (Uzefovsky et al., 2014). A significant gene by sex interaction was found for cognitive empathy (but not emotional empathy), and was

replicated in a second independent sample of adults. Specifically, it was found that women carriers of the 7R allele scored higher on cognitive empathy than women who were not carriers of the 7R allele, whereas for men, 7R carriers scored lower than non-carriers. This finding suggests that the polymorphism of *DRD4-III* is related to cognitive empathy and that sex differences are involved in this relationship.

The Current Research

As cognitive empathy has been shown to have a genetic basis in early childhood (Knafo et al., 2009), one could expect to replicate the findings by Uzefovsky et al. (2014) among children. However, it is important to note the differences in measure type (cognitive empathy questionnaire vs. a task measuring affective knowledge) as well as the notion that genetic effects are often age-specific (e.g., Choh et al., 2014). For example, an AVPR1a polymorphism is associated with generosity in both adults and children, but different alleles are responsible for this association in the two age groups (Knafo et al., 2008b; Avinun et al., 2011). It is therefore important to study the role of genes in children, to get a developmental perspective on the association between the *DRD4-III* genotype, cognitive empathy, and specifically affective knowledge.

The current study sought to expand the knowledge of the genetic basis of the cognitive component of empathy to the developmental context. Specifically, we examined the association between the *DRD4-III* polymorphism and affective knowledge in children 3.5 and 5 years of age. In view of previous findings, this study focused on the *DRD4-III* 7R allele, hypothesizing that the association between genotype and affective knowledge would be contingent on sex. We expected to replicate the findings in adults (Uzefovsky et al., 2014) whereby sex interacted with the *DRD4-III* polymorphism in the association with cognitive empathy. However, as we are looking at a different age group and a different phenotype of social cognition we did not make strong hypotheses, but rather were interested in investigating the associations between *DRD4-III*, sex, and affective knowledge in early childhood.

Materials and Methods

Participants

Families in this study were participants in the Longitudinal Study of Twins (LIST) that focuses on children's social development as influenced by genetics, abilities, and socialization (Knafo, 2006). All Hebrew-speaking families who were identified by the Israeli Ministry of the Interior as having twins born in 2004 and 2005, were contacted with mail surveys regarding children's development close to the twins' third birthday. See Avinun and Knafo (2013) for further details on the sample.

Families from the Greater Jerusalem area were invited to partake in an experimental session at the laboratory when twins reached 3.5 and 5 years of age. The laboratory session focused on evaluating empathy, pro-social behavior, cognitive abilities, and other social skills. In addition, DNA samples were taken from the twins and their parents, when parents' agreement was obtained.

The project was approved by the S. Herzog Hospital Institutional Review Board committee.

Of the initial lab sample we selected children with relevant data on both affective knowledge and DNA samples. Out of 447 individual participants from the first lab phase of the LIST, 128 were excluded because of a lack of DNA data, and 39 were excluded due to missing affective knowledge data. Similarly, out of 398 age 5 participants, 107 were excluded because of lack of DNA data, and eight were excluded due to missing affective knowledge data.

Therefore, the final sample from age 3.5 included 280 children (149 boys, 131 girls), aged 36–51 months ($M = 44.13$, SD $= 2.78$). The final age 5 sample included 283 children (149 boys, 134 girls), aged 59–71 months ($M = 61.73$, SD $= 2.15$). In total, 402 children participated at least once, of which 161 children participated in both phases.

Procedure

Around the age of 3.5, families (the twins and a parent, sometimes accompanied by another family member) arrived at the lab for an experimental session. In the lab they met two examiners, and each twin was asked to enter a separate testing room with one of the examiners. Visits were scheduled at a time when parents estimated children were likely to be at their best. Assessments of social and cognitive development skills were made through a number of tasks (Avinun and Knafo, 2013), separately for each twin. Most visits lasted for less than 2 h. During or prior to the visit, mothers filled out questionnaires which included questions on the pregnancy, twins' zygosity, twins' behavior, and demographic details, including socioeconomic status.

Toward children's fifth birthday all families from the twin sample living in the Greater Jerusalem area were invited to the lab again. The same laboratory procedure was performed when the twins were about 5 years old.

Measures

Affective knowledge was measured using the Jerusalem Story Test of Interpersonal Understanding (Knafo et al., 2011b). This instrument consists of an illustrated story that taps several socio-cognitive abilities with a single story narrative loaded with various emotional associations (Knafo et al., 2011b). We followed Denham's (1986, p. 194) recommendation, that to measure children's affective knowledge a measure has to be sensitive to the needs of "capturing young children's attention and of embedding tasks within an ongoing social context." We therefore measured affective knowledge with children's reactions to easy to understand situations, in a contextually valid setting of telling a story.

Stimuli and methods from existing relevant assessments (Denham, 1986; Ribordy et al., 1988) were integrated into the task. While reading the story, the experimenter asks the child predetermined questions tapping broad aspects of interpersonal understanding: affective knowledge, desire understanding, and false belief. Previously, measures from the test predicted children's observed prosocial behavior (Knafo et al., 2011b) and empathy (Knafo et al., 2009). In the current investigation we focus on affective knowledge, as measured by the illustrated

story assignment through three indices: emotion understanding, expression selection, and affective matching (see description below).

The story depicts emotional situations relevant to children's lives, involving story character Loulou (matched to the participating child's gender). Five situations (Ribordy et al., 1988) eliciting different emotions were examined: **happiness** (Loulou gets a long wished-for present), **fear** (a sudden darkness and a tree branch that appears like someone's hand touching the window), **anger** (Loulou is given a present in appreciation for his/her help, but then the giver changes his mind and requests the present back), **sadness** (Loulou is laughed-at by his/her friends after failing to play a game successfully), and **disgust** (Loulou finds a worm in his/her apple), (see **Figure 1**). The four negative emotions were used in the current investigations following up on Knafo et al. (2009), in consideration of the differences between perceiving negative emotions of the other (Roberts and Strayer, 1996; Simpson et al., 2003; Eisenberg et al., 2014) and perceiving other's positive emotions.

Three measures of affective knowledge were obtained:

(1) *Emotion understanding – Attribution of emotional states:*
    After hearing each emotionally relevant situation, the child was asked, "How does Loulou feel?" If the child did not respond correctly or failed to respond at all, the question was repeated, and three options (e.g., angry, sad, yuck/disgusted) were presented. Responses were scored on a five-point scale: 4 = correct unaided answer; 3 = initial response concerning an overall negative emotion, which was correctly identified when options were given; 2 = wrong answer on unaided question and a correct answer when options were given; 1 = general negative affective response on unaided question and wrong answer when options were given; and 0 = wrong answer on both questions.
    The emotion understanding measure was calculated for each participant as a sum across the four situations, resulting in a score that could range between 0 and 16.

(2) *Expression selection – Attribution of facial affective expressions*: Following the previous question the child was shown three facial expressions of Loulou (corresponding to the three options given in the earlier question). For example, "can you show me how Loulou looks now after the children laughed at him/her?" (see **Figure 2**). A correct (matching the situation) answer was rated 3. Incorrect responses received a score between 2 and 0, depending on the degree of dissimilarity; previous research has shown that emotional expressions can be ordered in a circular manner based on the similarity in their different facial configurations (anger, disgust, happiness, surprise, fear, sadness, anger; Susskind et al., 2007). Incorrect answers were therefore scored according to the degree of difference between the face chosen and the correct face on this circumplex (e.g., anger and disgust involve two relatively similar facial configurations and confusing them would incur the score 2; in contrast, confusing the very different expressions of fear and disgust would incur the score 0).
    The expression selection measure was calculated for each participant as a sum across the four situations, resulting in a score that could range from 0 to 12.

(3) *Affective matching – Matching facial expressions to attributed states:* Success in matching label to facial expression was measured by counting the times in which the child chose a facial expression matching the emotion he or she named in response to the verbal question ("How does Loulou feel?"), regardless of whether this choice was correct or not with regards to the situation. Correct match was coded as 1, and an incorrect match as 0. A sum score was calculated, and could range between 0 and 4.

Table 1 shows descriptive statistics for the three measures of affective knowledge. Each of the three indices of affective knowledge measures a different aspect of cognitive empathy. For example, a child could label an emotion correctly but fail to point to the right facial expression. In order to examine the factorial structure of all three measures we ran a factor analysis. As expected, since all three measures were designed to assess different aspects of affective knowledge, they were all significantly inter-correlated ($p < 0.01$, see correlations in **Table 2**) and loaded on a single factor. In the 3.5 year-olds sample, the factor accounted for 52.60% of the variance, with loadings ranging from

Fear
Sudden dark and a tree branch that appears like someone's hand touching the window.

Anger
Loulou is given a present in appreciation for his/her help, but then the giver changes their mind and requests the present back.

Sadness
Loulou is laughed at by his/her friends after failing to play a game successfully.

Disgust
Loulou finds a worm in his/her apple.
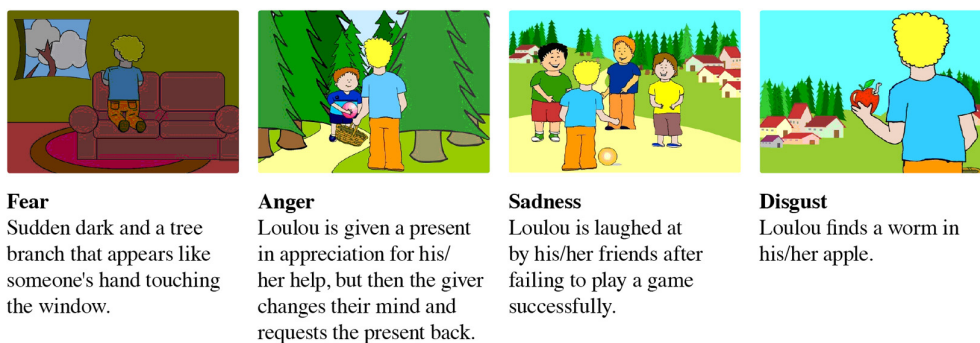
FIGURE 1 | An illustration of the four emotional situations described in the story.

Can you show me how Loulou looks now after the children laughed at him?
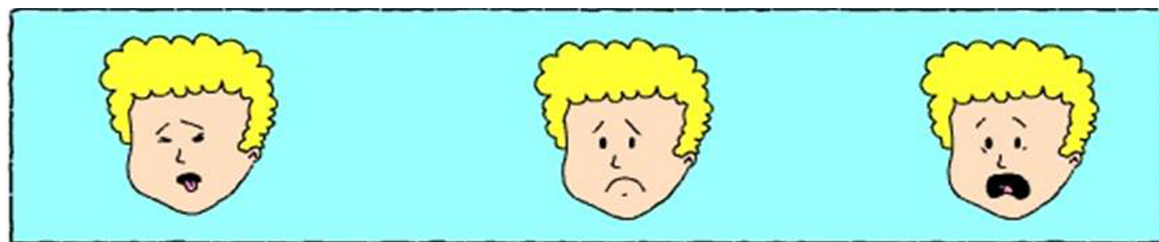
FIGURE 2 | Expression selection – an exemplary item assessing attribution of facial affective expression.

TABLE 1 | Descriptive statistics for the affective knowledge composite and measures.

		N	Min	Max	*M*	SD	SE
3.5 years old	Emotion understanding	280	0	16	5.73	3.31	0.19
	Expression selection	280	2	12	7.73	2.07	0.12
	Affective matching	280	0	4	2.02	1.35	0.08
	Affective knowledge	**280**	**−1.46**	**1.9**	**0.01**	**0.69**	**0.04**
5 years old	Emotion understanding	283	0	16	8.40	3.04	0.18
	Expression selection	283	4	12	9.32	1.97	0.12
	Affective matching	283	0	4	2.69	1.27	0.07
	Affective knowledge	**283**	**−1.81**	**1.45**	**0.00**	**0.70**	**0.03**

The three measures comprising the affective knowledge score and the affective knowledge composite computed by standardizing the raw scores of the measures and averaging them.

0.70 to 0.75. In the 5 year-olds sample the factor accounted for 52.43% of variance, with loadings ranging from 0.69 to 0.79. Thus, the factor structure remained relatively constant at both ages. Based on these results a total affective knowledge measure was computed by calculating Z-standardized scores for each measure, and averaging them. The descriptive statistics for the affective knowledge composite are shown in **Table 1**.

DRD4-III Polymorphism

DNA was extracted from 20 ml of mouthwash samples using Master Pure kit (Epicentre, Madison, WI, USA). PCR (Polymerase chain reaction) amplification was carried. The exon III repeat region of the *DRD4* receptor was characterized using PCR amplification procedure (using a Reddy Mix kit, AB gene, Surrey, UK), and genotyping was conducted as previously described by Knafo et al. (2011a).

In the 3.5 years old sample 94 participants (33.57%) were carriers of the 7R allele, and 186 were non-carriers. In the 5 years

TABLE 2 | Pearson's correlations.

		Expression selection	Affective matching
Emotion understanding	3.5 years old	0.26**	0.27**
	5 years old	0.29**	0.20**
Expression selection	3.5 years old		0.31**
	5 years old		0.33**

**P < 0.01.

TABLE 3 | Frequencies of Genotype (7R+, 7R−) by sex (Boys, Girls).

		7R+	7R−	Total
3.5 Years old	Boys	43	106	149
	Girls	51	80	131
	Total	**94 (33.57%)**	**186 (66.43%)**	**280 (100%)**
5 years old	Boys	38	111	149
	Girls	41	93	134
	Total	**79 (27.92%)**	**204 (72.08%)**	**283 (100%)**

7R+: carries of the 7 repeat (7R) allele (Participants with the presence of at least one 7R allele).
7R−: non-carries of the 7R allele (Participants with the 7R allele absent).

old sample 79 participants (27.92%) were carriers of the 7R allele, and 204 were non-carriers. In both ages genotypes were in Hardy–Weinberg equilibrium as tested with the PEDSTATS software (Wigginton and Abecasis, 2005) for a sample in which one participant from each family was chosen randomly (Age 3.5: $\chi^2 = 1.02$, $p = 0.79$; Age 5: $\chi^2 = 6.46$, $p = 0.09$). Results reflect a stable frequency of the *DRD4* repeat alleles in the population under study (see **Table 3** for descriptive data regarding the distribution of the genotype by sex, for both ages).

Statistical Analysis

The *DRD4-III* genotype was coded as a two level variable: "carriers of the 7R allele" (presence of at least one 7R allele), and "non-carriers of the 7R allele" (absence of a 7R allele), following Uzefovsky et al. (2014). Grouping of participants according to

7R-carrier status is common practice (e.g., Faraone et al., 2001; Bakermans-Kranenburg and van IJzendoorn, 2007; Knafo et al., 2011a; Uzefovsky et al., 2014) as the 7R allele is the second most common allele in Caucasian populations, and because of the difference in functionality between the 7R, specifically, and other alleles (Asghari et al., 1995). It would be interesting to test for an additive effect of sharing two 7R alleles, however, the homozygous (7-7R) genotype is relatively rare (in the current study 9 and 14 participants at ages 3.5 and 5, respectively).

The genotype, sex and their interaction served together as predictors of the affective knowledge score. Descriptive and preliminary statistics were carried out using SPSS v20 (Statistical Package for the Social Sciences). The main analyses were carried out using the GEE test (Generalized Estimating Equations test) in the SPSS. This test takes into account the dependency between twins, and enables using data from both twins. The results were verified by performing the analyses in the Mplus v5 software (Muthén and Muthén, 1998–2010). The procedure uses a slightly different analysis and a different set of assumptions to the GEE. Twins were considered as clustered within twin pairs, SE were computed using the TYPE = COMPLEX option, taking into account the fact that twin-data are non-independent of each other.

Results

Preliminary Analyses
Preliminary analyses showed that the sample of children, who participated at age 3.5 but not at age 5, did not differ on *DRD4* 7R genotype distribution, sex composition, or affective knowledge scores from those who participated at both time points. Similarly, children from families who joined the study at age 5 did not significantly differ from families who joined at age 3.5 on any study variable.

DRD4-III 7R Polymorphism and Sex
The main hypothesis in our investigation was that the *DRD4-III* polymorphism would be associated with children's affective knowledge, in a sex-contingent manner. We examined the association between the phenotype and the gene using the GEE procedure in SPSS v20. This procedure treats individual children as clustered by family using robust estimates of the SE, with affective knowledge regressed onto sex, *DRD4* genotype, and their interaction.

In the 3.5 years old sample the analysis yielded a significant effect of sex [Wald $\chi^2(1, N = 280) = 4.53, p = 0.03$], with boys scoring higher ($M = 0.10$, SE = 0.07) than girls ($M = -0.09$, SE = 0.06). Although there was no main effect for the *DRD4-III* polymorphism [Wald $\chi^2(1, N = 280) = 0.605, p = 0.44$], *DRD4-III* did qualify the sex difference, showing a significant interaction with sex [Wald $\chi^2(1, N = 280) = 4.97, p = 0.02$]. Results remained robust when the analysis was performed in the Mplus statistical package ($\beta = -0.15$, SE = 0.06, $p = 0.013$).

In order to better understand the nature of the interaction we examined simple effects. Mean comparison indicated that among carriers of the 7R allele, boys had higher affective knowledge scores ($M = 0.17$, SE = 0.13) than girls ($M = -0.23$, SE = 0.07), whereas no such effect was found for non-carriers ($M = 0.04$, SE = 0.07 vs. $M = 0.05$, SE = 0.09, respectively; see **Figure 3**). The sex effect on affective knowledge was significant for carriers of the 7R allele [Wald $\chi^2(1, N = 280) = 7.66, p = 0.006$], but not among non-carriers [Wald $\chi^2(1, N = 280) = 0.01, p = 0.92$]. Examination of the genotype effect separately for boys and girls yielded a significant effect among girls, as the 7R allele was significantly associated with lower affective knowledge scores and the absence of the 7R was significantly associated with higher affective knowledge scores [Wald $\chi^2(1, N = 280) = 5.85, p = 0.016$]. No significant genotype effect was found for boys [Wald $\chi^2(1, N = 280) = 0.88, p = 0.35$].

In an attempt to check whether these results are consistent in two different ages through childhood, we examined the association between the phenotype and the gene in the 5 year-olds sample. The main effects of genotype [Wald $\chi^2(1, N = 283) = 1.08, p = 0.30$] and sex [Wald $\chi^2(1, N = 283) = 0.28, p = 0.59$] were not significant, but the interaction between sex and *DRD4-III* polymorphism was again significant [Wald $\chi^2(1, N = 283) = 4.238, p = 0.04$], suggesting that the effect of the gene on affective knowledge is moderated by sex in both age groups. The interaction was significant in the Mplus analysis as well ($\beta = -0.12$, SE = 0.06, $p = 0.039$).

Furthermore, the direction of effects was similar in both ages. We examined simple effects in order to understand the nature of the interaction. Mean comparison indicated that among carriers of the 7R allele, boys had higher affective knowledge scores ($M = 0.13$, SE = 0.11) than girls ($M = -0.18$, SE = 0.09), whereas no such effect was found for non-carriers ($M = -0.01$, SE = 0.08 vs. $M = 0.04$, SE = 0.08, respectively; see **Figure 4**). The sex effect on affective knowledge was significant for carriers of the 7R allele [Wald $\chi^2(1, N = 283) = 4.816, p = 0.03$], showing that in the presence of the 7R allele boys scored significantly higher ($M = 0.13$, SE = 0.11) than girls ($M = -0.18$, SE = 0.09). For non-carriers of the 7R there was no significant sex effect on affective knowledge [Wald $\chi^2(1, N = 283) = 0.28, p = 0.56$]. Examination of the genotype effect separately for boys and girls yielded no significant effect for boys [Wald $\chi^2(1, N = 283) = 1.08, p = 0.30$] or for girls [Wald $\chi^2(1, N = 283) = 3.23, p = 0.07$].

In a subsample of children who only participated in the age 5 phase there was no significant effect for either sex, *DRD4-III* or their interaction, possibly reflecting the small sample size for this group ($N = 121$). Nevertheless, it is important to note that the pattern of findings, although not significant, was very similar for children from this new sample and for those retained from the age 3 phase, as seen in **Figure 5**.

Because there were age differences within each age group, and to account for the possible role of children's social and developmental background, we further tested several potential contributing variables. Gestational age (in weeks) or birthweight did not relate significantly to affective knowledge at either age group. Within-age group age differences (in months) were associated with affective knowledge at age 3.5 ($r = 0.16$,

7R+ : carriers of the 7R allele (participants with the presence of at least one 7 repeat allele);
7R- : non-carriers of the 7R allele (participants with the 7 repeat allele absent).
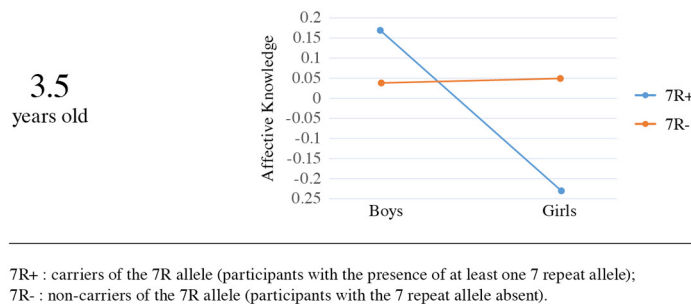
FIGURE 3 | The effect of *DRD4-III* polymorphism and sex on affective knowledge (3.5 years old).

7R+ : carriers of the 7R allele (participants with the presence of at least one 7 repeat allele);
7R- : non-carriers of the 7R allele (participants with the 7 repeat allele absent).
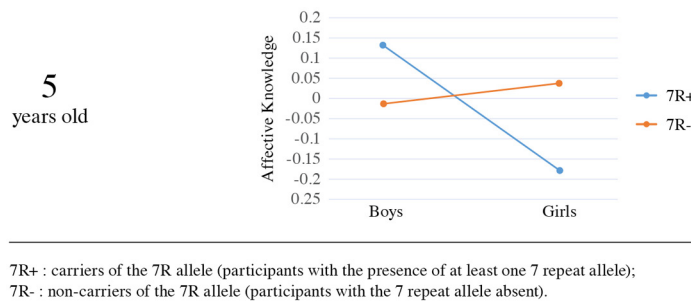
FIGURE 4 | The effect of *DRD4-III* polymorphism and sex on affective knowledge (5 years old).

7R+ : carriers of the 7R allele (participants with the presence of at least one 7 repeat allele);
7R- : non-carriers of the 7R allele (participants with the 7 repeat allele absent).
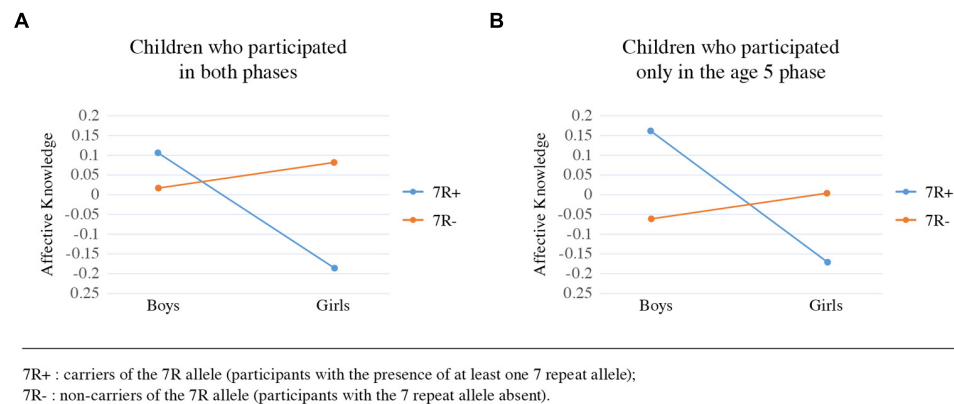
FIGURE 5 | Affective knowledge by *DRD4-III* polymorphism and sex of subsample of children. Mean comparison showing the same pattern of findings for children who participated in the 3.5 years old phase and retained for the 5 years old phase **(A)** and for children who participated only in the 5 years old phase **(B)**.

$p = 0.01$), and more weakly so at 5 ($r = 0.10$, *ns*). Family socio-economic status (SES, indicated by mothers' report on the family's income relative to the given national average, where $1 =$ much below national average, $3 =$ around average, and $5 =$ much above average) did not significantly correlate with affective knowledge at age 3.5 ($r = -0.12$, *ns*), but was significantly related to better performance at age 5 ($r = 0.25$, $p = 0.01$).

We therefore examined a covariate-adjusted model, controlling for within-age group age differences and SES. Importantly, the genotype × sex interaction remained significant when controlling for these variables, at both age 3.5, [Wald $\chi^2(1, N = 236) = 5.00$ $p = 0.025$], and age 5 [Wald $\chi^2(1, N = 210) = 3.92$, $p = 0.048$], attesting to the robustness of the findings.

Discussion

We examined the association between the *DRD4-III* polymorphism and affective knowledge among children 3.5 and 5 years of age. The findings demonstrate that the association

between *DRD4-III* and affective knowledge is contingent on sex. In both age groups, in the presence of the 7R allele boys scored significantly higher than girls, whereas in the absence of the 7R there was no significant sex effect on affective knowledge.

Due to the fact that socio-cognitive abilities, and especially cognitive empathy, develop dramatically during the preschool period (Lennon and Eisenberg, 1990), this consistent replication across two age groups is of special value. In other words, although children mature and change in these critical years, the interaction effect remains consistent, reflecting its robustness.

The results of the current study can be interpreted in the context of brain mechanisms underlying socio-cognitive abilities. The *DRD4* is widely expressed in the brain, particularly in the prefrontal cortex, hippocampus, hypothalamus, amygdala, and mesolimbic pathways (Matsumoto et al., 1995; Oak et al., 2000). These regions are considered to be a part of the "social brain" (Skuse and Gallagher, 2009), that includes the amygdala as one of the central parts of the reward system. *DRD4* is an integral part of the dopamine system, a system that is considered to be involved in making social interactions rewarding. The importance of the dopaminergic reward circuits in the regulation of social cognition is presented in the model of Skuse and Gallagher (2009, 2011). This model suggests that genetic variation in the receptors associated with OT, AVP, and dopamine may explain individual differences, as well as, deficits, in socio-cognitive processes and behaviors (Skuse and Gallagher, 2009, 2011). For example, the negative symptoms of schizophrenia (impairments in emotional processing, social perception and knowledge, ToM, and attributional bias) may be associated with abnormalities in OT and dopamine signaling in the amygdala (Rosenfeld et al., 2011). In this context it is important to note that the *DRD4* specifically was found to be expressed in excess in the striatum in postmortem brains of schizophrenia patients (Seeman et al., 1993). The current findings add to the literature by showing that dopamine, and especially *DRD4*, have an important role in typical social cognition, as well as in psychopathology.

Interestingly, the direction of the sex by gene interaction effect on affective knowledge was reversed compared to the effect found in adult populations on cognitive empathy (Uzefovsky et al., 2014). What may be the reasons for this reversal? One possibility is that genetic effects may be different across different developmental stages. Genetic effects can change from childhood to adulthood, both in the overall heritability of a trait (Haworth et al., 2008, 2010), and in the specific molecular genetic contributions to individual variability, as discussed above with regards to the genetics of generosity (Knafo et al., 2008b; Avinun et al., 2011). In addition, we must consider the fact that the sex effect can reflect both biological sex and/or gender, the social aspect of sex. Therefore, the gene by sex interaction can be understood in different ways (as elaborated by Uzefovsky et al., 2014).

Considering sex effects as related to biological as well as to social mechanisms, it is reasonable to assume that the effects of these mechanisms change from childhood to adulthood. Significant biological and social changes occur in the gap between these age periods, especially during adolescence. For example, the body undergoes significant externally visible changes due to surges in sex-hormones during puberty, and adolescents experience an increase in gender-differential socialization pressure (Hill and Lynch, 1983). Studies have shown that social behavior, such as pair bonding, romantic relationship, concern for others, and empathy, is influenced by sex-hormones (Hastings et al., 2006; van Anders and Gray, 2007). There are also some empirical data that support the notion that socialization contributes to gender differences in empathy (Lennon and Eisenberg, 1990). Taken together, it is possible that after a long period of gender-related changes (social and/or biological) the interactive contribution of gene and sex on cognitive empathy may change. These might explain the reversed direction of interaction between the different populations (children and adults).

Moreover, unless they are measured in the same sample, age differences often represent the fact that individuals of different ages were born and raised in different periods (i.e., cohort effects, e.g., Smits et al., 2011). The children in the present study were born two decades after the participants of the adult study (Uzefovsky et al., 2014), and changes in gender roles during this period may account in part for this difference (Twenge, 1997).

Finally, it is important to mention here that beyond the age difference between participants in the study of Uzefovsky et al. (2014) and the current study, there are also differences in the method used to assess the dependent variable. While adults' empathy was measured using a self-report questionnaire, in the current study we used a performance-based measure of affective knowledge. Although self-reported measures can reveal a complex and stable characteristic, they are more sensitive to reporter bias and demand characteristics. This might have an additional influence on the different findings.

Raising the possible reasons for the reversal in the interaction direction still leaves us with many unsolved questions: What is it about being a child-boy that, in the presence of the 7R allele, predisposes to higher affective knowledge scores? Why and when does this predisposition change during the maturational processes? Addressing these questions is very important and challenging due to the elusive nature of the notion "sex," as we note above. Carriers of the 7R allele are thought to be sensitive to environmental influence (e.g., Belsky and Pluess, 2009), and thus a combined effect of socialization effects across development with their tendency for being more strongly affected by such effects could contribute to the developmental change in the observed genotype × sex effects.

Conclusions and Suggestions for Future Research

In line with previous studies showing that the dopaminergic system is essential for social cognition, our findings suggest that *DRD4-III* polymorphism is associated with affective knowledge starting from early childhood, in interaction with sex. Being the first study to investigate the association between *DRD4-III* polymorphism and affective knowledge among children, this study provides novel evidence for the particular association of the genotype with cognitive empathy in early childhood.

A comparison with a recent study conducted on adult population reveals the possibility that the direction of the gender effect on the association between the genotype and phenotype changes throughout development.

Further studies are crucial in order to validate these findings and to expand our understanding of the molecular genetics of empathy and related variables. Future studies should use multiple measures, to better understand the role of measurement type on the results. In addition, it is important to examine the interaction effect of *DRD4-III* polymorphism and sex on the emotional component of empathy among children. Finally, the current findings emphasize the need to examine the role of genes in various age groups, from childhood, through adolescence to adulthood. This is especially important for understanding the reported interaction to determine whether and when the direction of the gender effect changes.

References

Adams, G. R., Schvaneveldt, J. D., and Jenson, G. O. (1979). Sex, age and perceived competency as correlates of empathic ability in adolescence. *Adolescence* 14, 811–818.

Asghari, V., Sanyal, S., Buchwaldt, S., Paterson, A., Jovanovic, V., and Van Tol, H. H. (1995). Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants. *J. Neurochem.* 65, 1157–1165. doi: 10.1046/j.1471-4159.1995.65031157.x

Avinun, R., Israel, S., Shalev, I., Gritsenko, I., Bornstein, G., Ebstein, R. P., et al. (2011). AVPR1A variant associated with preschoolers' lower altruistic behavior. *PLoS ONE* 6:e25274. doi: 10.1371/journal.pone.0025274

Avinun, R., and Knafo, A. (2013). The longitudinal Israeli study of twins (LIST) – an integrative view of social development. *Twin Res. Hum. Genet.* 16, 197–201. doi: 10.1017/thg.2012.73

Bachner-Melman, R., Gritsenko, I., Nemanov, L., Zohar, A. H., Dina, C., and Ebstein, R. P. (2005). Dopaminergic polymorphisms associated with self-report measures of human altruism: a fresh phenotype for the dopamine D4 receptor. *Mol. Psychiatry* 10, 333–335. doi: 10.1038/sj.mp.4001635

Bakermans-Kranenburg, M. J., and van IJzendoorn, M. H. (2007). Research review: genetic vulnerability or differential susceptibility in child development: the case of attachment. *J. Child Psychol. Psychiatry* 48, 1160–1173. doi: 10.1111/j.1469-7610.2007.01801.x

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or Asperger syndrome. *J. Child Psychol. Psychiatry* 38, 813–822. doi: 10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism. Dev. Disord.* 34, 163–175. doi: 10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1111/1469-7610.00715

Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: intervention outcomes. *J. Autism. Dev. Disord.* 32, 283–298. doi: 10.1023/A:1016378718278

Belsky, J., and Pluess, M. (2009). Beyond diathesis stress: differential susceptibility to environmental influences. *Psychological. Bull.* 135, 885–908. doi: 10.1037/a0017376

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Dev.* 53, 413–425. doi: 10.2307/1128984

Casey, R. J. (1993). Children's emotional experience: relations among expression, self-report, and understanding. *Dev. Psychol.* 29, 119–129. doi: 10.1037/0012-1649.29.1.119

Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., et al. (2009). Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and asperger syndrome. *Autism Res.* 2, 157–177. doi: 10.1002/aur.80

Chang, F. M., Kidd, J. R., Livak, K. J., Pakstis, A. J., and Kidd, K. K. (1996). The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. *Hum. Genet.* 98, 91–101. doi: 10.1007/s004390050166

Charman, T., Ruffman, T., and Clements, W. (2002). Is there a gender difference in false belief development? *Soc. Dev.* 11, 1–10. doi: 10.1111/1467-9507.00183

Choh, A. C., Lee, M., Kent, J. W., Diego, V. P., Johnson, W., Curran, J. E., et al. (2014). Gene-by-age effects on BMI from birth to adulthood: the fels longitudinal study. *Obesity* 22, 875–881. doi: 10.1002/oby.20517

Christ, C. C., Carlo, G., and Stoltenberg, S. F. (2015). Oxytocin receptor (OXTR) single nucleotide polymorphisms indirectly predict prosocial behavior through perspective taking and empathic concern. *J. Pers.* doi: 10.1111/jopy.12152 [Epub ahead of print].

Cutting, A. L., and Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Dev.* 70, 853–865. doi: 10.1111/1467-8624.00061

Davis, M. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catal. Select. Doc. Psychol.* 10, 85–104.

Decety, J., and Lamm, C. (2006). Human empathy through the lens of social neuroscience. *Sci. World J.* 6, 1146–1163. doi: 10.1100/tsw.2006.221

Denham, S. A. (1986). Social cognition, prosocial behavior, and emotion in preschoolers: contextual validation. *Child Dev.* 57, 194–201. doi: 10.2307/1130651

de Vignemont, F., and Singer, T. (2006). The empathic brain: how, when and why? *Trends Cogn. Sci.* 10, 435–441. doi: 10.1016/j.tics.2006.08.008

Eisenberg, N., and Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychol. Bull.* 94, 100–131. doi: 10.1037/0033-2909.94.1.100

Eisenberg, N., Shea, C. L., Carlo, G., and Knight, G. P. (2014). "Empathy-related responding and cognition: a chicken and the egg dilemma," in *Handbook of Moral Behavior and Development Research*, Vol. 2 eds W. M. Kurtines and J. L. Gewirtz (New York: Erlbaum).

Eisenberg, N., and Strayer, J. (1990). "Critical issues in the study of empathy," in *Empathy and Its Development*, eds N. Eisenberg and J. Strayer (Cambridge: Cambridge University Press), 7–13.

Faraone, S. V., Doyle, A. E., Mick, E., and Biederman, J. (2001). Meta-analysis of the association between the 7-repeat allele of the dopamine D4 receptor gene and attention deficit hyperactivity disorder. *Am. J. Psychiatry* 158, 1052–1057. doi: 10.1176/appi.ajp.158.7.1052

Garner, P. W., and Waajid, B. (2012). Emotion knowledge and self-regulation as predictors of preschoolers' cognitive ability, classroom behavior, and social competence. *J. Psychoeduc. Assess.* 30, 330–343. doi: 10.1177/0734282912449441

Gross, A. L., and Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: a review. *Dev. Rev.* 11, 368–398. doi: 10.1016/0273-2297(91)90019-K

Hastings, P. D., Zahn-Waxler, C., and McShane, K. (2006). "We are, by nature, moral creatures: biological bases of concern for others," in *Handbook of Moral Development*, ed. M. K. J. G. Smetana (Mahwah, NJ: Erlbaum Publisher), 483–516.

Haworth, C., Carnell, S., Meaburn, E. L., Davis, O. S., Plomin, R., and Wardle, J. (2008). Increasing heritability of BMI and stronger associations with the FTO gene over childhood. *Obesity* 16, 2663–2668. doi: 10.1038/oby.2008.434

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., De Geus, E. J. C., Van Beijsterveldt, C. E. M., et al. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Mol. Psychiatry* 15, 1112–1120. doi: 10.1038/mp.2009.55

Hill, J. P., and Lynch, M. E. (1983). "The intensification of gender-related role expectations during early adolescence," in *Girls at Puberty,* eds J. Brooks-Gunn et al. (Berlin: Springer), 201–228.

Hoffman, M. L. (1977). Sex differences in empathy and related behaviors. *Psychol. Bull.* 84, 712–722. doi: 10.1037/0033-2909.84.4.712

Hughes, C., and Cutting, A. L. (1999). Nature, nurture, and individual differences in early understanding of mind. *Psychol. Sci.* 10, 429–432. doi: 10.1111/1467-9280.00181

Hughes, C., Jaffee, S. R., Happé, F., Taylor, A., Caspi, A., and Moffitt, T. E. (2005). Origins of individual differences in theory of mind: from nature to nurture? *Child Dev.* 76, 356–370. doi: 10.1111/j.1467-8624.2005.00850_a.x

Israel, S., Hasenfratz, L., and Knafo-Noam, A. (2015). The genetics of morality and prosociality. *Curr. Opin. Psychol.* 6, 55–59. doi: 10.1016/j.copsyc.2015.03.027

Kang, J. I., Namkoong, K., and Kim, S. J. (2008). Association of DRD4 and COMT polymorphisms with anger and forgiveness traits in healthy volunteers. *Neurosci. Lett.* 430, 252–257. doi: 10.1016/j.neulet.2007.11.005

Knafo, A. (2006). The longitudinal Israeli study of twins (LIST): children's social development as influenced by genetics, abilities, and socialization. *Twin Res. Hum. Genet.* 9, 791–798. doi: 10.1375/twin.9.6.791

Knafo, A., Israel, S., and Ebstein, R. P. (2011a). Heritability of children's prosocial behavior and differential susceptibility to parenting by variation in the dopamine receptor D4 gene. *Dev. Psychopathol.* 23, 53–67. doi: 10.1017/S0954579410000647

Knafo, A., Steinberg, T., and Goldner, I. (2011b). Children's low affective perspective-taking ability is associated with low self-initiated pro-sociality. *Emotion* 11, 194–198. doi: 10.1037/a0021240

Knafo, A., and Uzefovsky, F. (2013). "Variation in empathy: the interplay of genetic and environmental factors," in: *The Infant Mind: Origins of the Social Brain*, eds M. Legerstee, D. W. Haley, and M. H. Bornstein (New York: The Guilford Press), 97–122.

Knafo, A., Zahn-Waxler, C., Davidov, M., Van Hulle, C., Robinson, J. L., and Rhee, S. H. (2009). Empathy in early childhood. *Ann. N. Y. Acad. Sci.* 1167, 103–114. doi: 10.1111/j.1749-6632.2009.04540.x

Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., and Rhee, S. H. (2008a). The developmental origins of a disposition toward empathy: genetic and environmental contributions. *Emotion* 8, 737–752. doi: 10.1037/a0014179

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L., et al. (2008b). Individual differences in allocation of funds in the Dictator Game and postmortem hippocampal mRNA levels are correlated with length of the arginine vasopressin 1a receptor (AVPR1a) RS3 promoter-region repeat. *Genes Brain Behav.* 7, 266–275. doi: 10.1111/j.1601-183X.2007.00341.x

Knafo-Noam, A., Uzefovsky, F., Israel, S., Davidov, M., and Zahn-Waxler, C. (2015). The prosocial personality and its facets: genetic and environmental architecture of mother-reported behavior of 7-year-old twins. *Front Psychol.* 6:112. doi: 10.3389/fpsyg.2015.00112

Lackner, C. L., Bowman, L. C., and Sabbagh, M. A. (2010). Dopaminergic functioning and preschoolers' theory of mind. *Neuropsychologia* 48, 1767–1774. doi: 10.1016/j.neuropsychologia.2010.02.027

Lackner, C., Sabbagh, M. A., Hallinan, E., Liu, X., and Holden, J. J. (2012). Dopamine receptor D4 gene variation predicts preschoolers' developing theory of mind. *Dev. Sci.* 15, 272–280. doi: 10.1111/j.1467-7687.2011.01124.x

Lau, J. Y., Burt, M., Leibenluft, E., Pine, D. S., Rijsdijk, F., Shiffrin, N., et al. (2009). Individual differences in children's facial expression recognition ability: the role of nature and nurture. *Dev. Neuropsychol.* 34, 37–51. doi: 10.1080/87565640802564424

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., and David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychol. Med.* 34, 911–920. doi: 10.1017/S0033291703001624

Lennon, R., and Eisenberg, N. (1990). "Gender and age differences in empathy and sympathy," in *Empathy and its Development*, eds N. Eisenberg and J. Strayer (Cambridge: Cambridge University Press), 195–213.

Lucht, M. J., Barnow, S., Sonnenfeld, C., Ulrich, I., Grabe, H. J., Schroeder, W., et al. (2013). Associations between the oxytocin receptor gene (OXTR) and "mind-reading" in humans – an exploratory study. *Nord. J. Psychiatry* 67, 15–21. doi: 10.3109/08039488.2012.700731

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., et al. (2000). A new test to measure emotion recognition ability: matsumoto and Ekman's Japanese and caucasian brief affect recognition test (JACBART). *J. Nonverbal Behav.* 24, 179–209. doi: 10.1023/A:100666812 0583

Matsumoto, M., Hidaka, K., Tada, S., Tasaki, Y., and Yamaguchi, T. (1995). Full-length cDNA cloning and distribution of human dopamine D4 receptor. *Mol. Brain Res.* 29, 157–162. doi: 10.1016/0169-328X(94)00245-A

McKone, E., and Palermo, R. (2010). A strong role for nature in face recognition. *Proc. Natl. Acad. Sci.* 107, 4795–4796. doi: 10.1073/pnas.1000567107

Muthén, L. K., and Muthén, B. O. (1998–2010). *Mplus User's Guide*, 6th edn. Los Angeles, CA: Authors.

Oak, J. N., Oldenhof, J., and Van Tol, H. H. (2000). The dopamine D 4 receptor: one decade of research. *Eur. J. Pharmacol.* 405, 303–327. doi: 10.1016/S0014-2999(00)00562-8

Preston, S. D. (2013). The origins of altruism in offspring care. *Psychol. Bull.* 139, 1305–1341. doi: 10.1037/a0031755

Ribordy, S. C., Camras, L. A., Stefani, R., and Spaccarelli, S. (1988). Vignettes for emotion recognition research and affective therapy with children. *J. Clin. Child Psychol.* 17, 322–325. doi: 10.1207/s15374424jccp1704_4

Roberts, W., and Strayer, J. (1996). Empathy, emotional expressiveness, and prosocial behavior. *Child Dev.* 67, 449–470. doi: 10.2307/1131826

Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., and Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21437–21441. doi: 10.1073/pnas.0909579106

Ronald, A., Happé, F., Hughes, C., and Plomin, R. (2005). Nice and nasty theory of mind in preschool children: nature and nurture. *Soc. Dev.* 14, 664–684. doi: 10.1111/j.1467-9507.2005.00323.x

Ronald, A., Viding, E., Happé, F., and Plomin, R. (2006). Individual differences in theory of mind ability in middle childhood and links with verbal ability and autistic traits: a twin study. *Soc. Neurosci.* 1, 412–425. doi: 10.1080/17470910601068088

Rosenfeld, A. J., Lieberman, J. A., and Jarskog, L. F. (2011). Oxytocin, dopamine, and the amygdala: a neurofunctional model of social cognitive deficits in schizophrenia. *Schizophr. Bull.* 37, 1077–1087. doi: 10.1093/schbul/sbq015

Schneiderman, I., Kanat-Maymon, Y., Ebstein, R. P., and Feldman, R. (2013). Cumulative risk on the oxytocin receptor gene (OXTR) underpins empathic communication difficulties at the first stages of romantic love. *Soc. Cogn. Affect. Neurosci.* 1524–1529. doi: 10.1093/scan/nst142

Seeman, P., Guan, H. C., and Van Tol, H. H. (1993). Dopamine D4 receptors elevated in schizophrenia. *Nature* 365, 441–445. doi: 10.1038/365441a0

Shamay-Tsoory, S. G., Aharon-Peretz, J., and Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain* 132, 617–627. doi: 10.1093/brain/awn279

Simpson, J. A., Oriña, M. M., and Ickes, W. (2003). When accuracy hurts, and when it helps: a test of the empathic accuracy model in marital interactions. *J. Pers. Soc. Psychol.* 85, 881–893. doi: 10.1037/0022-3514.85.5.881

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863. doi: 10.1016/j.neubiorev.2006.06.011

Skuse, D. H., and Gallagher, L. (2009). Dopaminergic-neuropeptide interactions in the social brain. *Trends Cogn. Sci.* 13, 27–35. doi: 10.1016/j.tics.2008.09.007

Skuse, D. H., and Gallagher, L. (2011). Genetic influences on social cognition. *Pediatr. Res.* 69, 85–91. doi: 10.1203/PDR.0b013e318212f562

Smits, I. A., Dolan, C. V., Vorst, H., Wicherts, J. M., and Timmerman, M. E. (2011). Cohort differences in Big Five personality factors over a period of 25 years. *J. Pers. Soc. Psychol.* 100, 1124–1138. doi: 10.1037/a0022874

Susskind, J. M., Littlewort, G., Bartlett, M. S., Movellan, J., and Anderson, A. K. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia* 45, 152–162. doi: 10.1016/j.neuropsychologia.2006.05.001

Twenge, J. M. (1997). Changes in masculine and feminine traits over time: a meta-analysis. *Sex Roles* 36, 305–325. doi: 10.1007/BF02766650

Uekermann, J., Kraemer, M., Abdel-Hamid, M., Schimmelmann, B. G., Hebebrand, J., Daum, I., et al. (2010). Social cognition in attention-deficit hyperactivity disorder (ADHD). *Neurosci. Biobehav. Rev.* 34, 734–743. doi: 10.1016/j.neubiorev.2009.10.009

Uzefovsky, F., Shalev, I., Israel, S., Edelman, S., Raz, Y., Mankuta, D., et al. (2015). Oxytocin receptor and vasopressin receptor 1a genes are, respectively, associated with emotional and cognitive empathy. *Horm. Behav.* 67, 60–65. doi: 10.1016/j.yhbeh.2014.11.007

Uzefovsky, F., Shalev, I., Israel, S., Edelman, S., Raz, Y., Perach-Barzilay, N., et al. (2014). The Dopamine D4 receptor gene shows a gender-sensitive association with cognitive empathy: evidence from two independent samples. *Emotion* 14, 712–721. doi: 10.1037/a0036555

van Anders, S. M., and Gray, P. B. (2007). Hormones and human partnering. *Annu. Rev. Sex Res.* 18, 60–93.

Volbrecht, M. M., Lemery-Chalfant, K., Aksan, N., Zahn-Waxler, C., and Goldsmith, H. H. (2007). Examining the familial link between positive affect and empathy development in the second year. *J. Genet. Psychol.* 168, 105–130. doi: 10.3200/GNTP.168.2.105-130

Walker, S. (2005). Gender differences in the relationship between young children's peer-related social competence and individual differences in theory of mind. *J. Genet. Psychol.* 166, 297–312. doi: 10.3200/GNTP.166.3.297-312

Walter, H. (2012). Social cognitive neuroscience of empathy: concepts, circuits, and genes. *Emotion Rev.* 4, 9–17. doi: 10.1177/1754073911421379

Wigginton, J., and Abecasis, G. (2005). PEDSTATS: descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics* 21, 3445–3447. doi: 10.1093/bioinformatics/bti529

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., et al. (2010). Human face recognition ability is specific and highly heritable. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5238–5241. doi: 10.1073/pnas.0913053107

Wu, N., Li, Z., and Su, Y. (2012). The association between oxytocin receptor gene polymorphism (OXTR) and trait empathy. *J. Affect. Disord.* 138, 468–472. doi: 10.1016/j.jad.2012.01.009

Wu, N., and Su, Y. (2014). Oxytocin receptor gene relates to theory of mind and prosocial behavior in children. *J. Cogn. Dev.* 16, 302–313. doi: 10.1080/15248372.2013.858042

Yirmiya, N., Sigman, M. D., Kasari, C., and Mundy, P. (1992). Empathy and cognition in high-functioning children with Autism. *Child Dev.* 63, 150–160. doi: 10.2307/1130909

Young, L. J., and Wang, Z. (2004). The neurobiology of pair bonding. *Nat. Neurosci.* 7, 1048–1054. doi: 10.1038/nn1327

Zahn-Waxler, C., Cole, P. M., and Barrett, K. C. (1991). "Guilt and empathy: sex differences and implications for the development of depression," in *The Development of Emotion Regulation and Dysregulation,* eds J. Garber and K. A. Dodge (Cambridge: Cambridge University Press), 243–272. doi: 10.1017/CBO9780511663963.012

Zahn-Waxler, C., Cummings, E. M., and Cooperman, C. (1984). "Emotional development in childhood," in *Annals of Child Development,* Vol. 1, ed. C. Whitehurst (New York: JAI), 45–106

Zahn-Waxler, C., Shirtcliff, E. A., and Marceau, K. (2008). Disorders of childhood and adolescence: gender and psychopathology. *Annu. Rev. Clin. Psychol.* 4, 275–303. doi: 10.1146/annurev.clinpsy.3.022806.091358

Zhong, S., Israel, S., Shalev, I., Xue, H., Ebstein, R. P., and Chew, S. H. (2010). Dopamine D4 receptor gene associated with fairness preference in ultimatum game. *PLoS ONE* 5:e13765. doi: 10.1371/journal.pone.0013765

Zhu, B., Chen, C., Moyzis, R. K., Dong, Q., Chen, C., He, Q., et al. (2011). Genetic variations in the dopamine system and facial expression recognition in healthy chinese college students. *Neuropsychobiology* 65, 83–89. doi: 10.1159/000329555

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Curr. Biol.* 20, 137–142. doi: 10.1016/j.cub.2009.11.067

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Do semantic contextual cues facilitate transfer learning from video in toddlers?

Laura Zimmermann[1], Alecia Moser[2], Amanda Grenell[3], Kelly Dickerson[4], Qianwen Yao [1], Peter Gerhardstein[2]* and Rachel Barr[1]

[1] Department of Psychology, Georgetown University, Washington, DC, USA, [2] Department of Psychology, Binghamton University, Binghamton, NY, USA, [3] Institute of Child Development, University of Minnesota, Minneapolis, MN, USA, [4] Army Research Laboratory, Human Research and Engineering Directorate, Aberdeen Proving Ground, Aberdeen, MD, USA

Young children typically demonstrate a *transfer deficit*, learning less from video than live presentations. Semantically meaningful context has been demonstrated to enhance learning in young children. We examined the effect of a semantically meaningful context on toddlers' imitation performance. Two- and 2.5-year-olds participated in a puzzle imitation task to examine learning from either a live or televised model. The model demonstrated how to assemble a three-piece puzzle to make a fish or a boat, with the puzzle demonstration occurring against a semantically meaningful background context (ocean) or a yellow background (no context). Participants in the video condition performed significantly worse than participants in the live condition, demonstrating the typical *transfer deficit effect*. While the context helped improve overall levels of imitation, especially for the boat puzzle, only individual differences in the ability to self-generate a stimulus label were associated with a reduction in the transfer deficit.

Keywords: transfer deficit, context learning, imitation, social learning, learning from screen media, memory binding

Introduction

Infants and young children perform more poorly on tasks involving transfer of learning from television to real-life situations than in direct face-to-face interactions. This finding, which has been termed the *transfer deficit* (Barr, 2010, 2013), is supported by data from multiple investigations including imitation (Barr and Hayne, 1999; Flynn and Whiten, 2008; Nielsen et al., 2008; Zack et al., 2009; Simcock et al., 2011; Dickerson et al., 2013), object retrieval (Troseth and DeLoache, 1998; Troseth et al., 2006), self-recognition (Suddendorf et al., 2007), and object recognition tasks (Carver et al., 2006; Simcock and Dooley, 2007). For example, Dickerson et al. (2013) reported a transfer deficit that persisted across early childhood. Using an imitation procedure, they modeled assembly of a 3-piece puzzle of a fish or a boat via either a live or televised demonstration. Toddlers (2- and 2.5-year-olds) imitated significantly fewer gestures and goals following a video demonstration than a live demonstration. This transfer deficit is problematic for early childhood learning, especially with the increased popularity of computers, television, and other interactive media as teaching tools for infants and toddlers (Rideout, 2013, 2014).

One account of the deficit notes that it may be challenging for children to perceptually match features between encoding and retrieval when the features undergo changes in color, brightness, motion, and depth information between the demonstration (e.g., video) and the test. These changes increase the *transfer distance* (Barnett and Ceci, 2002) between the training

and test situations; that is, the degree of similarity between encoding and retrieval of new information. Other accounts focus on the effect that transfer distance has on specific aspects of memory processing including symbolic understanding (Troseth and DeLoache, 1998), memory flexibility (Hayne, 2004) and memory binding (Olson and Newcombe, 2014). The transfer deficit can be ameliorated by manipulations that reduce transfer distance and increase memory flexibility (see Barr, 2010, 2013; Troseth, 2010 for review and discussion) through repetition (Barr et al., 2007), social engagement (Tennie et al., 2006; Nielsen et al., 2008; Subiaul et al., 2012), contingency cues (eye contact, directed gaze, directed pointing; Csibra and Gergely, 2006), and increased perceptual realism (Simcock and DeLoache, 2006; Simcock et al., 2011). Given the well-established beneficial role of context in learning, the present study sought to address whether the inclusion of a semantically meaningful context would ameliorate the transfer deficit on an established imitation task in 2- and 2.5-year-olds.

Effect of Context

Context is the physical, temporal, and affective or internal environment within which an event occurs (Bouton, 1993). The role of context in learning is well established in both the animal and human learning literature (Bouton, 1993; Boller et al., 1996; see also Rescorla and Wagner, 1972; Barnat et al., 1996; Learmonth et al., 2004). Bouton (1993) found that consistency between the context present at training (encoding) and the context present at test facilitates memory retrieval. The benefit of contextual consistency has also been found in studies of infants using the mobile conjugate reinforcement procedure (e.g., Borovsky and Rovee-Collier, 1990) and a deferred imitation paradigm (Hayne et al., 2000). While the context and cues are often discussed experimentally as separate and relatively independent entities, early in development context and cues are thought to be parts of a single encoded event (Spear and McKinzie, 1994). When these parts are congruent, object recognition should be more precise (Oliva and Torralba, 2007). Thus, children may be better able to identify a fish in the ocean than on a mountaintop, or against a solid (non-specific) background.

An encoded event is generally seen as the result of memory binding. Memory binding is the process of encoding the relations among stimuli that co-occur spatially or temporally (Cohen and Eichenbaum, 1993). This process is critical to the ability to integrate visual background context into a memory for central foreground object details. There is a long developmental trajectory of memory binding across childhood that has been linked to hippocampal development (Raj and Bell, 2010; Olson and Newcombe, 2014), but investigation across early childhood has not been systematic due to differences in approaches and measurement across age.

Early in development, cue information appears to be inextricably bound to other memory attributes, including attributes of the context in which the event occurs, such as the background scene, making transfer of learning outside a particular context challenging (Spear and McKinzie, 1994). Boller et al. (1996) reported that infants' memory retrieval was robust in the presence of the training context, but degraded when the context was changed

or removed. Additional work has demonstrated that memory retrieval of 6-month-olds is highly context-specific, such that a contextual change (i.e., original mobile cue in a novel context) disrupts retention following as little as a 24-h delay (Borovsky and Rovee-Collier, 1990; Hartshorn and Rovee-Collier, 1997). Development then is characterized by a decrease in contextualized learning and a subsequent increase in flexibility of memories to withstand changes in context.

Hayne et al. (2000) found similar effects with 6-month-olds using a deferred imitation paradigm. They found that when the context changed between demonstration and test (e.g., from home to laboratory), 6-month-old infants were no longer able to imitate the target actions, while older children, 12- and 18-month-olds, were successful at transferring learning across a context change from the home to the laboratory setting. These studies suggest that learning is highly context-specific early in infancy, and that context features might bind with other memory attributes to form a single cue representation. Hayne (2004) noted that this high degree of memory specificity constrains memory flexibility and generalization of learning to new settings, and argued that the ability to use memory more flexibly develops across infancy and childhood.

Memory binding has also been examined using visual recognition memory paradigms during infancy and has revealed evidence of fragile memory binding (Richmond and Nelson, 2009). Using precise eye-tracking techniques, Richmond and Nelson (2009) demonstrated that infants could encode memories based on relationships between images. However, with age-dependent experience, children learn to disregard or de-emphasize less relevant contextual information and focus more on central cues (Bornstein et al., 2011). An increase in hippocampal volume in infancy may help explain the rapid changes in memory binding (Olson and Newcombe, 2014). By age 2, there is a shift in spatial coding and representation as children are able to encode multiple spatial locations and maintain them across a delay (Sluzenski et al., 2004).

Less work investigating the memory binding capacities of toddlers is available. Studies with 4- to 6-year-olds have used protocols adapted from studies of adults. In particular, children display difficulty with tasks that involve reporting the combination of visual foreground object and contextual background cues, suggesting that memory binding continues to develop into the preschool years (Sluzenski et al., 2006; Lloyd et al., 2009). Specifically, there were age-related increases in performance between 4 and 6 years (Lloyd et al., 2009). There were also age-related differences between children and adults; 4- to 6-year-old children performed significantly worse on the combined condition than adults (Sluzenski et al., 2006). One possible explanation for the poor performance of 4- to 6-year-olds in these studies was memory binding or retrieval deficits, but another possibility is that the task (verbal report) was too taxing for this age range. A non-verbal measure is likely to provide a better index of memory binding in younger children.

More recently, Newcombe et al. (2014) used an episodic memory search paradigm to examine memory binding in young children, and found systematic age-related increases in search performance among 15- to 72-month-olds. Older children

remembered more items (toys) across different rooms (contexts). Younger children (under 26 months) remembered more locations when they were given a label than when they were not, but older children (34- to 56-month-olds) did not benefit from a label cue. Other studies have demonstrated that the scale of the contextual cue also critically determines whether toddlers will use the cue effectively or not. For example, Deloache et al. (2004) demonstrated that transfer from a small-scale model to a larger-scale model was significantly easier than transfer from a small-scale model to a real room and vice versa for 2.5- and 3-year-olds. In the present study, we therefore reduced the demands on toddlers by testing them with fewer items and by testing the effect of contextual cues on transfer within a much smaller space.

Taken together, the studies discussed above demonstrate age-related changes in memory binding and processing of contextual cues from infancy to school-age. These changes are associated with a host of developmental changes in memory processing. Older children have better memory capabilities; that is, they encode more efficiently and are better able to equate and integrate information across different contexts compared to younger children, showing better memory flexibility across time (Hayne, 2004; Barr, 2013). Contextual cues may be weighted and bound differently as a function of age and complexity (Deloache et al., 2004; Olson and Newcombe, 2014). Infants may encode background contextual information at the expense of central information, resulting in disruption in memory processing when the context changes (e.g., Borovsky and Rovee-Collier, 1990; Shields and Rovee-Collier, 1992; Boller et al., 1996; Hayne et al., 2000). Toddlers may progress from fused memory representations that have both central and background information, to memory representations that contain primarily central cue information, resulting in neither a disruptive nor a facilitative effect of context. Later in development, they may progress to more flexible adult-like memory that has both background and central information stored in a relational network that can be accessed depending upon the specific situation (Sluzenski et al., 2006; Lloyd et al., 2009; Olson and Newcombe, 2014).

Effect of Visually Meaningful Cues on Memory

Manipulations of context in the research discussed above were highly distinctive (large changes to brightly colored crib liners, different physical locations), but not *iconic*. An iconic – or semantically related – visual context is thought to tap into the rich background knowledge and the extensive visual experience of the observer, and thus facilitate performance (Simcock and DeLoache, 2006; Pereira and Smith, 2009). An example would be depicting a car on a street, as compared to an arbitrary context (Biederman, 1972; Oliva and Torralba, 2007). A related visual context can direct spatial attention to important features in a display and facilitate adult memory for a visual context (Chun and Jiang, 1998). This is especially relevant during early childhood as visual context may be a semantically meaningful cue for young children, who have a smaller verbal semantic network available to them. A semantically meaningful context has been shown to facilitate recognition and object search in

2-year-olds (Pereira and Smith, 2009). Finally, 24-month-olds perform significantly more target actions from a picture book when drawings are iconic photographs than when they are line drawings (Simcock and DeLoache, 2006). The potential advantage conveyed by related context may be especially relevant during early childhood, as contextual cues may increase the probability of retrieving a semantically meaningful target. Young children have a smaller verbal semantic network available to them, and thus the presence of an iconic visual context may produce a greater level of performance increase by providing more cues with which to access the memory. The role that visual context – specifically semantically meaningful scenes – plays on learning and memory in toddlers is explored here.

The Present Study

The present study adopted methodology from Dickerson et al. (2013), using the same 3-piece boat and fish puzzle apparatus. The primary research question was whether the presence of a semantically meaningful visual context would ameliorate the transfer deficit on the puzzle imitation task. The reproduction of demonstrated gestures and final goal state of the puzzle (fish or boat) during the task, were coded in the present study. Groups of 2- and 2.5-year-old children were tested on the puzzle imitation task following a live or video demonstration. These ages were selected because the Dickerson et al. (2013) test demonstrated that performance in this age range is neither at floor nor at ceiling for the puzzle task. Half of the children were assigned to a meaningful semantic context condition and the other half were not. Performance was compared to baseline controls that never saw a demonstration. The current study extends previous work by manipulating the presence of a semantically meaningful context to examine whether increasing semantic congruence can ameliorate the transfer deficit.

We sought to link context to confines of a smaller space than previous memory binding studies in large rooms (Newcombe et al., 2014) using a task that 2- to 3-year-olds has been successful on. Additionally, we intended to extend previous work on the transfer deficit to include the role of context. Consistent with the memory binding accounts of context, we hypothesized that the presence of a visual semantic context (i.e., ocean) would facilitate imitation of a demonstrated goal and gestures. Given previously documented age-related changes in imitation on this task (Dickerson et al., 2013), we hypothesized that older children would be more successful in transfer tasks compared to younger children. Furthermore, applying the transfer deficit concept (Barr, 2010, 2013) to this design, we predicted that the addition of a semantically meaningful visual context would ameliorate the transfer deficit.

Materials and Methods

Participants

The study included 165 typically developing children (87 boys) from two metropolitan areas. Independent groups of children were tested at 2 years ($N = 88$, M age = 24 months 16 days, $SD = 11.46$ days, range 23–25 months) and 2.5 years ($N = 77$,

M age = 30 months 16 days, SD = 25 days, range 28–31 months). Participants were primarily Caucasian (79.9%) and from college-educated families (*M* years of education = 17.26, SD = 1.26). The remaining 20% of the sample included the following races: Mixed (14.6%), African–American (1.8%), Asian (0.9%), and not reported (2.8%). Additionally, 6.5% of the sample was Latino. The mean rank of socioeconomic index (SEI; Nakao and Treas, 1994) was 74.14 (SD = 19.12) based on 127 families (76%). Additional children were excluded from the analysis for the following reasons: eight due to experimenter error, three for technical error, five for failure to interact with the experimental stimuli, 14 due to parental interference, and 20 for interacting with the stimuli prior to test.

Apparatus

This study used a metal board inserted into a rectangular black case. The case was 35 cm tall, 42 cm wide, and 23 cm deep. The metal board could be easily slid in and out of the black case. The black case behind the metal board contained an LCD monitor that was only visible when the metal board was not in place (see **Figure 1**). The metal board was either completely school bus yellow or displayed a cartoon of the ocean. The caricature of the ocean had a light blue sky, with dark blue waves representing the ocean, and a yellow sun located at the center left of the sky. The sun was composed of one semi-circle and three triangles (see **Figure 2**).

Stimuli

The stimuli consisted of three magnet pieces that were various shapes and colors but were the same thickness (0.5 cm). These magnets were strong enough that they stuck to the metal board, but they were weak enough so that they could be easily moved around. The pieces, when moved and connected correctly, formed either a "boat" or a "fish." At the beginning of the trial,
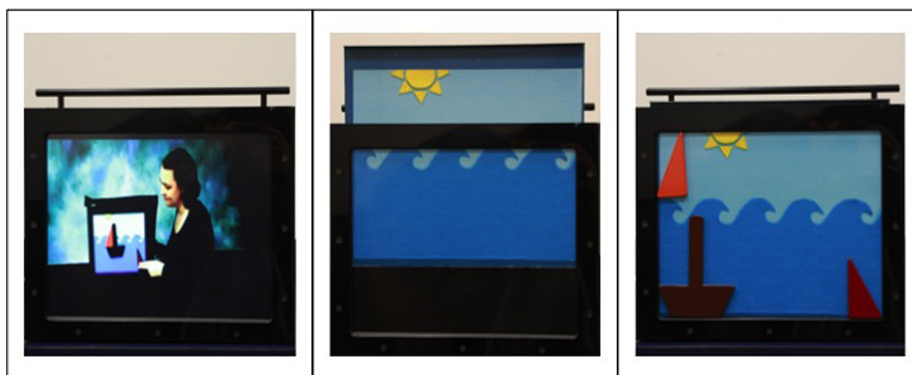
FIGURE 1 | The image on the far left depicts the apparatus with the video image displayed. The center image depicts the apparatus with both the magnet board and video screen. The image on the far right depicts the magnet board with the boat puzzle pieces affixed to magnet board.
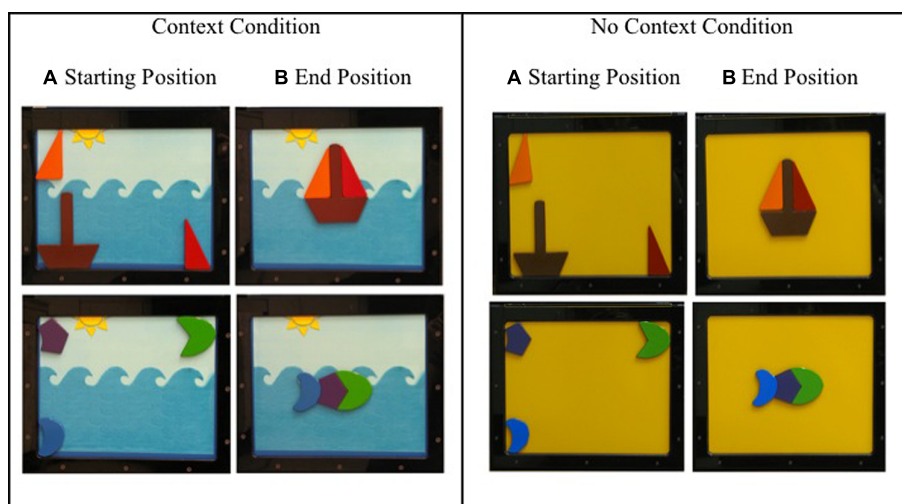
FIGURE 2 | (Left) Context condition. The cluster of four images on the left shows the stimuli with the ocean background. **(Right)** No context condition. The cluster of four images on the right shows the stimuli with the schoolbus-yellow background. Within each context, **(A)** shows the starting position of the stimuli for the boat at the top and fish at the bottom and **(B)** shows the end position for each puzzle.

each piece was placed in a different corner of the metal board. For each puzzle, there were two predetermined placements for the pieces.

Boat

The boat puzzle consisted of three pieces: one red right triangle piece and one orange right triangle piece that represented the sails of the sailboat, and one trapezoid piece with a long thin rectangle attached at the center that represented the hull and the mask of the sail boat (see **Figure 2**).

Fish

The fish puzzle consisted of three pieces: one green moon shaped piece that represented the head of the fish, one purple pentagon piece that represented the body of the fish, and one blue moon shaped piece that represented the tail of the fish (see **Figure 2**).

Vocabulary and Demographics Information

The caregiver was asked to complete a general information questionnaire (assessing SES, parental education, childcare, and language) as well as the MacArthur Communicative Development Inventory: Words and Sentences Short Form (MCDI) to measure children's productive vocabulary (Fenson et al., 2000).

Design

Children were randomly assigned to independent groups in order to conduct a 3 (Condition: Live, Video, Baseline) × 2 (Context: no context or context) × 2 (Age: 2.0 or 2.5 years) between-subjects design. Stimulus type (boat or fish) was counterbalanced across participants. The pieces for each stimulus set were placed in one of two arbitrarily predetermined positions that were counterbalanced across participants. Children in the baseline group did not receive a demonstration session and thus were not exposed to the stimuli until the start of the test phase.

Procedure

All protocols were approved by the Georgetown and Binghamton University IRBs. Testing primarily occurred in the home and a small subset ($n = 39$) was tested in the laboratory. The protocol was described to parents prior to obtaining informed consent from all parents. All of the children in the study were given a brief (5–10 min) warm up play session to ensure that they were familiar and comfortable with the experimenter. The apparatus was placed on a small table about one foot high. Before the task began, the apparatus was covered by a black cloth.

Live Demonstration Groups

The pieces were placed on the board behind the black cloth. The experimenter lifted the cloth and showed the toddler how to put the magnet pieces together to make the "boat" or "fish." The experimenter slid each piece by putting two fingers on the center of the piece. Every time a piece was moved, the experimenter made non-specific, fully scripted comments ("Look at this!," "What was that?," and "Isn't that fun?") to orient the child to the demonstration. After moving the pieces into place to create the "boat" or the "fish," the experimenter covered the apparatus

with the black cloth and moved the pieces back into their original locations. The demonstration was repeated three times in total; the three demonstrations together lasted approximately 50 s. After the demonstration was finished the experimenter covered the apparatus with the black cloth again and placed the pieces back into their original locations.

Video Demonstration Group

For this group the apparatus had the metal board removed. The experimenter lifted the cloth to reveal a monitor and played a video of another experimenter demonstrating how to put the puzzle together with the semantically meaningful context. The experimenter in the video presented the same demonstration as the experimenter in the live demonstration condition, including the use of the same scripted language. The video lasts 60 s. After the video was finished, the experimenter inserted the metal board back into the apparatus case, put the black cloth in front of the board, and placed the pieces on the board.

Test Phase

The test phase was the same for the video, live, and baseline groups. A short delay occurred between the end of the demonstration and the start of test. The transition was slightly longer in the transfer condition from video to magnet board ($M = 23.12$ s) than the no transfer condition resetting the pieces on the magnet ($M = 6.81$ s). The experimenter then lifted the black cloth up away from the apparatus and told the child "Now it's your turn!" The test lasted 60 s from the first time the child touched the magnet board or any of the magnet pieces. Following the 60 s test period, the experimenter conducted a manipulation check (demonstrated the target actions one time) and then gave the child the opportunity to reproduce them. The purpose of the manipulation check was to confirm that children were capable of sliding the puzzle pieces. As part of the manipulation check, experimenters asked the child, "What did you make?" (see labeling section) to assess whether children could identify the final puzzle state as a boat or fish. The purpose of the baseline was to assess whether children spontaneously produced the target gestures or goal of connecting the puzzle pieces when they are presented with the stimuli without a demonstration.

Results

Coding

Imitation is operationally defined as duplicating the demonstrated actions at a rate significantly above baseline.

On-Task Behaviors

Each contact with a puzzle piece (beginning when a piece was touched and ending when the touch ended) was coded. Each contact was coded along two dimensions: gesture and goal. On-task behaviors excluded exploratory play (interactions where the piece was removed from the board for more than 3 s) and micro-gestures (a piece was 'nudged,' meaning that it was moved less than 1/6 of the board) that did not result in any type of connection.

Gesture Coding

Coded actions included the following categories of gestures: *correct slide*, *incorrect slide*, *strategy switch*, and *pick up and move*.

Goal Coding

Coded actions that connected puzzle pieces included the following categories of goals: *correct connection, target error connection,* and *connect other*. Based on 30% of all test sessions rescored by a second coder, inter-rater reliability was very good (kappas on each of the subscales; $\kappa_{gesture} = 0.76$, $\kappa_{goal} = 0.81$).

The coded goals and gestures were used to compute four dependent measures (gesture imitation, action fidelity, goal imitation, and goal efficiency). Analyses of gesture imitation and action fidelity (both derived from gesture-coded actions) are presented first followed by analyses of goal imitation and goal efficiency (both derived from goal-directed actions). The coding of action fidelity and goal efficiency is included to more precisely characterize the participants' overall behavior during the test phase. Additionally, labeling performance and a vocabulary measure based on parental report (MCDI) was coded. **Table 1** shows the mean proportion score for each dependent measure for each condition and age group.

Gesture Imitation Score

Following Dickerson et al. (2013), children received credit for each target puzzle piece that they correctly slid, up to a maximum of 3, during the 60 s test period. The resulting gesture imitation score was then converted to a proportion to allow for cross measure comparison. No additional points were given for multiple correct slides with the same puzzle piece.

Action Fidelity Score

To assess the rate at which correct slides were reproduced relative to other, less faithful actions, an action fidelity measure was calculated by taking the sum of all correct slides produced in the testing period (prior to reset following first puzzle completion) and dividing by all on-task behaviors produced (prior to reset following first puzzle completion). Higher proportions indicate more faithful reproduction of demonstrated actions; lower proportions indicate increasing numbers of non-demonstrated actions were produced during the test.

Goal Imitation Score

Following Dickerson et al. (2013), children received one point for each correct connection (maximum = 2). As with the gesture imitation score, the goal imitation score was then converted to a proportion (out of two). The goal imitation score is distinct from the gesture imitation and the action fidelity scores in that if a child used an incorrect gesture to correctly connect two puzzle pieces, they still received a point for the goal.

Goal Efficiency Score

This measure is calculated as all correct connections performed as a proportion of all on-task behaviors prior to first puzzle completion. This measure allows participants to be classified on a continuum, with higher proportions being indicative of highly efficient puzzle reproduction on one end to failure to reproduce the puzzle at all on the other end. For example, for the boat puzzle, a child might simply move the two sails to most efficiently complete the puzzle, but another child might imitate by first moving the brown mast and then the sails. Even less efficiently, another child may produce 20 on-task behaviors in the course of making the puzzle.

Vocabulary Measure

For the MCDI parental report questionnaire, percentile rank scores were calculated from raw scores using age and gender norms (Fenson et al., 2000). The mean percentile ranks were in the average range for 2 year olds ($M = 42.31$, SD $= 30.59$) and for 2.5-year-olds ($M = 38.11$, SD $= 27.05$).

Labeling

Coders recorded if the child generated the object label, either "fish" or "boat" (or synonyms of the object), and when identification first occurred, either during the demonstration, test, or post-test phase with minimal prompting (i.e., "What did you make?"). A score of 0 was given if no label was produced during any phase; a score of 1 was given if a child produced a label during any phase. Parental report from the MCDI collected prior to test indicated that the majority of children had fish (79%) and boat (85%) in their vocabulary.

Data Analysis Plan

First we conducted a preliminary analysis on experimental groups on each of the four dependent measures (**Table 1**). Next

---

TABLE 1 | Mean proportions and SEs for each dependent measure across age, transfer type, and context.

		Measures							
		Gesture proportion		Action fidelity		Goal proportion		Goal efficiency	
Transfer type	Context	2.0 years	2.5 years	2.0 years	2.5 years	2.0 years	2.5 years	2.0 years	2.5 years
Baseline	None	0.06 (0.03)	0.08 (0.06)	0.07 (0.04)	0.05 (0.03)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	Ocean	0.00 (0.00)	0.14 (0.06)	0.00 (0.00)	0.11 (0.06)	0.00 (0.00)	0.04 (0.04)	0.00 (0.00)	0.03 (0.03)
Video	None	0.24 (0.11)	0.12 (0.06)	0.11 (0.06)	0.05 (0.03)	0.23 (0.12)	0.39 (0.13)	0.16 (0.10)	0.25 (0.10)
	Ocean	0.19 (0.08)	0.25 (0.09)	0.08 (0.04)	0.10 (0.05)	0.32 (0.12)	0.42 (0.15)	0.22 (0.09)	0.36 (0.13)
Live	None	0.40 (0.09)	0.28 (0.08)	0.24 (0.07)	0.24 (0.07)	0.24 (0.09)	0.81 (0.09)	0.11 (0.05)	0.42 (0.08)
	Ocean	0.45 (0.09)	0.31 (0.10)	0.36 (0.08)	0.12 (0.04)	0.61 (0.12)	0.57 (0.13)	0.25 (0.06)	0.24 (0.06)
Measure totals		0.23 (0.04)	0.20 (0.03)	0.15 (0.03)	0.11 (0.02)	0.23 (0.04)	0.38 (0.05)	0.12 (0.03)	0.22 (0.04)

---

we assessed whether performance between groups, both experimental and baseline, differ as a function of age (2.0, 2.5 years), transfer type (video, live), and context (ocean, none). Third, excluding baseline participants, we conducted a first order correlational analysis to assess which of our demographic factors, experimental conditions, labeling behavior, and vocabulary were associated with performance on the four dependent measures (see **Table 2**). Based on the pattern of results in our correlational analysis, we conducted a multivariate linear regression on the goal imitation measure (see **Table 3**).

Preliminary analyses

Preliminary analyses on gesture imitation and action fidelity revealed no main effects of gender, stimulus type, or latency between the demonstration and test session and only entered one interaction, which did not survive follow-up analyses. Therefore, gender, stimulus, and latency between demonstration and test will not be considered further for gesture imitation or action fidelity.

Preliminary analyses on goal imitation and goal efficiency, revealed no main effects of gender or latency between the demonstration and test session. These variables only entered one interaction, which did not survive follow-up analyses. For

goal imitation, stimulus type did enter into significant 2-way interactions and will be analyzed further. There was a main effect of stimulus type, $F(1,95) = 18.82$, $p < 0.001$, $\eta_p^2 = 0.17$, and an interaction between context and stimulus type, $F(1,95) = 5.90$, $p < 0.05$, $\eta_p^2 = 0.06$; performance was highest for the boat puzzle in the ocean context ($M = 0.74$, SD $= 0.42$), which was significantly higher than the boat without context ($M = 0.50$, SD $= 0.48$), $p < 0.01$. Performance with the fish puzzle was not affected by context (context: $M = 0.33$, SD $= 0.43$; none: $M = 0.22$, SD $= 0.38$; see **Figure 3**). A similar pattern of results emerged for goal efficiency. These effects involving stimulus type will be discussed further; see Correlational analysis section.

Gesture Imitation and Action Fidelity Analysis
Gesture Imitation

A 3 (transfer type: baseline, video, live) × 2 (age: 2.0, 2.5 years) × 2 (context: ocean, none) ANOVA on gesture imitation yielded a main effect of transfer type, $F(2,153) = 14.45$, $p < 0.001$, $\eta_p^2 = 0.16$. Gesture imitation following a 'live' demonstration ($M = 0.37$, SD $= 0.35$) was significantly higher than following a video demonstration ($M = 0.20$, SD $= 0.30$), which did not differ from baseline ($M = 0.07$, SD $= 0.16$). There was no

TABLE 2 | First order correlation between context, stimulus, age, gender, labeling, and the four dependent variables.

	Goal imitation	Goal efficiency	Gesture imitation	Action fidelity	Transfer	Context	Age	Gender	Stimulus	Label
Goal efficiency	0.854**	–	–	–	–	–	–	–	–	–
Gesture imitation	0.169	0.014	–	–	–	–	–	–	–	–
Action fidelity	0.198*	0.058	0.772**	–	–	–	–	–	–	–
Transfer condition	–0.202*	0.009	–0.260**	–0.336**	–	–	–	–	–	–
Context	0.088	0.063	0.044	–0.012	0.064	–	–	–	–	–
Age	0.202*	0.201*	–0.161	–0.178	0.060	0.009	–	–	–	–
Gender	0.013	0.033	–0.060	–0.070	–0.006	0.009	–0.063	–	–	–
Stimulus type	0.359**	0.397**	–0.127	–0.040	0.114	0.064	0.061	–0.118	–	–
Labeling	0.122	0.204*	–0.004	–0.006	0.005	0.083	0.187	–0.210**	0.315**	–
MCDI	0.187	0.251**	0.116	0.195*	0.116	0.074	–0.108	–0.168	0.232*	0.101

$*p \leq 0.05$, $**p \leq 0.01$.

TABLE 3 | The regression models for goal imitation and gesture imitation performance.

	Goal imitation				Gesture imitation			
	Unstandardized coefficient		Standardized coefficient		Unstandardized coefficient		Standardized coefficient	
	B	SE	β	*t*	*B*	SE	β	*t*
(Constant)	0.443	0.039		11.33**	0.287	0.031		9.15**
Transfer	–0.246	0.079	–0.264	–3.10**	–0.167	0.064	–0.246	–2.62*
Context	0.082	0.079	0.088	1.04	0.052	0.063	0.078	0.826
Age	0.203	0.080	0.219	2.54*	–0.094	0.064	–0.140	–1.47
Stimulus	0.365	0.083	0.394	4.38**	–0.068	0.067	–0.101	–1.02
Labeling	–0.055	0.086	–0.057	–0.63	0.035	0.069	0.050	0.500
Transfer × labeling	0.400	0.165	0.209	2.43*	0.199	0.132	0.144	1.51
Transfer × age	–0.211	0.161	–0.113	–1.31	0.038	0.129	0.028	0.292

$*p \leq 0.05$, $**p \leq 0.01$.

FIGURE 3 | The effect of context and stimulus type on goal proportion score.
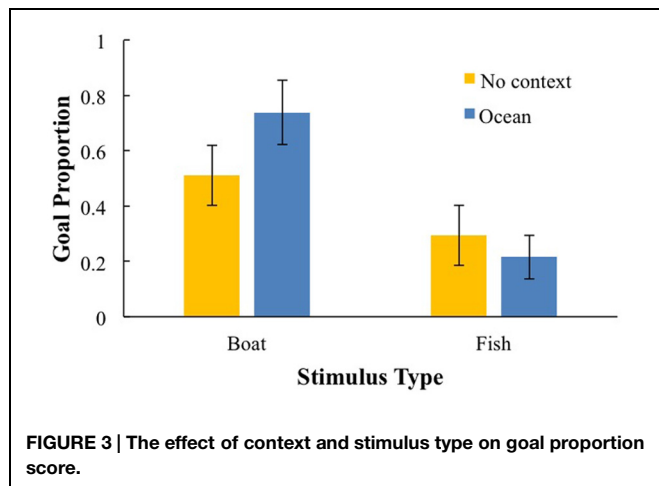
main effect of age, $F < 1$, or context, $F < 1$, and no significant interactions.

Action Fidelity

A 3 (transfer type: baseline, video, live) × 2 (age: 2.0, 2.5 years) × 2 (context: ocean, none) ANOVA on action fidelity revealed a main effect of transfer type, $F(2,153) = 13.99$, $p < 0.001$, $\eta_p^2 = 0.15$. As with gesture imitation, action fidelity was higher following live ($M = 0.24$, SD = 0.27) than video demonstrations ($M = 0.08$, SD = 0.14) and baseline ($M = 0.06$, SD = 0.14). Again, the video group did not significantly exceed baseline performance, a clear demonstration of poor learning from video. Neither age nor context were significant ($F < 1$), but a 3-way interaction between age, transfer type and context was observed; $F(2,153) = 3.74$, $p < 0.05$, $\eta_p^2 = 0.05$. To follow up this 3-way interaction, we conducted Tukey HSD *post hoc* tests ($p < 0.01$). There were no differences among the baseline and video groups. Among the live demonstration groups, although it did not reach statistical significance, the effect appears to be driven by the 2-year-old context group ($M = 0.36$, SD = 0.30) that showed elevated action fidelity performance relative to the other groups: 2- and 2.5-year-olds without context (2.0: $M = 0.24$, SD = 0.32; 2.5: $M = 0.24$, SD = 0.26) and 2.5-year-olds with context ($M = 0.12$, SD = 0.14). In summary, these analyses indicate that context did not ameliorate the transfer deficit.

Goal Imitation Score and Goal Efficiency Score Analysis

Goal Imitation

A 3 (transfer type: baseline, video, live) × 2 (age: 2.0, 2.5 years) × 2 (context: ocean, none) ANOVA performed on goal imitation yielded a main effect of age, $F(1,153) = 5.96$, $p < 0.05$, $\eta_p^2 = 0.04$, and transfer type, $F(2,153) = 31.60$, $p < 0.001$, $\eta_p^2 = 0.29$, but no effect of context ($F < 1$). A follow-up Tukey HSD test on transfer type demonstrated a clear transfer deficit; the live demonstration group imitated significantly more goal actions ($M = 0.52$, SD = 0.45) compared to the video group ($M = 0.34$, SD = 0.46), which was above baseline levels

($M = 0.01$, SD = 0.07). This effect was qualified by a significant three-way interaction between age, context, and condition, $F(2,153) = 3.20$, $p < 0.05$, $\eta_p^2 = 0.04$. To follow-up the 3-way interaction, we conducted Tukey HSD *post hoc* tests ($p < 0.01$). The 3-way effect was largely confined to an age-related difference in the live condition; 2-year-olds who received a context-backed demonstration ($M = 0.61$, SD = 0.45) showed higher goal performance relative to all other groups involving 2-year-olds (M's ranged from 0.23–0.32). 2.5-year-olds showed the standard transfer deficit, with live conditions eliciting generally better performance than video conditions. The baseline conditions did not significantly rise above zero. Importantly, the addition of context did not ameliorate the transfer deficit at either age.

Goal Efficiency

A 3 (transfer type: baseline, video, live) × 2 (age: 2.0, 2.5 years) × 2 (context: ocean, none) ANOVA on goal imitation produced a main effect of age, $F(1,153) = 5.28$, $p < 0.05$, $\eta_p^2 = 0.03$ and condition, $F(2,153) = 16.55$, $p < 0.001$, $\eta_p^2 = 0.18$; live ($M = 0.24$, SD = 0.26) and video ($M = 0.25$, SD = 0.37) groups did not differ from one another, but both were significantly above baseline ($M = 0.01$, SD = 0.04). No other main effects or interactions emerged.

Correlational Analysis

In order to assess which factors were associated with imitation performance across the four dependent measures, a first-order correlation matrix was constructed. This included demographic factors, experimental factors, and naming and vocabulary variables. Review of the correlation matrix reveals that gesture imitation and action fidelity, not surprisingly, are associated with one another, $r(109) = 0.77$, $p < 0.01$, but are associated with few of the other variables except for transfer. Goal imitation and goal efficiency, as expected, are also associated with one another, $r(109) = 0.85$, $p < 0.01$. These two are also associated with action fidelity, age, transfer, and stimulus type. This pattern of results suggests that factors predicting goal imitation may differ from those predicting gesture imitations. To explore this idea further, a regression model with goal imitation as the outcome variable was constructed. A second model was constructed with gesture imitation using the same predictors as well, to enable examination of these measures separately.

Predicting Goal Imitation

This analysis was conducted to identify factors associated with enhanced transfer performance. Transfer type (live, video), context (none or ocean), stimulus type (fish or boat), age (2, 2.5 years) and labeling (yes or no) were included in a multivariate linear regression on goal imitation performance. A labeling × transfer interaction term and age × transfer interaction term were entered simultaneously as well. All predictor means were centered. Interaction terms were calculated using the centered means. Although a number of first order correlations were significant (see **Table 2**), there was no multi-collinearity in the model; VIFs range from 1.03 to 1.16. Given that our prior ANOVA analyses had previously determined that neither context nor stimulus entered into significant interactions with transfer,

these interaction terms were not included in the final regression model. Results from the regression are presented in **Table 3**. The overall model was significant, $F(7,101) = 5.67$, $p < 0.001$, $R = 0.53$, $R^2 = 0.28$. As expected, transfer type (live, video) was a significant predictor, indicating the transfer deficit. There was a main effect of age; older children showed higher goal imitation overall, and of stimulus type (fish, boat), demonstrating that children connected more pieces with the boat puzzle than the fish puzzle, an effect that has been reported in prior work.

Labeling alone did not predict goal imitation. There was, however, a significant interaction between transfer type and labeling: The advantage of a self-generated label was 0.40 points greater when it was combined with the transfer condition than when it was not. Follow-up regressions were conducted to examine the simple slopes for toddlers who had labeled and those who had not, as a function of transfer. Although there was a significantly negative effect when children did not self-generate a label, $F(1,66) = 7.53$, $p < 0.04$, $B = -0.29$, $\beta = -0.32$, there was no difference in the slope when children did generate a label, $F(1,42) < 1$. This analysis supports the interpretation that the *ability to label enabled these children to make the far transfer "jump."* Thus, the impact of the transfer deficit was ameliorated for children who generated an object label during the test phase. The transfer type by self-generated label interaction is depicted in **Figure 4**. As shown, children in the far transfer (video) condition who generated a label for the puzzle produced significantly higher imitation scores than children in the video condition who did not (see **Figure 4**). No other effects were significant.

The same model was marginally significant for gesture imitation, $F(7,101) = 2.10$, $p = 0.051$, $R = 0.36$, $R^2 = 0.13$. As shown in **Table 3**, transfer type was the only significant predictor, once again demonstrating that performance of the video group was significantly worse than the live group. No other associations were significant, including interaction terms. Comparison of the models suggests that factors that reduce the transfer deficit for goal imitation are not the same as for gesture imitation. Other models including additional interaction terms were conducted,

but these interactions were not significant and the overall model did not explain more of the variance.

General Discussion

Consistent with previous findings, this study showed that young children displayed a significant transfer deficit. Two- and 2.5-year-old children who received a video demonstration reproduced significantly fewer gestures and goals than children receiving a live demonstration. Consistent with our hypothesis, we did find an age-related effect of context in the live condition. Contrary to our hypothesis, we found that the addition of a semantically meaningful visual context did not ameliorate the transfer deficit. Importantly though, the context did not interfere with learning either. This finding is consistent with a point in development where children may form representations that contain primarily central cue information, resulting in neither a disruptive nor a facilitative effect of background context. There were individual differences in self-generation of a label that were associated with better performance for the transfer group on the puzzle task.

Rather than being impacted by context, the transfer deficit was ameliorated when children were able to generate a verbal label for the puzzle. There was a significant positive correlation between a child's ability to generate a label for the completed puzzle and their ability to correctly connect the puzzle pieces (goal proportion). This labeling effect only facilitated performance for children in the video group, however. That is, it was the self-generated labeling of the object, and not the semantically relevant context that facilitated transfer of goal learning, highlighting the importance of a pre-existing object representation that facilitated transfer across 2D and 3D demonstration and test phases. This finding adds to the growing body of research suggesting that self-generating the object label enhances young children's performance (e.g., Miller and Marcovitch, 2011). Other studies have also demonstrated that vocabulary size predicts object recognition, such as in Smith (2003), who reported a positive correlation between language and recognition. Smith found that although 18- to 24-month-olds (with smaller vocabularies) and children (with larger vocabularies) were able to recognize richly detailed instances of an object equally well, children with smaller *noun* vocabularies performed at chance levels when presented with a more perceptually challenging recognition task that included less iconic images of shapes. Simcock and Hayne (2002) used a 'magic shrinking machine' task to assess children's understanding of the actions required to operate the box and objects that were made smaller, following a long delay. Their results, that children's verbal reports following the delay matched their verbal skill during the encoding event rather than their verbal skill when tested 6 months or one year later (Simcock and Hayne, 2002), highlight the importance of children's productive vocabulary at the time of encoding.

Taken together, prior research suggests that object labels may help establish abstract and dual representations of objects, as well as direct attention to relevant task details (Miller and Marcovitch, 2011). This research is consistent with our finding that the label

FIGURE 4 | Goal imitation proportion as a function of transfer condition (near live or far video condition) and production of self-generated label.
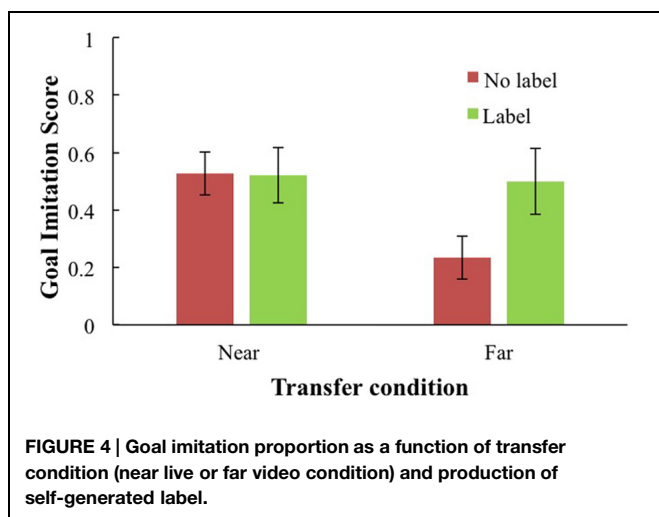
serves as an effective retrieval cue for children in the present study on a far transfer task. Language can enhance recognition and learning under perceptually impoverished conditions and high cognitive load. Transfer distance increases cognitive load and the label acts as a cue that facilitates both encoding and retrieval (Simcock and Hayne, 2002; Hayne and Herbert, 2004; Troseth, 2010; Miller and Marcovitch, 2011).

Corresponding research on experimenter-generated verbal cues suggests that these cues are not as robust under challenging learning conditions. Bates et al. (1989) found support for the argument that congruent experimenter-generated language cues facilitated imitation performance in 1-year-olds (see also Gerson and Woodward, 2013). These results suggest that the use of relevant language enhances object recognition and imitation under conditions where there is no transfer. The same was not necessarily true, however, for a transfer task. Zack et al. (2013) found that neither a nonsense nor a meaningful object label facilitated 15-month-olds' imitation on a touchscreen transfer task (2D to 3D or 3D to 2D). There are, however, age-related differences in the effectiveness of verbal cues. Studies of narration effects are important to consider as verbal cues have semantic or referential meaning and may be more effective retrieval cues than non-verbal auditory cues. Studies of the effect of narrative cues during an imitation task with 18- and 24-month-olds suggest infants can imitate from TV or books when verbal descriptions are not available. Additionally they can rely on verbal cues when images of the objects are absent (Simcock et al., 2011). There is likely to be a bidirectional effect whereby language affects learning and vice versa. Language development is also associated with domain-general processes such as individual differences in working memory and long-term retention. Understanding how these factors are related to transfer learning requires further empirical investigation.

The label may not be the only factor that facilitates transfer. The presence of the label suggests that the child possesses a representation of the object; not just as a single encoded exemplar, but rather, what Rosch et al. (1976) first called an *entry level category*. This indicates that the child possesses a generalizable representation of "boat" or "fish" in the present test, potentially allowing these children to access this category from either the 2D image or the 3D puzzle, in agreement with Hayne's (2004) concept of memory flexibility. In other words, the children do not have to recognize the two instantiations as the same thing precisely, but only as exemplars of the same category. The stimulus effects described above (the boat puzzle, in general, invoked better performance than the fish puzzle, and ameliorated the transfer deficit when the child could produce the label) supports the argument that construction of the puzzle, independent of the other manipulations, affected access to a representation. The boat puzzle, with its clearly parsed sails and recognizable mast, displays a set of parts that map onto a mid-level visual representation of the type described by Biederman (1987; Hummel and Stankiewicz, 1996; see also Schacter et al., 1990; Biederman and Cooper, 1991; Biederman and Gerhardstein, 1993). The puzzle pieces that make up the fish, however, do not clearly correspond to certain recognizable parts of a fish (head or fins). Further,

the pieces are of different colors, which are highly unlikely to correspond to any prior 'fish' exemplars that the majority of the children tested would have experienced, making access to a category more difficult even in cases where a child does possess such a category. This interpretation is bolstered by the finding that the presentation of context facilitated performance when the boat, but not the fish puzzle, was demonstrated.

In the present study, individual differences in the self-generated labels were associated with transfer performance. This outcome provides a potential explanation for how the semantic context facilitated performance on the puzzle task. This interpretation has limitations. It is possible that children who did not generate the label spontaneously may have known the label but did not express it, or that they may still benefit from a label or verbal cue being provided by the experimenter during the demonstration. Future studies could address this systematically by including labels during the demonstration phase. Also, future studies should seek to investigate whether other individual differences such as working memory or experience with puzzles are associated with performance.

Experimenter-generated nonverbal cues (i.e., visual context) in the present study did not reduce the transfer deficit but did improve overall goal performance. The lack of a main effect of nonverbal semantic context was surprising, but is consistent with similar difficulty in utilization of experimenter-generated verbal cues as discussed above and with the (non-iconic) perceptual properties of the "ocean" context used in the present study. Alternatively, this lack of a semantic context effect could be explained by accounts of developmental changes in memory binding. Research on memory binding suggests that after infancy, central and peripheral details are no longer fused, and children may disregard peripheral and contextual information and focus on more central details. A more salient foreground object may prevent toddlers from utilizing the background cues available because these cues are less salient. Consequently, children may not automatically bind the context to the memory as they did earlier in development. The attention system may focus on central details with overall less binding, resulting in neither a facilitative nor a disruptive effect of context. Processing of central and peripheral details and binding may become more flexible with further development (e.g., Sluzenski et al., 2006; Lloyd et al., 2009; Bornstein et al., 2011; Olson and Newcombe, 2014). This progression would ultimately result in a facilitatory effect of context without disruptive effects under conditions of context change. This progression may track developmental changes in the hippocampus (Olson and Newcombe, 2014; see also Chalfonte and Johnson, 1996; Mitchell et al., 2000 for discussion of age-related decline in hippocampal functioning and flexible memory binding). The developmental trajectory of memory binding during early childhood requires additional empirical attention. Additional research is also necessary to ascertain whether older children use contextual information to form flexible adult-like memories that contain both background and central information that can be used under similar complex transfer learning conditions (see also Olson and Newcombe, 2014).

Other factors more proximal to the puzzle imitation task may have limited children's ability to utilize the contextual cues.

Puzzle complexity in this task was high. It is important to note that to make our task ecologically valid we deliberately used cartoon-like and abstract representations for both the puzzle pieces and the background context. Many educational applications include animated (low iconicity) images because these images are easier to program. However, the lower iconicity of the context in the present task may have limited children's ability to utilize contextual cues. Future studies could include more iconic representations of both stimuli (boats and fish) and context (e.g., fins and eyes on the fish or photographic images of the ocean). There are also likely to be individual differences in attention to pieces, gestures, and the context background; assessing visual attention to the context and puzzle pieces using eye-tracking may prove fruitful in this regard. The present study adds to a growing body of literature showing that the transfer deficit persists into toddlerhood (Dickerson et al., 2013, see also McGuigan et al., 2007; Moser et al., 2015). The ocean context facilitated completion of the boat puzzle relative to the fish puzzle. In addition, self-generated labeling of the puzzle (boat or fish) elevated goal performance of those in the video demonstration condition. This suggests that object identification can ameliorate the transfer deficit during toddlerhood. Understanding the nature of visuo-spatial integration (Bremner, 1978; Lockman, 2000; Kirkorian and Pempek, 2013) and spatial development more generally in early childhood has important implications for both parents and educators (see Levine et al., 2012 for related discussion). This puzzle imitation far transfer task provides a unique opportunity to examine the role of multiple factors that influence cognitive development.

Acknowledgments

References

Barnat, S. B., Klein, P. J., and Meltzoff, A. N. (1996). Deferred imitation across changes in context and object: memory and generalization in 14-month-old infants. *Infant Behav. Dev.* 19, 241–251. doi: 10.1016/S0163-6383(96)90023-5

Barnett, S. M., and Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychol. Bull.* 128, 612–637. doi: 10.1037/0033-2909.128.4.612

Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: informing theory and practice. *Dev. Rev.* 30, 128–154. doi: 10.1016/j.dr.2010.03.001

Barr, R. (2013). Memory constraints on infant learning from picture books, television, and touchscreens. *Child Dev. Perspect.* 7, 205–210. doi: 10.1111/cdep.12041

Barr, R., and Hayne, H. (1999). Developmental changes in imitation from television during infancy. *Child Dev.* 70, 1067–1081. doi: 10.1111/1467-8624.00079

Barr, R., Muentener, P., Garcia, A., Fujimoto, M., and Chavez, V. (2007). Age-related changes in deferred imitation from television by 6- to 18-month-olds. *Dev. Sci.* 10, 910–921. doi: 10.1111/j.1467-7687.2007.00641.x

Bates, E., Thal, D., Whitesell, K., Fenson, L., and Oakes, L. (1989). Integrating language and gesture in infancy. *Dev. Psychol.* 25, 1004–1019. doi: 10.1037/0012-1649.25.6.1004

Biederman, I. (1972). Perceiving real-world scenes. *Science* 177, 77–80. doi: 10.1126/science.177.4043.77

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147. doi: 10.1037/0033-295X.94.2.115

Biederman, I., and Cooper, E. E. (1991). Priming contour-deleted images: evidence for intermediate representations in visual object priming. *Cogn. Psychol.* 23, 393–419. doi: 10.1016/0010-0285(91)90014-F

Biederman, I., and Gerhardstein, P. (1993). Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 1162–1182. doi: 10.1037/0096-1523.19.6.1162

Boller, K., Rovee-Collier, C., Gulya, M., and Prete, K. (1996). Infants' memory for context: timing effects of postevent information. *J. Exp. Child Psychol.* 63, 583–602. doi: 10.1006/jecp.1996.0063

Bornstein, M. H., Mash, C., and Arterberry, M. E. (2011). Perception of object–context relations: eye-movement analyses in infants and adults. *Dev. Psychol.* 47, 364. doi: 10.1037/a0021059

Borovsky, D., and Rovee-Collier, C. (1990). Contextual constraints on memory retrieval at 6 months. *Child Dev.* 61, 1569–1583. doi: 10.2307/1130765

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychol. Bull.* 114, 80–99. doi: 10.1037/0033-2909.114.1.80

Bremner, J. G. (1978). Egocentric versus allocentric spatial coding in nine-month-old infants: factors influencing the choice of code. *Dev. Psychol.* 14, 346. doi: 10.1037/0012-1649.14.4.346

Carver, L. J., Meltzoff, A. N., and Dawson, G. (2006). Event-related potential (ERP) indices of infants' recognition of familiar and unfamiliar objects in two and three dimensions. *Dev. Sci.* 9, 51–62. doi: 10.1111/j.1467-7687.2005.00463.x

Chalfonte, B. L., and Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Mem. Cognit.* 24, 403–416. doi: 10.3758/BF03200930

Chun, M. M., and Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cogn. Psychol.* 36, 28–71. doi: 10.1006/cogp.1998.0681

Cohen, N. J., and Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System*. Cambridge, MA: MIT press.

Csibra, G., and Gergely, G. (2006). "Social learning and social cognition: the case for pedagogy," in *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*, eds Y. Munakata and M. H. Johnson (Oxford: Oxford University Press), 249–274.

Deloache, J., Simcock, G., and Marzolf, D. P. (2004). Transfer by very young children in the symbolic retrieval task. *Child Dev.* 75, 1708–1718. doi: 10.1111/j.1467-8624.2004.00811.x

Dickerson, K., Gerhardstein, P., Zack, E., and Barr, R. (2013). Age-related changes in learning across early childhood: a new imitation task. *Dev. Psychobiol.* 55, 719–732. doi: 10.1002/dev.21068

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., and Thal, D. (2000). Reply: measuring variability in early child language: don't shoot the messenger. *Child Dev.* 71, 323–328. doi: 10.1111/1467-8624.00147

Flynn, E., and Whiten, A. (2008). Imitation of hierarchical structure versus component details of complex actions by 3- and 5-year-olds. *J. Exp. Child Psychol.* 101, 228–240. doi: 10.1016/j.jecp.2008.05.009

Gerson, S. A., and Woodward, A. L. (2013). The goal trumps the means: highlighting goals is more beneficial than highlighting means in means-end training. *Infancy* 18, 289–302. doi: 10.1111/j.1532-7078.2012.00112.x

Hartshorn, K., and Rovee-Collier, C. (1997). Infant learning and long-term memory at 6 months: a confirming analysis. *Dev. Psychobiol.* 30, 71–85. doi: 10.1002/(SICI)1098-2302(199701)30:1<71::AID-DEV7>3.0.CO;2-S

Hayne, H. (2004). Infant memory development: implications for childhood amnesia. *Dev. Rev.* 24, 33–73. doi: 10.1016/j.dr.2003.09.007

Hayne, H., Boniface, J., and Barr, R. (2000). The development of declarative memory in human infants: age-related changes in deferred imitation. *Behav. Neurosci.* 114, 77–83. doi: 10.1037/0735-7044.114.1.77

Hayne, H., and Herbert, J. (2004). Verbal cues facilitate memory retrieval during infancy. *J. Exp. Child Psychol.* 89, 127–139. doi: 10.1016/j.jecp.2004.06.002

Hummel, J. E., and Stankiewicz, B. J. (1996). "An architecture for rapid, hierarchical structural description," in *Attention and Performance XVI: Information*

Integration in Perception and Communication, eds T. Inui and J. L. McClelland (Cambridge, MA: MIT Press), 93–121.

Kirkorian, H. L., and Pempek, T. A. (2013). Toddlers and touch screens: potential for early learning? *Zero three* 33, 32–37.

Learmonth, A. E., Lamberth, R., and Rovee Collier, C. (2004). Generalization of deferred imitation during the first year of life. *J. Exp. Child Psychol.* 88, 297–318. doi: 10.1016/j.jecp.2004.04.004

Levine, S. C., Ratliff, K. R., Huttenlocher, J., and Cannon, J. (2012). Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Dev. Psychol.* 48, 530–542. doi: 10.1037/a0025913

Lloyd, M. E., Doydum, A. O., and Newcombe, N. S. (2009). Memory binding in early childhood: evidence for a retrieval deficit. *Child Dev.* 80, 1321–1328. doi: 10.1111/j.1467-8624.2009.01353.x

Lockman, J. J. (2000). A perception–action perspective on tool use development. *Child Dev.* 71, 137–144. doi: 10.1111/1467-8624.00127

McGuigan, N., Whiten, A., Flynn, E., and Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. *Cogn. Dev.* 22, 353–364. doi: 10.1016/j.cogdev.2007.01.001

Miller, S. E., and Marcovitch, S. (2011). Toddlers benefit from labeling on an executive function search task. *J. Exp. Child Psychol.* 108, 580–592. doi: 10.1016/j.jecp.2010.10.008

Mitchell, K. J., Johnson, M. K., Raye, C. L., Mather, M., and D'Esposito, M. (2000). Aging and reflective processes of working memory: binding and test load deficits. *Psychol. Aging* 15, 527–541. doi: 10.1037/0882-7974.15.3.527

Moser, A., Zimmermann, L., Dickerson, K., Grenell, A., Barr, R. and Gerhardstein, P. (2015). They can interact, but can they learn? Toddlers' transfer learning from touchscreens and television. *J. Exp. Child Psychol.* doi: 10.1016/j.jecp.2015.04.002

Nakao, K., and Treas, J. (1994). Updating occupational prestige and socioeconomic scores: how the new measures measure up. *Sociol. Methodol.* 24, 1–72.

Newcombe, N. S., Balcomb, F., Ferrara, K., Hansen, M., and Koski, J. (2014). Two rooms, two representations? Episodic-like memory in toddlers and preschoolers. *Dev. Sci.* 17, 743–756. doi: 10.1111/desc.12162

Nielsen, M., Simcock, G., and Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Dev. Sci.* 11, 722–731. doi: 10.1111/j.1467-7687.2008.00722.x

Oliva, A., and Torralba, A. (2007). The role of context in object recognition. *Trends Cogn. Sci.* 11, 520–537. doi: 10.1016/j.tics.2007.09.009

Olson, I. R., and Newcombe, N. S. (2014). "Binding together the elements of episodes: relational memory and the developmental trajectory of the hippocampus," in *The Wiley Handbook on the Development of Children's Memory*, eds P. J. Bauer and R. Fivush (Chichester: Wiley-Blackwell), 285–308. doi: 10.1002/9781118597705.ch13

Pereira, A. F., and Smith, L. B. (2009). Developmental changes in visual object recognition between 18 and 24 months of age. *Dev. Sci.* 12, 67–80. doi: 10.1111/j.1467-7687.2008.00747.x

Raj, V., and Bell, M. A. (2010). Cognitive processes supporting episodic memory formation in childhood: the role of source memory, binding, and executive functioning. *Dev. Rev.* 30, 384–402. doi: 10.1016/j.dr.2011.02.001

Rescorla, R. A., and Wagner, A. W. (1972). "A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement," in *Classical Conditioning II: Current Research and Theory*, eds A. H. Black and W. F. Prokasy (New York, NY: Appleton-Century-Crofts), 64–99.

Richmond, J., and Nelson, C. A. (2009). Relational memory during infancy: evidence from eye tracking. *Dev. Sci.* 12, 549–556. doi: 10.1111/j.1467-7687.2009.00795.x

Rideout, V. (2013). *Zero to Eight: Children's Media Use in America 2013*. San Francisco, CA: Common Sense Media.

Rideout, V. (2014). *Learning at Home: Families' Educational Media Use in America*. New York, NY: Joan Ganz Cooney Center.

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., and Boyes-Braem, P. (1976). Basic objects in natural categories. *Cogn. Psychol.* 8, 382–439. doi: 10.1016/0010-0285(76)90013-X

Schacter, D. L., Cooper, L. A., and Delaney, S. M. (1990). Implicit memory for visual objects and the structural description system. *Bull. Psychon. Soc.* 28, 367–372. doi: 10.3758/BF03334043

Shields, P. J., and Rovee-Collier, C. (1992). Long-term memory for context-specific category information at 6 months. *Child Dev.* 63, 245–259. doi: 10.2307/1131476

Simcock, G., and DeLoache, J. (2006). Get the picture? The effects of iconicity on toddlers' reenactment from picture books. *Dev. Psychol.* 42, 1352–1357. doi: 10.1037/0012-1649.42.6.1352

Simcock, G., and Dooley, M. (2007). Generalization of learning from picture books to novel test conditions by 18-and 24-month-old children. *Dev. Psychol.* 43, 1568–1578. doi: 10.1037/0012-1649.43.6.1568

Simcock, G., Garrity, K., and Barr, R. (2011). The effect of narrative cues on infants' imitation from television and picture books. *Child Dev.* 82, 1607–1619. doi: 10.1111/j.1467-8624.2011.01636.x

Simcock, G., and Hayne, H. (2002). Breaking the barrier? Children fail to translate their preverbal memories into language. *Psychol. Sci.* 13, 225–231. doi: 10.1111/1467-9280.00442

Sluzenski, J., Newcombe, N. S., and Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: a developmental perspective. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 89–100. doi: 10.1037/0278-7393.32.1.89

Sluzenski, J., Newcombe, N. S., and Satlow, E. (2004). Knowing where things are in the second year of life: implications for hippocampal development. *J. Cogn. Neurosci.* 16, 1443–1451. doi: 10.1162/0898929042304804

Smith, L. B. (2003). Learning to recognize objects. *Psychol. Sci.* 14, 244–250. doi: 10.1111/1467-9280.03439

Spear, N. E., and McKinzie, D. L. (1994). "Intersensory integration in the infant rat," in *The Development of Intersensory Perception: Comparative Perspectives*, eds D. J. Lewkowicz and R. Lickliter (Hillsdale, NJ: Erlbaum), 133–161.

Subiaul, F., Anderson, S., Brandt, J., and Elkins, J. (2012). Multiple imitation mechanisms in children. *Dev. Psychol.* 48, 1165–1179. doi: 10.1037/a0026646

Suddendorf, T., Simcock, G., and Nielsen, M. (2007). Visual self-recognition in mirrors and live videos: evidence for a developmental asynchrony. *Cogn. Dev.* 22, 185–196. doi: 10.1016/j.cogdev.2006.09.003

Tennie, C., Call, J., and Tomasello, M. (2006). Push or pull: imitation vs. emulation in great apes and human children. *Ethology* 112, 1159–1169. doi: 10.1111/j.1439-0310.2006.01269.x

Troseth, G. L. (2010). Is it life or is it Memorex? Video as a representation of reality. *Dev. Rev.* 30, 155–175. doi: 10.1016/j.dr.2010.03.007

Troseth, G. L., and DeLoache, J. S. (1998). The medium can obscure the message: young children's understanding of video. *Child Dev.* 69, 950–965. doi: 10.1111/j.1467-8624.1998.tb06153.x

Troseth, G. L., Saylor, M. M., and Archer, A. H. (2006). Young children's use of video as a source of socially relevant information. *Child Dev.* 77, 786–799. doi: 10.1111/j.1467-8624.2006.00903.x

Zack, E., Barr, R., Gerhardstein, P., Dickerson, K., and Meltzoff, A. N. (2009). Infant imitation from television using novel touch screen technology. *Br. J. Dev. Psychol.* 27, 13–26. doi: 10.1348/026151008X334700

Zack, E., Gerhardstein, P., Meltzoff, A. N., and Barr, R. (2013). 15-month-olds' transfer of learning between touch screen and real-world displays: language cues and cognitive loads. *Scand. J. Psychol.* 54, 20–25. doi: 10.1111/sjop.12001

frontiers
in Psychology

She called that thing a mido, but should you call it a mido too? Linguistic experience influences infants' expectations of conventionality

Annette M. E. Henderson * *and Jessica C. Scott*

Early Learning Laboratory, School of Psychology, The University of Auckland, Auckland, New Zealand

Words are powerful communicative tools because of conventionality—their meanings are shared among same language users. Although evidence demonstrates that an understanding of conventionality is present early in life, this work has focused on infants being raised in English-speaking monolingual environments. As such, little is known about the role that experience in multilingual environments plays in the development of an understanding of conventionality. We addressed this gap with 13-month-old infants regularly exposed to more than one language. Infants were familiarized to two speakers who either spoke the same (English), or different (French vs. English) languages. Next, infants were habituated to a video in which one of the speakers provided a new word and selected one of two unfamiliar objects. Infants were then shown test events in which the other speaker provided the same label and selected either the same object or a different object. Our results demonstrate that exposure to at least one other language influences infants' expectations about conventionality. Unlike monolinguals, bilingual infants do not assume that word meanings are shared across speakers who use the same language. Interestingly, when shown speakers who use different languages, bilingual infants looked longer toward the test trials in which the second speaker labeled the object consistently with the first speaker. This finding suggests that exposure to multiple languages enhances infants' understanding that speakers who use different languages should not use the same word for the same object. This is the first known evidence that experience in multilingual environments influences infants' expectations surrounding the shared nature of word meanings. An increased sensitivity to the constraints of conventionality represents a fairly sophisticated understanding of language as a conventional system and may shape bilingual infants' language development in a number of important ways.

Keywords: conventionality, bilingualism, experience, infant cognition, visual habituation, infants

Introduction

Words are powerful communicative tools because their meanings are shared within a particular linguistic community (e.g., Clark, 1992, 1993, 2009). A basic understanding of this fact about

language emerges early in development (Henderson and Graham, 2005; Graham et al., 2006; Buresh and Woodward, 2007; Henderson and Woodward, 2012) and has been argued to play an important role in children's word learning (for reviews see Sabbagh and Henderson, 2007, 2013; Diesendruck and Markson, 2011). Although the existing evidence clearly suggests that an understanding of conventionality emerges early in life, the focus of these studies has been on infants being raised primarily in monolingual English-speaking environments. As such, very little is known about the role that experience in multilingual environments plays in the development of an understanding of conventionality. The present research begins to fill this gap by providing the first known investigation of the role that exposure to multilingual environments plays in infants' expectations surrounding the shared nature of words.

The conventional nature of language refers to the fact that "For certain meanings there is a form that speakers expect to be used in a language community" (Clark, 1992, p. 171). This shared nature of words ensures consistency in word meanings across language-users and thus, regulates communication within linguistic communities. Critically however, conventionality also has constraints; linguistic conventions are bounded by linguistic groups (Clark, 2007). To illustrate, although all English speakers would be expected to share knowledge that the word "shoes" refers to a category of items that are put on one's feet for protection, users of other languages are not bound by the same expectation. For example, French speakers would instead be expected to say "les chaussures" to refer to footwear. Thus, a key challenge facing language learners is to acquire the word meanings that are known and used within their linguistic community.

There is now converging evidence from a number of studies suggesting that children are sensitive to various aspects relevant to the conventional nature of language early in their lives. By 12 months of age, infants are aware that speech is the tool that people use to communicate (Martin et al., 2012; Vouloumanos et al., 2014; Pitts et al., 2015). By 16 months, infants are sensitive to the fact that objects have conventional names and expect other people to use the conventionally appropriate names for objects (i.e., call a "ball" a ball and not "a shoe"; Koenig and Echols, 2003). By their second birthday, infants focus on learning words, and not other symbolic behaviors, such as sounds or gestures, as appropriate names for objects (Namy and Waxman, 1998; Woodward and Hoyne, 1999; Graham and Kilbreath, 2007).

The most direct evidence of an understanding of the conventional nature of words comes from studies testing the age at which children understand that object labels are shared across people who use the same language (Woodward et al., 1994; Henderson and Graham, 2005; Graham et al., 2006; Buresh and Woodward, 2007; Henderson and Woodward, 2012). To test this question in infants, Buresh and Woodward (2007) developed a visual habituation paradigm in which infants were repeatedly shown an event in which a speaker either provided a novel label (i.e., "medo") or expressed positive affect (i.e., "ooh. Mmmmm.") while holding one of two novel objects. After habituating to that event, infants were shown test trials in which a speaker produced the same label while holding either the previously labeled target object (target trials) or a different object (distractor trials). The key

manipulation was whether the test speaker was the same speaker from habituation, or a different speaker who had been shown to use the same language as the habituation speaker. Buresh and Woodward's results revealed that, regardless of test speaker, 12-month-old infants looked longer toward the distractor test trials in the word conditions but not the positive affect condition (see Henderson and Woodward, 2012 for similar results with 9-month-olds). Thus, by 9 months, infants demonstrate an understanding of the shared nature of word meanings—they expect object labels, but not object preferences to be generalized across speakers.

However, word meanings are also tied to specific linguistic communities. Thus, a sophisticated understanding of conventionality requires understanding of the scope of its application; that word meanings are only shared by individuals from *the same* linguistic community. Au and Glusman (1990) demonstrated that preschool-aged children show some understanding of this concept in a study in which monolingual English speaking children are taught a novel label for an animal in two languages (i.e., English and Spanish). The children were then asked a question ("Can you guess which one Spanish speaking children would call a theri?"). The results showed that 3- to 6-year-olds would accept two labels for the same object, but only if the evidence was clear that the labels came from different languages. This finding suggests that preschool-aged children understand that labels are constrained by linguistic community insofar as speakers of different languages use different labels.

Scott and Henderson (2013) provided the first evidence of an understanding of the fact that linguistic community constrains conventionality in monolingual infants. In this study, 13-month-olds being raised in monolingual-English environments were familiarized to two speakers singing nursery rhymes in different languages (one actor sang in French and the other in English). Infants were then habituated to one of the speakers providing a new word (i.e., "A modi. A modi.") while holding one of two unfamiliar objects. After habituation, the second speaker provided the test events in which he/she uttered the same word as in habituation and picked up either the same or different object. Contrary to the results of the different speaker conditions reported in past research in which the speakers had been shown to use the same language (i.e., Buresh and Woodward, 2007; Henderson and Woodward, 2012), infants in Scott and Henderson's study did not look longer toward either test event. That is, infants did not generalize the word-referent link across two speakers who had been shown to speak different languages. These findings suggest that infants as young as 13 months of age have a fairly nuanced understanding of conventionality; they are sensitive to the fact that linguistic community constrains conventionality.

Taken together, the existing evidence provides a clear picture that infants understand several facets of the conventional nature of language. However, because this work has focused on infants being raised in monolingual environments, very little is known about the extent to which experience in multilingual environments influences the development of an understanding of conventionality. Given that as many as half the world's children grow up exposed to more than one language (Hoff, 2009), investigating the role that multilinguistic experience plays in the early development

of an understanding of conventionality represents a significant gap in the literature.

Similar to infants who are exposed to one language, bilingual infants are exposed to people (within a linguistic community) labeling objects consistently. However, unlike monolingual infants, bilingual infants also receive direct evidence that an object can have more than one label. For instance, while a bilingual infant's English-speaking father will always use the word "shoes" when placing shoes on his infant's feet, the infant's French-speaking mom will always say "les chaussures" while doing so. In addition to receiving direct evidence that objects can have two names, bilingual language learners are likely to have firsthand experience with the fact that linguistic community constrains conventionality. Infants may have already tried to use a word in one of their languages with a speaker from an entirely different linguistic community and thus, encountered a situation in which a word they know is not shared by another person. Such experiences might encourage bilingual infants to develop an early appreciation of the fact that not all people are likely to share many of the words they might be learning. Given that bilingual infants are provided with regular exposure to multiple labels for an object and may be more likely to have had experiences in which other language users do not understand the words that they use, it seems reasonable to expect that they may develop different expectations surrounding conventionality.

Experience with early bilingualism has been shown to influence young children's expectations about language in a number of ways. For example, early bilingualism has been shown to result in an increased awareness of the arbitrary nature of language. Evidence supporting this point comes from Eviatar and Ibrahim (2000) who explained an exchanging words game (e.g., "We'll call the sun the moon and the moon the sun") and then asked the children to answer a question (e.g., "When you go to sleep at night what do you see in the sky?"). The results of this study revealed that 4- to 7-year-old bilingual children were more likely to adhere to this new relationship than were monolingual children on this task suggesting a greater appreciation of the arbitrary nature of word-referent links. Similarly Bialystok (1988) showed that bilingual children perform better than their monolingual counterparts in tasks of metalinguistic awareness, or knowledge about language. However, only a handful of studies have examined whether early exposure to more than one language influences children's expectations about conventionality.

One such study was conducted by Diesendruck (2005) in which monolingual and bilingual 3-year-old children were taught a new word for one of two objects (e.g., "This is a Teega"). In a subsequent task, children were asked to select the object that was the referent of a second novel label (e.g., "Can you give me the patoo?") by a speaker who was absent when the original word was taught. Consistent with past research conducted by Diesendruck and Markson (2001), monolingual 3-year-olds assumed conventionality; they assumed that the second speaker was aware of the previously labeled object's name and when a different term was used, they inferred that the second novel label was used to refer to the unlabeled object. Interestingly, bilingual children did not select the unlabeled object at levels greater than chance. These findings suggest that bilingual children do

not assume that a speaker who was absent when an object had been labeled will know (and use) the same word to refer to the same object and thus, did not assume that the second novel label would refer to the unlabeled object. These results suggest that that bilingual preschoolers are cautious about making assumptions that other people will share knowledge of the linguistic terms that they know. Diesendruck concluded that bilingual children might believe that there are conventional ways to refer to objects, but do not assume that everybody knows them. Consistent with this possibility are the recent findings bilingual preschool-aged children (Kalashnikova et al., 2014) and toddlers (Byers-Heinlein et al., 2014) will accept a second label for a novel object from a second speaker if that speaker has been shown to use a different language. Together, these findings suggest that bilingual children do not assume conventionality and are sensitive to the fact that object labels do not generalize across linguistic groups.

Taken together, existing evidence suggests that early exposure to both consistent labeling within a linguistic community and divergent labeling from different language speakers influences bilingual children's assumptions about the use of labels. As noted above, bilingual children are cautious in making assumptions that speakers share linguistic terms (Diesendruck, 2005) and understand that object labels are not shared across speakers of different languages (Byers-Heinlein et al., 2014; Kalashnikova et al., 2014). Considering these findings and in light of increasing evidence of an understanding of conventionality in infancy, it is possible that bilingual infants may show similar tendencies. Some reason to suspect that bilingual infants might be attuned to the role that linguistic community plays in conventionality comes from evidence suggesting that bilingual infants are able to distinguish between different languages early in their lives. There is now a solid body of evidence demonstrating that, early in infancy, there are perceptual discrimination abilities, which assist infants in differentiating between languages (e.g., Bahrick and Pickens, 1988; Mehler et al., 1988; Moon et al., 1993; Bosch and Sebastián-Gallés, 1997). Some researchers have argued that these discriminative abilities allow bilingual infants to form separate representations for the languages they are acquiring (for a review see Werker and Byers-Heinlein, 2008). Regularly making this kind of distinction when receiving language input may result in bilingual infants being particularly sensitive to the presence of different languages. Supporting this sensitivity is evidence suggesting that bilingual toddlers adjust their language use based on the language most relevant to the present context, even when their communicative partner is an unfamiliar adult, which suggests a well-developed understanding of how and when to use their different languages (e.g., Genesee et al., 1996; Deuchar and Quay, 1999). These findings indicate that from early on, bilingual language learners show an awareness of the linguistic community of the speakers around them and raise the possibility that infants might be particularly sensitive to the fact that speakers of different languages should not use the same word meanings.

We investigate whether experience in a bilingual environment influences infants' understanding of conventionality in the present research by using a visual habituation paradigm. Thirteen-month-old infants who are being raised in bilingual environments

were familiarized to two speakers singing nursery rhymes in one of two conditions. Infants were exposed to two speakers singing nursery rhymes either in the same language (i.e., both speakers sang in English) or in a different language (i.e., one speaker sang in English and the other in French). Infants were then habituated to one of the speakers providing a novel label (i.e., "medo") while holding one of two novel objects. After habituation, infants were shown test trials in which the other speaker (from familiarization) produced the same label while holding either the previously labeled object (target trials) or a different object (distractor trials). If infants being raised in bilingual environments have the same expectations of conventionality as infants being raised in monolingual environments, we expected our findings to be consistent with previous research. Specifically, we expected that: (1) infants in the same language condition would look significantly longer toward the distractor test trials thereby demonstrating an expectation that word-referent links are shared across speakers who have been shown to use the same language and (2) infants in the different language condition would not look significantly longer toward the distractor test events thereby demonstrating an understanding that word-referent links are not shared by speakers who do not use the same language. To our knowledge, this is the first investigation of an understanding of the constraints of conventionality in bilingual children under the age of 3.

Materials and Methods

Participants

Thirty 13-month-old infants (M_{age} = 13 months, 5 days; SD = 0.39; range = 12;2–13;29; 16 males) being raised in multilingual environments were recruited from a large database of families who have volunteered to take part in studies on infant development managed by a cognitive development lab in an urban center in New Zealand. Parents reported their infant as being exposed to English between 40 and 65% of the time (M = 53.57%, SD = 7.6%). Thus, infants were exposed to at least one other language a minimum of 35% and a maximum of 65% of the time. The other languages to which infants were exposed were: German (n = 5), Dutch (n = 2), Samoan (n = 2), Portuguese (n = 3), Chinese (n = 3), Maori (n = 2), French (n = 2), Farsi (n = 2), Serbian (n = 1), Italian, (n = 1), Korean (n = 1), Turkish (n = 1), Spanish (n = 1), Afrikaans (n = 1), Hindi (n = 1), Polish (n = 1), and Japanese (n = 1)[1]. Parents also reported the ethnicities of their infant, which resulted in the following breakdown: New Zealand European (n = 6), Pacific Islander (n = 1), Asian (n = 3), Middle Eastern (n = 1), and other European (n = 5). Thirteen infants were reported as belonging to more than one ethnic group. One parent did not complete the demographic questionnaire.

Infants were randomly assigned to either the *same language condition* (n = 14, 8 males, 6 females) or the *different language condition* (n = 16, 8 males, 8 females). An additional seven

---

[1]The diversity in languages to which infants in this sample were exposed was a result of our participant recruitment approach. Many of these infants were called in to be included in the sample described in Scott and Henderson (2013). However, parents' responses on our demographic questionnaire revealed that these infants did not meet the monolingual criteria.

infants participated but were excluded from the final sample due to technical errors (n = 3) or because the infant received more than 65% of English exposure and thus did not meet the bilingual language criteria (n = 4).

Infants were given a small prize for their participation at the end of the study; parents were given a parking ticket and a $10 gift voucher for petrol or groceries.

Materials, Stimuli, and Procedure

After a warm-up play period during which infants were given time to become comfortable in the laboratory environment and the experimenter completed the informed consent procedures with the parent, infants and their parents were escorted to the experimental testing room.

Infants were seated on their parents lap approximately 168 cm from a projector screen on which the video stimuli would be shown. The presentation of the video stimuli was controlled by the experimenter who stood behind a curtain via a MacBook Pro laptop. The software Looking Time X (Hannigan, 2008) was used to present the video stimuli. Infants' gaze was recorded using a camera that was hidden underneath the projection screen, which was connected to a mixer that consolidated the video stimuli with the view of the infant from the Baby Camera into one video file. This video file was recorded using a HyperDeck Studio SSD recording device. All of the recording equipment was hidden behind a curtain out of infants' view. The live feed from the baby view camera was also transmitted via HDMI to a monitor in an adjacent room in which the coder, who was blind to condition and trial, sat and coded infants' attention.

Once the infant was seated on his/her parent's lap, the experimental session began. All infants participated in the following six phases (as per Scott and Henderson, 2013): language familiarization, habituation, baseline, test familiarization, and test.

Language Familiarization

During this 90-s phase infants were introduced to two speakers, a male and a female, who alternated singing nursery rhymes (see **Figure 1**). For infants in the *same language condition*, both speakers sang in English; the male speaker sang "Mary Had a Little Lamb" and "Itsy Bitsy Spider," the female speaker sang "Row, Row, Row Your Boat" and "Twinkle Twinkle Little Star." Consistent with Scott and Henderson (2013), infants in the *different language condition* were shown the male speaker singing in French (e.g., "Frere Jacques" and "Alouette") and the female speaker singing in English (e.g., "Row, Row, Row Your Boat" and "Twinkle Twinkle Little Star"). The female actor was a native English speaker. The male actor was both a native French and English speaker and thus, did not have a French-accent when singing the English nursery rhymes. Although each song differed slightly in duration, the total duration of time infants were exposed to each speaker was consistent within and across conditions. To ensure that there were no differences across conditions in infants' attention during this phase we ran a 2 (song: first, second) × 2 (speaker: male, female) × 2 (condition: same language, different language) mixed-design ANOVA on the percentage of time that infants looked toward the display for each song with song and speaker as within subject factors. This analysis revealed a

FIGURE 1 | Sample image of the language familiarization, habituation, test familiarization, and test trials used in this research.

significant main effect of speaker, $F(1,23) = 19.37$, $p < 0.001$, $\eta^2 = 0.46$; infants spent a significantly smaller percentage of time attending to the display during the male speaker's songs ($M = 92.4\%$, $SD = 1.52$) than they did the female speaker's songs ($M = 99.1\%$, $SD = 0.49$). Importantly, no other effects reached statistical significance confirming that there were no differences between conditions in the percentage of time that infants attended to either the songs and/or the speakers during this phase. An independent samples t-test further confirmed that the duration of time (seconds) that infants spent looking toward

the speakers during this phase was not significantly different across conditions ($M_{\text{same language}} = 77.0 = SE_{\text{same language}} = 2.04$; $M_{\text{different language}} = 73.33$, $SE_{\text{different language}} = 1.02$), $t(25) = 1.75$, $p > 0.05$, Cohen's $d = 0.66$.

Habituation

Infants in both conditions were shown the same habituation event in which the male speaker looked up from his lap, smiled, looked at one of the two objects on the table, provided a new word (i.e., "medo"), picked up the object, and said "medo" a second time

while looking at the object in his hand (see **Figure 1**). Infants were shown this video until the sum of their attention toward three consecutive trials was less than the sum of the first three habituation trials divided by two. (i.e., the habituation criterion), or until 14 habituation trials had elapsed.

Baseline

After habituation, infants were shown the habituation event one last time before entering the next phase.

Test familiarization

The purpose of this trial was to introduce infants to the set-up for the test trials. The second speaker (from the language familiarization phase) was seated at the table between the two objects from habituation. Consistent with other habituation studies (e.g., Woodward, 1998; Buresh and Woodward, 2007; Henderson and Woodward, 2012; Scott and Henderson, 2013), the side on which each object appeared was switched. During this trial, the speaker looked up smiled, looked at each object and then back toward the infant, lifted her arms up and shrugged (i.e., as if to say "which one?"). The non-verbal nature of this trial ensured that infants were not reminded of the language used by the second speaker ensuring that any condition differences were a result of the language information provided to infants during the language familiarization phase.

Test trials

Infants in both conditions were shown the same test trials in which the second speaker provided the word "medo" and grasped either the same object that the first speaker had grasped during habituation (i.e., target test trials) or the other object that was present during habituation but never grasped (i.e., distractor test trials). All infants were shown three test trials of each type in alternation.

After the last test trial, infants and their parents were escorted back to the family room. Parents were told about the hypotheses of the study and were given the opportunity to ask any questions. After answering their questions, the experimenter thanked parents for their time, gave the infant his/her prize and parents their voucher and parking ticket and then walked the family back to the carpark.

All phases of the experiment, with the exception of language familiarization, were infant controlled. Thus, the video paused after the actor had provided the second label and/or stopped moving and remained on the screen until the infant looked away for 2 s, or until 120 s had elapsed. The paused frame marked the onset of the calculation of infants' looking time. Infants' looking time and habituation criterion was calculated using the software jHab (Casstevens, 2007). The target object and type of first test trial were counterbalanced across conditions.

A second coder reliability coded all of the habituation and test trials. The original coder and the second coder agreed on 93% of the test trials. Importantly, the direction of disagreements was not systematic across the types of test trials (Fisher's Exact Test, $p = 0.56$, two tailed).

TABLE 1 | Mean looking times and standard errors for the habituation, baseline, and familiarization phases for each condition.

	Habituation			Test familiarization
	Sum first 3	Sum last 3	Baseline	
Condition				
Same language	49.90(4.91)	18.54(1.97)	7.21(1.50)	21.34(2.33)
Different language	55.08(8.14)	20.15(2.61)	13.32(2.00)	23.71(6.53)

Results

Preliminary independent samples t-tests revealed no significant differences between conditions in infants' age [$t(28) < 1$], percentage of time exposed to English [$t(28) = 1.34$, $p > 0.1$], or number of languages to which infants were exposed [$t(28) = 1.74$, $p = 0.08$]. Next, we investigated whether infants' attention during the habituation phase differed depending on the condition to which infants were assigned. **Table 1** shows infants' average looking time during habituation and toward the test familiarization trial. As expected, a 2 (habituation trial: sum first three trials, sum last three trials) × 2 (condition: same language, different language) mixed-design ANOVA revealed that infants looked significantly longer on the first three habituation trials ($M = 52.49$, SE = 4.96) than they did on the last three habituation trials ($M = 19.34$, SE = 1.67), $F(1,28) = 66.09$, $p < 0.001$, $\eta^2 = 0.70$. Critically, there were no differences between the conditions in infants' attention during the habituation phase.

Independent samples t-tests revealed no significant differences between conditions in the average number of habituation trials or the average duration of the test familiarization phase, t's(28) < 1. Surprisingly, infants in the different language condition ($M = 13.33$, SE = 2.00) looked significantly longer toward the baseline trial than did the infants in the same language condition ($M = 7.21$, SE = 1.51), $t(28) = 2.38$, $p = 0.02$, Cohen's $d = 0.87$. In light of this finding, the main analyses were conducted controlling for infants' attention toward the baseline condition.

Of key interest was whether infants' attention toward the target and distractor test trials differed depending on whether the speaker conducting the test trials had previously been shown to speak a language that was either the same as, or different from, the speaker who conducted the habituation trials. This question was examined by running a 2 (test trial type: target, distractor) × 2 (condition: same language, different language) × 2 (first test trial: target, distractor) mixed-design ANCOVA on infants' attention toward the test trials with test trial type as the within-subjects factor and attention toward baseline as the covariate[2]. The results revealed a statistically significant interaction between test trial type and condition, $F(1,25) = 4.31$, $p < 0.05$, $\eta^2 = 0.15$ and no other significant effects (see **Figure 2**).

A follow-up independent samples t-test revealed that infants in the same language condition did not look significantly longer toward either type of test trial, $t(13) = 1.37$, $p > 0.1$, suggesting

---

[2]These analyses were collapsed across gender, target object, and test pair as preliminary analyses revealed no significant interactions between these factors and condition or test trial type.
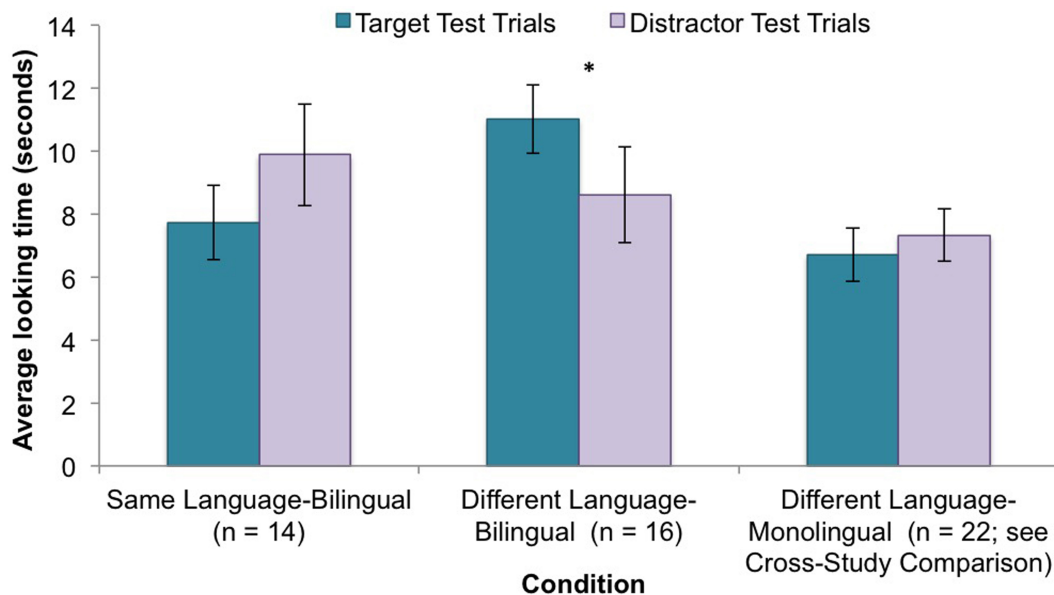
FIGURE 2 | Infants' average looking time (± SE) toward each type of test trial for each condition * = *p* < 0.05.

that infants did not expect two speakers who had previously been shown to use the same language to use the same word-object pairings. Thus, the infants in this study did not generalize the word-referent link across speakers from the same linguistic community. In contrast, infants in the different language condition looked significantly longer toward the target test trials than they did the distractor test trials, $t(15) = 2.20$, $p < 0.05$, Cohen's $d = 0.60$, $r = 0.63$, suggesting that infants found it unexpected when two speakers who had previously been shown to use different languages used the same word to refer to the same object, but not when the speakers used the same word to refer to a different object.

A sign test revealed that 79% of infants in the same language condition looked longer toward the *distractor* trials than they did the target test trials, $p = 0.057$. The fact that the *p*-value for this analysis is approaching significance contrasts with the results of the ANCOVA results reported above, which suggests that the infants in this study may have some, but not a robust, expectation that word-referent links are to be used consistently across two speakers who have been shown to use the same language. Conversely, 81% of the infants in the different language condition looked longer toward the *target* test trials than they did toward the distractor test trials ($p = 0.02$, sign test). This finding further suggests that bilingual infants do not generalize word-referent links across speakers who have been shown to speak different languages and, in fact, they find it surprising when word-referent links are used consistently across two speakers that had been shown to use different languages.

In sum, the results of the same language condition show that infants being raised in bilingual environments do not have a robust expectation that word-referent links are shared across speakers who have been shown to speak the same language. The results of the different language condition demonstrate that 13-month-old infants being raised in bilingual environments do

not expect two speakers who had previously been shown to speak different languages to use the same word to refer to the same object.

Comparison of Infants being Raised in Multilingual vs. Monolingual Contexts: Different Language Condition

To further examine the role that linguistic experience plays in infants' expectations surrounding the constraints of conventionality, we conducted a final set of analyses comparing the looking time data of the bilingual infants in this study to a group of monolingual infants ($n = 22$; 12 males, 10 females; mean age = 12 months, 27 days; range = 12 months, 9 days to 13 months, 21 days). Eighteen of the monolingual infants were from the data reported in Scott and Henderson (2013, Experiment 1)[3] and the remaining four were the infants who had been excluded from the final sample of this study because they did not meet the language criteria (i.e., infant had more than 65% of English exposure). As expected, the bilingual infants were exposed to significantly less English ($M = 51\%$, SE = 2.19) than were the monolingual infants ($M = 94\%$, SE = 2.34), $t(36) = 12.98$, $p < 0.001$, Cohen's $d = 4.38$. The bilingual infants were also exposed to a significantly greater number of languages ($M = 2.19$, SE = 0.10) than were the monolingual infants ($M = 1.55$, SE = 0.14), $t(36) = 3.40$, $p < 0.002$, Cohen's $d = 1.15$.

Preliminary analyses did not reveal any statistically significant differences between the language experience groups on infants' habituation and test familiarization trial looking times, or the average number of habituation trials. However, bilingual infants ($M = 13.33$, SE = 2.00) looked significantly longer toward the

---

[3]Scott and Henderson (2013) used the same videos as per the different language condition in the present study with infants who were primarily exposed to English (English exposure (> 80%).

baseline trial than did monolingual infants ($M = 5.90$, SE $= 0.83$), $t(36) = 3.79$, $p = 0.001$, Cohen's $d = 1.24$.

The main question of interest was whether infants' attention toward the target and distractor test trials in the different language condition differed depending on whether infants regularly received monolingual or bilingual linguistic experience. To investigate this question, we ran a 2 (test trial type: target, distractor) × 2 (linguistic experience: monolingual, bilingual) × 2 (first test trial: target, distractor) mixed-design ANOVA on infants' attention toward the test trials with test trial type as the within-subjects factor. The results revealed a significant two-way interaction between linguistic experience and test trial type, $F(1,34) = 6.38$, $p = 0.02$, $\eta^2 = 0.16$ (see **Figure 2**), and no other significant effects. Infants who are regularly exposed to more than one language looked significantly longer toward the target test trials than they did toward the distractor test trials, whereas infants who are only regularly exposed to one language did not look significantly longer toward either type of test trial.[4]

A chi-square analysis on the number of infants in each linguistic group who demonstrated longer looking toward (i.e., a preference for) the target test trials revealed that 81% of the infants being raised in multilingual environments, but only 41% of the infants being raised in monolingual environments, showed a looking time preference toward the target test trials, Pearson Chi-Square $= 6.18$, df $= 1$, $p = 0.01$. These results further confirm that experience influences infants' expectations regarding the extent to which linguistic community constrains conventionality.

Lastly, a Pearson's bivariate correlation revealed a significant negative correlation between the percentage of time infants are exposed to English and infants' total looking time toward the target test trials, Pearson correlation $= -0.34$, $p = 0.04$. Consistent with the above results, infants who are exposed to English a lower percentage of time showed greater looking toward the target test trials.

The above set of analyses examined the role that varying levels of exposure to another language (i.e., from 0 to 60% exposure to a language other than English) plays on 13-month-old infants' expectations of conventionality. The results demonstrate that linguistic experience influences infants' expectation surrounding the extent to which linguistic community constrains conventionality. While infants being raised in monolingual environments do not generalize new words across speakers who use different languages, infants being raised in multilingual environments find it particularly surprising when speakers of different languages use the same word for the same object. These findings suggest that infants being exposed to more than one language might have an enhanced understanding of the fact that word meanings are generally unique to individual languages.

Discussion

An understanding of the shared nature of word meanings emerges early in life (Woodward et al., 1994; Henderson and Graham, 2005;

Graham et al., 2006; Buresh and Woodward, 2007; Henderson and Woodward, 2012). However, most of the evidence supporting this point comes from monolingual infants. As such, little is known about how linguistic experience influences this development. Given the growing number of families throughout the world that are raising their infants in multilingual environments, it is essential to understand how such experiences influence language development. The present research addresses this gap by examining whether exposure to more than one language influences infants' developing expectations of conventionality. Specifically, we tested whether bilingual infants expect words to be shared across speakers who have been shown to use the same, or a different, language. If bilingual infants' understanding of conventionality is similar to monolingual infants, we expected that bilingual 13-month-olds would assume that word-object associations would be consistent across users of the same language, but not users of different languages. To the best of our knowledge, our findings provide the first evidence that experience in bilingual environments influences the expectations that are formed about the shared nature of word meanings within the first 13 months of our lives.

Firstly, and contrary to our expectations, our findings suggest that bilingual infants do not have a robust expectation that word-object links should be consistent across speakers of the same language. This comes from our finding that infants in the same language condition did not look reliably longer toward either type of test trial and thus, demonstrating that they do not expect two speakers who use the same language to provide the same word-object pairings. This result contrasts with the results of previous research in which monolingual infants in the same condition look longer toward distractor test trials thereby demonstrating an expectation that word-referent pairings should be consistent across speakers who have been shown to use the same language (Henderson and Graham, 2005; Graham et al., 2006; Buresh and Woodward, 2007; Henderson and Woodward, 2012). This finding also appears to contrast with the findings reported by Byers-Heinlein et al. (2014) who posited that bilingual 2-year-olds' tendency to avoid attaching a second label to an already labeled object revealed an assumption of conventionality. However, because the test trials in their study were completed by the same speaker and not a second speaker who spoke the same language as the first speaker, the extent to which the bilingual toddlers were performing in line with an assumption of conventionality remains unclear.

The finding that bilingual infants in the present study did not generalize word-object pairings across speakers of the same language suggests that exposure to more than one language encourages infants to be more conservative in assuming conventionality. Why might exposure to more than one language influence infants' expectations of conventionality regarding two people who use the same language? One possibility is that bilingual infants might assume that other language users are also exposed to multiple languages, even though infants are only shown the speakers speaking one language. As such, bilingual infants might be more hesitant to generalize word meanings across speakers. This possibility is consistent with previous research in which bilingual preschool-aged children were less likely to assume conventionality and

---

[4]An ANCOVA with attention toward baseline as the covariate, revealed a similar pattern however, the 2-way interaction between linguistic experience and test trial type was only approaching significance ($p = 0.09$).

more likely to accept second labels for the same object than were the monolingual children (e.g., Diesendruck, 2005; Kalashnikova et al., 2014). Further, Pitts et al. (2015) revealed that 20-month-old monolingual infants assumed that unfamiliar people would only understand one language, whereas bilingual infants did not. Pitts et al. (2015) argued that their findings suggest that bilingual infants are more open to the possibility that an unfamiliar person could understand more than one language than are their same-aged monolingual peers. Taken together, our findings and those reported in previous research suggest that early consistent exposure to more than one language influences the extent to which children will assume conventionality.

Our finding that bilingual infants do not have the same expectations of conventionality as their same-aged monolingual peers aligns with research demonstrating that bilingual and monolingual infants have different expectations about the possible meanings of new words. For example, by 18 months monolingual infants demonstrate a robust expectation that novel words map on to novel objects (Halberda, 2003, 2006; Markman et al., 2003; Byers-Heinlein and Werker, 2009; Xu et al., 2011), whereas infants of the same age who are regularly exposed to more than one language do not (Byers-Heinlein and Werker, 2009; Houston-Price et al., 2010). Together with the present findings, existing evidence suggests that exposure to more than one language affects infants' developing expectations about the meanings of words in several ways. In our future work we will seek to clarify the nature of the relationship between multilingual infants' developing expectations about word meanings and the speakers who share them.

The second key finding from the present study is that bilingual infants do not expect two speakers who had been shown to speak different languages to use the same word to refer to the same object. Consistent with the findings reported by Scott and Henderson (2013) our results further confirm that 13-month-old infants appreciate that linguistic community constrains conventionality. However, our findings extend this past work by demonstrating that infants who are regularly exposed to more than one language are particularly sensitive to the constraints that the language an individual speaks constrains conventionality. This conclusion is supported by the fact that bilingual infants in the present study looked significantly longer toward the target test trials suggesting that they were particularly surprised when the two speakers who had been shown to use different languages knew (and used) the same word. This pattern contrasts with the pattern demonstrated by the monolingual infants in Scott and Henderson's study who did not look reliably longer toward either type of test trial.

Longer looking toward the target test trials as a function of language experience was further confirmed in our final set of analyses directly comparing monolingual and bilingual infants. Infants who were exposed to at least one other language a greater percentage of time were more likely to look longer when users of different languages label objects consistently than were infants who were exposed only to English a greater percentage of time. The comparison of bilingual infants in the different language condition with a group of monolingual infants in the same condition provides converging evidence that exposure to more than one

language enhances infants' expectation that word meanings are tied to particular languages. Bilingual infants' enhanced sensitivity to the fact that users of different languages do not share word-referent links is consistent with the recent findings reported by Byers-Heinlein et al. (2014) who showed that bilingual 2-year-olds have an enhanced understanding of the nature of foreign language words, compared to their same aged-monolingual peers, in a mutual exclusivity paradigm.

The unexpected finding that the bilingual infants in the different language condition looked longer toward the baseline trial than did the bilingual infants in the same language condition and the monolingual infants in the different language condition warrants some attention. This finding suggests that providing bilingual infants with a context in which they had been shown two speakers using two different languages heightened their attention toward the labeling event, but only during the baseline trial. One possibility is that the bilingual infants in the different language condition needed extra time after habituating to make sure that the speaker was labeling the object consistently to ensure that they could learn the new label for the object[5]. This finding was surprising and the reason for this difference is unclear. However, it is important to note that none of the other pre-test trial measures revealed significant differences between conditions and perhaps most importantly, the key difference between the two bilingual conditions in infants' looking times toward the different test trials held after controlling for infants' attention toward the baseline trial.

These findings raise interesting questions about which aspects of bilingual infants' linguistic experience contribute to their enhanced understanding that word-referent links are not shared across different languages. One possibility is that bilingual infants' own communicative experiences, perhaps of producing words in one language to speakers of a different language and possibly being misunderstood, play a key role in shaping their reduced tendency assume conventionality. However, the fact that infants in the present study were only 13 months of age and thus, would not have had significant amounts of experience producing their own words, suggests that simply being exposed to more than one language on a regular basis might be sufficient to raise questions about conventionality in bilingual infants. If this were true, then younger infants being raised in bilingual environments might also have different expectations about conventionality than their same-aged monolingual peers who have been shown to expect words to be shared across speakers of the same language (Henderson and Woodward, 2012). Indeed, future work investigating the developmental trajectory of this understanding will shed important insights into the kinds of experiences that support infants' developing expectations about conventionality.

Another open question is whether the bilingual infants in this study would have performed differently if they were familiarized to two speakers who spoke both of the languages to which they are regularly exposed. In the present study all infants in the different language condition were exposed to an English speaker and a French speaker. As such most of the infants in the present research were only familiar with one of the languages used in

---

[5]We thank a reviewer for offering this suggestion.

the study (i.e., English). One interesting question is whether infants would have generalized the word-referent link if the other speaker had been shown to speak a language consistent with the other language to which infants were exposed. However, evidence from past research gives us reason to suspect that testing this possibility would not result in a different pattern of results. For example, in Scott and Henderson (2013), monolingual infants did not generalize word-referent links across users of different languages even when the English speaker completed the habituation phase. Thus, learning a new word-object pairing from a speaker from the same linguistic group as the infant did not result in the monolingual infants being more likely generalize the word-referent link across speakers who use different languages. Further, in their study on bilingual infants' expectations of the communicative nature of foreign languages, Pitts et al. (2015) directly tested whether infants' performance differed depending on their specific language-learning combinations. Their results revealed that bilingual infants' performance did not differ when the test languages used in the study were more, or less, similar to infants' own language-learning combination. Given these findings, we think it unlikely that our results would have been different if our participants had been familiar to both languages.

Our findings raise interesting questions about the extent to which bilingual infants' expectations about conventionality influences their subsequent word learning. One way in which an understanding of conventionality has been argued to help children's word learning is by enabling them to rapidly generalize words across both individuals and contexts (Sabbagh and Henderson, 2007, 2013). If bilingual infants do not expect words to generalize across individuals who use the same language, this might mean that they would require explicit information about a speaker's awareness of particular word-referent links, a requirement that might result in slower word learning relative to their monolingual same-aged peers. However, the strategy of not assuming conventionality might also be adaptive as it might help bilingual word learners to keep track of specific word-referent links and the language to which they belong. Such a possibility is related to a second way in which an understanding of conventionality might be important for word learning; it might help children focus on learning the new word-referent links that will be relevant (i.e., shared) within the their own linguistic community (Sabbagh and Henderson, 2007, 2013). Regarding this point, our finding that bilingual infants might be more attuned to the fact that people who use different languages should not produce the same word-referent links, might mean that they might be better able to identify words that are unlikely to be shared by other language users. In so doing, they might be better able to streamline their word learning toward learning the word meanings that are most likely to be relevant within one of their linguistic groups. The extent to which an understanding of conventionality influences children's word learning remains an open question. Future work examining the links between the development of an understanding of conventionality and its ties to word learning in children from both monolingual and bilingual backgrounds would help tease apart these potential consequences of multilingual exposure. Such work would determine whether conservatism when it comes to assuming conventionality is adaptive, or hinders subsequent language development.

In order for words to be effective communicative tools, their meanings must be known and used by all members within a given linguistic community. Adults readily appreciate this fact about language as we construct sentences using conventionally appropriate meanings and understand that communication will likely be difficult when speaking with people from other linguistic groups. There is now a substantial body of evidence suggesting that monolingual infants are sensitive to the conventional nature of language early in their lives. However, little is known about how diverse linguistic experiences influence infants' understanding of conventionality. The present research provides the first evidence that linguistic experience influences the assumptions that children develop about the conventional nature of word meanings within the first year of their lives. Exposure to more than one language encourages infants to be more restrictive in their assumptions regarding conventionality within users of the same language and enhances their understanding that different languages constrain conventionality. This increased sensitivity to the constraints of conventionality represents a fairly sophisticated understanding of language as a conventional system and may play a role in shaping bilingual infants' language development in a number of important ways.

Acknowledgments

References

Au, T. K., and Glusman, M. (1990). The principle of mutual exclusivity in word learning: to honor or not to honor? *Child Dev.* 61, 1474–1490. doi: 10.1111/j.1467-8624.1990.tb02876.x

Bahrick, L. E., and Pickens, J. N. (1988). Classification of bimodal English and Spanish language passages by infants. *Infant Behav. Dev.* 11, 277–296. doi: 10.1016/0163-6383(88)90014-8

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567. doi: 10.1037/0012-1649.24.4.560

Bosch, L., and Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition* 65, 33–69. doi: 10.1016/S0010-0277(97)00040-1

Buresh, J. S., and Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition* 104, 287–314. doi: 10.1016/j.cognition.2006.07.001

Byers-Heinlein, K., Chen, K. H., and Xu, F. (2014). Surmounting the tower of babel: monolingual and bilingual 2-year-olds' understanding of the nature of foreign language words. *J. Exp. Child Psychol.* 119, 87–100. doi: 10.1016/j.jecp.2013.09.011

Byers-Heinlein, K., and Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Dev. Sci.* 12, 815–823. doi: 10.1111/j.1467-7687.2009.00902.x

Casstevens, R. M. (2007). *jHab: Java Habituation Software (Version 1.0.2).* [Software]. Available at: http://woodwardlab.uchicago.edu/resources/

Clark, E. V. (1992). "Conventionality and contrast: pragmatic principles with lexical consequences," in *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*, eds A. Lehrer and E. F. Kittay (Mahwah: Lawrence Erlbaum Associates), 171–188.

Clark, E. V. (1993). *The Lexicon in Acquisition.* Cambridge: Cambridge University Press.

Clark, E. V. (2009). "What shapes children's language? Child-directed speech, conventionality, and the process of acquisition," in *Routes to Language: Studies in Honor of Melissa Bowerman*, ed. V. C. Mueller-Gathercole (New York: Psychology Press), 233–254.

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New Dir. Child Adolesc. Dev.* 115, 11–23. doi: 10.1002/cd.179

Deuchar, M., and Quay, S. (1999). Language choice in the earliest utterances: a case study with methodological implications. *J. Child Lang.* 26, 461–475. doi: 10.1017/S0305000999003852

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: an empirical examination. *Dev. Psychol.* 41, 451–463. doi: 10.1037/0012-1649.41.3.451

Diesendruck, G., and Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Dev. Psychol.* 37, 530–641. doi: 10.1037/0012-1649.37.5.630

Diesendruck, G., and Markson, L. (2011). Children's assumption of the conventionality of culture. *Child Dev. Perspect.* 5, 189–195. doi: 10.1111/j.1750-8606.2010.00156.x

Eviatar, Z., and Ibrahim, R. (2000). Bilingual is as bilingual does: metalinguistic abilities of arabic-speaking children. *Appl. Psycholinguist.* 21, 451–471. doi: 10.1017/S0142716400004021

Genesee, F., Boivin, I., and Nicoladis, E. (1996). Talking with strangers: a study of bilingual children's communicative competence. *Appl. Psycholinguist.* 17, 427–442. doi: 10.1017/S0142716400008183

Graham, S. A., and Kilbreath, C. S. (2007). It's a sign of the kind: gestures and words guide infants' inductive inferences. *Dev. Psychol.* 43, 1111–1123. doi: 10.1037/0012-1649.43.5.1111

Graham, S. A., Stock, H., and Henderson, A. M. E. (2006). Nineteen-month-olds' understanding ofthe conventionality of object labels versus desires. *Infancy* 9, 341–351. doi: 10.1207/s15327078in0903_5

Halberda, J. (2003). The development of a word-learning strategy. *Cognition* 87, B23–B34. doi: 10.1016/S0010-0277(02)00186-5

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cogn. Psychol.* 53, 310–344. doi: 10.1016/j.cogpsych.2006.04.003

Hannigan, T. (2008). *Looking TimeX Software (Version 2.6).* Kingston, ON, Canada.

Henderson, A. M. E., and Graham, S. A. (2005). Shared nature of novel object labels. *J. Cogn. Dev.* 6, 381–402. doi: 10.1207/s15327647jcd0603_4

Henderson, A. M. E., and Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Dev. Sci.* 15, 641–652. doi: 10.1111/j.1467-7687.2012.01157.x

Hoff, E. (2009). *Language Development* 4th Edn. Belmont: Wadsworth.

Houston-Price, C., Caloghiris, Z., and Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy* 15, 125–150. doi: 10.1111/j.1532-7078.2009.00009.x

Kalashnikova, M., Mattock, K., and Monaghan, P. (2014). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Biling. Lang. Cogn.* doi: 10.1017/S1366728914000364 [Epub ahead of print].

Koenig, M. A., and Echols, C. H. (2003). Infants' understanding of false labeling events: the referential roles of words and the speakers who use them. *Cognition* 87, 179–208. doi: 10.1016/S0010-0277(03)00002-7

Markman, E. M., Wasow, J. L., and Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognit. Psychol.* 47, 241–275. doi: 10.1016/S0010-0285(03)00034-3

Martin, A., Onishi, K. H., and Vouloumanos, A. (2012). Understanding the abstract role of speech in communication at 12 months. *Cognition* 123, 50–60. doi: 10.1016/j.cognition.2011.12.003

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178. doi: 10.1016/0010-0277(88)90035-2

Moon, C., Cooper, R. P., and Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behav. Dev.* 16, 495–500. doi: 10.1016/0163-6383(93)80007-U

Namy, L. L., and Waxman, S. R. (1998). Words and gestures: infants' interpretations of different forms of symbolic reference. *Child Dev* 69, 295–308. doi: 10.1111/j.1467-8624.1998.tb06189.x

Pitts, C. E., Onishi, K. H., and Vouloumanos, A. (2015). Who can communicate with whom? Language experience affects infants' evaluation of others as monolingual or multilingual. *Cognition* 134, 185–192. doi: 10.1016/j.cognition.2014.10.003

Sabbagh, M. A., and Henderson, A. M. E. (2007). How an appreciation of conventionality shapes early word learning. *New Dir. Child Adolesc. Dev.* 115, 25–37. doi: 10.1002/cd.180

Sabbagh, M. A., and Henderson, A. M. E. (2013). "Preschoolers are selective word learners," in *Navigating the Social World: What Infants, Children, and Other Species Can Teach us*, eds M. R. Banaji and S. A. Gelman (New York, NY: Oxford University Press), 177–180.

Scott, J. C., and Henderson, A. M. E. (2013). Language matters: thirteen-month-olds understand that the language a speaker uses constrains conventionality. *Dev. Psychol.* 49, 2102–2111. doi: 10.1037/a0031981

Vouloumanos, A., Martin, A., and Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Dev. Sci.* 17, 872–879. doi: 10.1111/desc.12170

Werker, J. F., and Byers-Heinlein, K. (2008). Bilingualism in infancy: first steps in perception and comprehension. *Trends Cogn. Sci.* 12, 144–151. doi: 10.1016/j.tics.2008.01.008

Woodward, A., and Hoyne, K. (1999). Infants' Learning about words and sounds in relation to objects. *Child Dev.* 70, 65–77. doi: 10.1111/1467-8624.00006

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34. doi: 10.1016/S0010-0277(98)00058-4

Woodward, A. L., Markman, E. M., and Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Dev. Psychol.* 30, 553–566. doi: 10.1037/0012-1649.30.4.553

Xu, F., Cote, M., and Baker, A. (2011). Labeling guides object individuation in 12-month-old infants. *Psychol. Sci.* 16, 372–377. doi: 10.1111/j.0956-7976.2005.01543.x

frontiers
in Psychology

Parents' empathic perspective taking and altruistic behavior predicts infants' arousal to others' emotions

Michaela B. Upshaw[1]*, Cheryl R. Kaiser[2] and Jessica A. Sommerville[1]

[1] Early Childhood Cognition Lab, Department of Psychology, Center for Child and Family Well-being, University of Washington, Seattle, WA, USA, [2] Social Identity Lab, Department of Psychology, University of Washington, Seattle, WA, USA

Empathy emerges in children's overt behavior around the middle of the second year of life. Younger infants, however, exhibit arousal in response to others' emotional displays, which is considered to be a precursor to fully developed empathy. The goal of the present study was to investigate individual variability in infants' arousal toward others' emotional displays, as indexed by 12- and 15-month-old infants' ($n = 49$) pupillary changes in response to another infant's emotions, and to determine whether such variability is linked to parental empathy and prosociality, as indexed via self-report questionnaires. We found that increases in infants' pupil dilation in response to others' emotional displays were associated with aspects of parental empathy and prosociality. Specifically, infants who exhibited the greatest arousal in response to others' emotions had parents who scored highly on empathic perspective taking and self-reported altruism. These relations may have been found because arousal toward others' emotions shares certain characteristics with empathic and prosocial dispositions. Together, these results demonstrate the presence of early variability in a precursor to mature empathic responding in infancy, which is meaningfully linked to parents' empathic dispositions and prosocial behaviors.

Keywords: infancy, pupil dilation, empathy, arousal, parental dispositions

Introduction

Empathy, or the understanding and experiencing of another's affective or psychological state, is integral to fostering positive social interactions and healthy interpersonal relationships. Indeed, empathy is considered to be vital to the emergence of prosocial behaviors (Hoffman, 1982; de Waal, 2008; Knafo et al., 2008). This theoretical claim is bolstered by empirical work demonstrating that higher levels of empathy in both children and adults are associated with an increased likelihood of offering help to a stranger (Dovidio et al., 1990), donating money to charity (Miller, 1979; Davis, 1983), and an increased willingness to encounter and help needy individuals (e.g., volunteering at a shelter; Davis et al., 1999; Davis, 2005; see also Eisenberg and Miller, 1987; Batson, 1991, 2002 for reviews). A critical question, then, concerns the developmental origins of empathy, as well as when in ontogeny individual differences emerge in empathic responses. The goal of this paper is to investigate an early precursor to empathy, particularly infants' arousal in response to others' emotions, and to examine whether variability in such arousal is associated with parental dispositions, such as empathy, and theoretically aligned characteristics, such as prosocial behavior.

In its mature form, empathy is considered to have both affective and cognitive components (Davis, 1983; Zahn-Waxler and Radke-Yarrow, 1990; Eisenberg, 2000; Preston and de Waal, 2002; Knafo et al., 2008). The cognitive component of empathy involves apprehending or understanding another person's experience and differentiating that from one's own (i.e., putting oneself in another person's 'shoes'; Davis, 1983; Zahn-Waxler and Radke-Yarrow, 1990). The affective component of empathy involves one's emotional response toward another person's experience (i.e., feelings of warmth, compassion, and concern toward others; Davis, 1983; Batson et al., 1987; Zahn-Waxler and Radke-Yarrow, 1990; Eisenberg, 2000). Previous work has found evidence for both cognitive and affective components of empathy by 18 months of age: for example, toddlers attempt to actively comfort an upset experimenter and actively seek information regarding the source of the experimenter's distress (Zahn-Waxler et al., 1979, 1992b). Importantly, past work also reveals that there is variability in young children's empathic responses by the second year of life (Zahn-Waxler et al., 1979, 1992b; Eisenberg et al., 1989). This individual variability in early empathic responding appears to be stable across contexts (Young et al., 1999; Robinson et al., 2001; Spinrad and Stifter, 2006; Moreno et al., 2008) and time (Zahn-Waxler et al., 1992a,b, 2001; Volbrecht et al., 2007). Nevertheless, because many previous paradigms (e.g., Zahn-Waxler et al., 1992a,b) have relied on children's overt verbal and behavioral responses to another person's distress, existing work may overestimate the age at which children first begin to demonstrate variability in empathic responses. This raises the possibility that individual differences in precursors to empathy may be present even earlier in development.

Indeed, well before overt signs of empathy emerge in development, infants exhibit arousal in response to others' emotional expressions. For example, newborns cry in response to hearing another infant's cry, but not in response to their own cry or to sounds with matched synthetic frequencies (Sagi and Hoffman, 1976; Martin and Clark, 1982; Dondi et al., 1999). Between 2 and 3 months of age, infants change their affective state in response to their mother's emotional expressions (e.g., angry facial expressions and freezing in response to their mother's expression of anger; Haviland and Lelwica, 1987), and by about 9 months of age, infants tailor their behavior to correspond with their mother's emotions, such as playing and smiling more frequently when their mother exhibits joy as opposed to sadness (Termine and Izard, 1988). These behaviors are considered to be precursors to more mature empathic responding because, in order to respond empathically, one must first register and be moved by another's emotional expression (Hoffman, 1982, 2000, 2008). Critically, infants' ability to register others' emotions has been linked to the later emergence of empathic behaviors (Roth-Hanania et al., 2011). For example, infants' facial and vocal signs of concern in response to their mother's distress at 10 months is positively associated with their attempts to help and comfort their mother during simulated expressions of pain at 12–16 months of age. Altogether, this work motivates a closer investigation into precursors to later empathic responses.

In order to track the development of empathy and related precursors, researchers are increasingly turning to physiological measures. One particularly promising physiological measure is pupil dilation, or changes in pupil size that occur during stimulus processing. Pupil dilation reflects activity of the sympathetic nervous system and is thought to index increased attention and arousal in response to the observed stimuli (Beatty and Lucero-Wagoner, 2000; Porter et al., 2007; Laeng et al., 2012). Of central importance to the present study, pupil dilation has been used to measure arousal in response to others' emotions (Partala and Surakka, 2003; Bradley et al., 2008; Geangu et al., 2011b). For example, infants and adults exhibit greater pupil dilation during the processing of emotional stimuli (i.e., stimuli with a positive or negative valence) relative to the processing of neutral stimuli (Partala and Surakka, 2003; Bradley et al., 2008). In addition, this work has shown that the processing of negative emotional stimuli leads to greater pupil dilation than the processing of positive emotional stimuli (e.g., Geangu et al., 2011b). Thus, changes in pupil size, and in particular, pupil dilation, reflect one's degree of arousal in response to others' emotional displays. Pupil dilation is ideal for investigating arousal in response to others' emotions in infancy, as pupil dilation is an automatically elicited, non-verbal response, that does not require infants to produce complex, overt behavior, nor possess a sophisticated understanding of the observed situation (Laeng et al., 2012). In addition, pupil dilation in response to others' emotions is variable across individuals, suggesting that this measure is well suited to capture individual differences or variability in arousal toward others' emotions (Vanderhasselt et al., 2014; see also Bitsios et al., 2004).

Past research has investigated factors that contribute to the development of empathy in childhood, most notably investigating the impact of parental behaviors. Perhaps unsurprisingly, this work has largely focused on parental behaviors in the context of parent–child interactions, and has generally found that parents who exhibit sensitivity and concern in response to their children's distress are more likely to have children who exhibit greater empathic responding toward others (Zhou et al., 2002; Strayer and Roberts, 2004; Taylor et al., 2013; Newton et al., 2014). In contrast, the impact of parental dispositions, broadly construed, such as the parents' personality and their tendencies to engage in certain behaviors, on children's empathy has been less well studied. Nevertheless, existing research has demonstrated relations between parents' dispositional empathy and children's empathy. For example, parents who report feeling more empathic concern toward others in their everyday lives (e.g., endorsing statements such as, "When I see someone being taken advantage of, I feel kind of protective toward them") have children who exhibit more empathic behaviors toward others in need (e.g., child tries to comfort or reassure another in distress; Eisenberg et al., 1991; Volling et al., 2008). In addition, mothers who score high on measures of dispositional empathy and low on measures of personal distress in response to others' misfortunes have children who exhibit more empathic concern toward needy others as well as an enhanced capacity to adopt others' perspectives (Davidov and Grusec, 2006). Thus, how parents think, feel, and act toward other people influences the development

of children's empathy, even when such thoughts, feelings, and actions occur outside of the parent–child relationship. This association between parents and children could be due to genetic similarities, or to parental dispositions influencing their everyday behavior, of which children are frequently exposed to. Regardless, the relation between parents' empathic dispositions and related behaviors and children's developing empathy warrants closer investigation.

In the present study, we used pupil dilation to index infants' arousal in response to others' emotions and investigated whether individual variability in infants' pupillary changes in response to others' emotional displays was related to parents' empathic dispositions and prosocial tendencies. Prior work investigating infants' pupillary responses toward other infants' emotions has demonstrated group level increases in infants' pupil diameter in response to happy and sad emotional expressions relative to pupil diameter during neutral emotional expressions (Geangu et al., 2011b). Thus, in the current study, infants between 12- and 15-months of age watched videos of other infants expressing happiness and sadness, as well as neutral emotionality, while changes in their pupil diameter were recorded using an eye-tracker. In order to assess parents' empathic dispositions and prosocial tendencies, infants' primary caregiver completed two, widely used questionnaires that measure self-reported dispositional empathy and prosociality: the Interpersonal Reactivity Index (IRI; Davis, 1983) and the Prosocial Personality Battery (PSB; Penner et al., 1995), respectively. We predicted that parents who report greater levels of dispositional empathy, and who report a higher frequency of performing helpful behaviors toward others, would have infants who exhibit greater arousal, as assessed via changes in pupil diameter, during observation of another infant's emotional displays.

Materials and Methods

Participants
The final sample included 22 ($n = 13$ female), 12-month-olds ($M = 12$ months and 5 days; range: 11 months and 23 days to 12 months and 16 days) and 27 ($n = 14$ female), 15-month-olds ($M = 15$ months and 12 days; range: 14 months and 25 days to 16 months and 10 days), who were recruited from a database maintained by a large university in the Pacific Northwest of the United States. Thirteen additional infants participated but were excluded from analysis due to insufficient data stemming from technical problems with the eye-tracking camera and software ($n = 8$) or because of fussiness ($n = 5$). Of these infants, 41 were Caucasian, one was Native American, one was Hispanic, nine were of mixed ethnic backgrounds, and two parents chose not to disclose this information. Before participating, all parents provided informed consent for their infants and themselves to participate in the study.

Stimuli
The video stimuli were adapted from Geangu et al. (2011b; see article for full details and description of the original stimuli) and were presented using Presentation® software (Version

0.70, www.neurobs.com). In the neutral video, a male infant displayed neutral facial expressions and produced neutral babbling vocalizations (without emotional prosody). In the happy video, a different male infant displayed happy facial expressions and produced laughing vocalizations. In the sad video, a third male infant displayed facial expressions of sadness and frustration and produced strong crying vocalizations. Each video was 25 s in length (reduced from 50 s, as prior work found that infants' attention wandered during the second half of the video; Geangu et al., 2011b). In shortening the video length, care was taken to select segments of the original video that contained the least amount of infant movement, in order to reduce luminance differences. As an additional control for luminance differences, the videos were presented in black and white. Lastly, the videos were cropped to reduce the amount of background imagery and to enhance focus on the infants' emotional expressions. After these adaptations, we extracted the spatial average of the RGB values for each frame of each video, and calculated the weighted sum of the RGB values to estimate photometric luminance for each video (i.e., luminance = (0.2126 * R) + (0.7152 * G) + (0.0722 * B); see Jackson and Sirois, 2009). This analysis confirmed that the videos did not differ in photometric luminance: 8.65 = neutral, 8.66 = happy, and 8.36 = sad (all comparisons *ns*).

Infants were also shown a 10 s baseline video which consisted of a red and white rattle moving back-and-forth against a black background accompanied by soft music. The baseline video served to break up and transition infants' attention between the emotional videos. In addition, the baseline video provided a baseline assessment of infants' pupil size, which was used to perform baseline corrections prior to data analysis. We used the same baseline video as in Geangu et al. (2011b) in order to aid comparability between the two studies.

Apparatus
Infants' pupil diameter was measured using an Applied Science Laboratories (ASL) Eye-Trac 6 Control Unit and Desktop Optics D6 camera; accuracy 0.5°, resolution 0.26°, and collected at a frequency of 60 Hz. The eye-tracking camera was positioned beneath the stimulus displaying monitor (measuring 68.6 cm diagonally), and both the camera and monitor were placed directly in front of a plain beige wall. Dark curtains surrounded the stimulus displaying monitor and the infant, in order to focus infants' attention, and no other stimuli were present that could distract infants' attention. The lighting in the experimental room was held constant across participants, in order to prevent pupil size changes as a function of ambient lighting differences.

Procedure
All study procedures were approved by the university's Internal Review Board before the research was conducted. Infants sat in a car seat, approximately 76.2 cm from the stimulus displaying monitor, and the infant's parent sat behind them, out of the infant's sight. After the infant was in position, and before pupil data were recorded, a five-point calibration was performed (see Gredebäck et al., 2010). For approximately half of participants ($n = 22$), the procedure began with the presentation of the neutral

video, followed by the happy video. For the other half of participants ($n = 27$), the procedure began with the presentation of the happy video, followed by the neutral video. Infants were always shown the sad video last in the series of three videos, as prior work has found that negative stimuli can have carryover effects (Geangu et al., 2011a,b). Before each emotional video, infants were shown the 10 s baseline video.

Data Processing

Infants' pupil diameter was filtered off-line using Matlab (version 7.11 0.584, R2010b, Natick, MA, USA). A 20-point moving average window was applied to the data (i.e., pupil diameter at each time point was calculated as the average diameter of the surrounding 20 data points) in order to remove sudden brief increases and decreases in pupil diameter that normally occur and are considered to be artifacts (Beatty and Lucero-Wagoner, 2000; Geangu et al., 2011b). For analysis purposes, pupil diameter was calculated as the average pupil diameter during the last 23 s of each video. Infants' pupil diameter during the first 2 s of each video was excluded from analysis because of pupillary reflexes related to the baseline video to stimulus transition. Before data analysis, infants' pupil diameter was baseline-corrected by subtracting the average pupil diameter during the last second of the preceding baseline video from the average pupil diameter during (the last 23 s of) each emotional video. Using a baseline-corrected measure of pupil diameter controls for differences in tonic pupil size, as well as circumvents 'drift' in pupillary size during the task, which can occur due to the emotional nature of the stimuli (Hess and Polt, 1960; Geangu et al., 2011b).

Parental Questionnaire Measures

Before participating in the study, infants' primary caregivers completed the IRI (Davis, 1983) and the PSB (Penner et al., 1995). Among the 47 primary caregivers who completed the questionnaires, $n = 43$ were mothers of the infant participants, and $n = 4$ were fathers of the infant participants. Scores on the IRI were exclusively used to investigate relations with parents' dispositional empathy, as the IRI is the most widely used assessment of dispositional empathy (e.g., Beven et al., 2004; Pulos et al., 2004; Koller and Lamm, 2014), and because there is considerable overlap between items assessing dispositional empathy on the PSB and the IRI. Scores on the PSB were used to assess parents' prosocial behavioral tendencies.

The IRI is composed of four subscales (seven items for each subscale) that assess cognitive and affective components of dispositional empathy. Two subscales that assess affective aspects of empathy (empathic concern and personal distress) and one subscale that assesses the cognitive aspect of empathy (perspective taking) were analyzed for the present study. Empathic concern measures the tendency for one to experience other-oriented feelings of empathy and compassion for less fortunate individuals (e.g., "I often have tender, concerned feelings for people less fortunate than me"). In contrast to empathic concern, personal distress measures the tendency to experience *self*-oriented feelings of distress during others' misfortunes (e.g., "When I see someone who badly needs help in an emergency, I go to pieces"). Thus, personal

distress is thought to impede one's ability to behave empathically during others' misfortunes, as one must first overcome self-oriented feelings of anxiety and discomfort. On the cognitive side of empathy, the perspective taking subscale measures the tendency for one to adopt the point-of-view of another person when appraising a social situation (e.g., "I try to look at everybody's side of a disagreement before I make a decision"). Items on the IRI were rated on a five-point scale ranging from 1 (*does not describe me well*) to 5 (*describes me very well*). Internal consistency (Cronbach's alpha) was 0.75 for the empathic concern subscale, 0.82 for the perspective taking subscale, and 0.79 for the personal distress subscale.

The PSB is a self-report measure of dispositional empathy and the frequency of performing prosocial and helpful behaviors toward others. Scores on subscales of the PSB are combined in order to represent two higher-order factors: Other-Oriented Empathy, which measures the tendency to feel concern, pity, or sorrow in response to others' distress (e.g., "I am often quite touched by things that I see happen.") and Helpfulness, which measures the tendency for one to engage in helpful and altruistic behaviors toward others (e.g., "I have helped carry a stranger's belongings."). The higher-order factor of Helpfulness is constructed from scores on two subscales: self-reported altruism (composed of five items) and personal distress (composed of three items). However, because the items used to assess personal distress on the PSB are a verbatim subset of the items used to assess personal distress on the IRI, only parental scores on the self-reported altruism subscale of the PSB were used in the present analysis. Items on the self-reported altruism subscale of the PSB were rated on a five-point scale ranging from 1 (*never*) to 5 (*very often*). Internal consistency (Cronbach's alpha) was 0.74 for the self-reported altruism subscale.

Results

Infants' Visual Attention to the Emotional Videos

Our first course of action was to ensure that infants visually attended to the emotional videos. Accordingly, we calculated the percentage of time infants spent looking toward (versus away from) each emotional video relative to the total video duration. We conducted a 3×2 ANOVA with emotion (happy, neutral, sad) as a within-subjects factor and infants' age (12 or 15 months) as a between-subjects factor on the percentage of time infants spent looking toward the emotional videos. We found a significant main effect of emotion, $F(2,94) = 5.88$, $p = 0.004$, $\eta_p^2 = 0.11$, and no other main effects or interactions. Accordingly, we collapsed across infant age in order to further explore the main effect of emotion. Paired samples t-tests reveal that infants spent significantly more time looking toward the sad video ($M = 76.6\%$ of the total video duration, $SE = 3.4\%$) relative to the happy ($M = 67.2\%$, $SE = 3.3\%$), $t(48) = 2.32$, $p = 0.03$, $d = 0.50$, and neutral videos, ($M = 64.2\%$, $SE = 3.6\%$), $t(48) = 3.25$, $p = 0.002$, $d = 0.40$. There was no difference in the amount of time infants

looked toward the happy and neutral videos, $t(48) = -0.99$, $p = 0.33$.

Infants' Pupil Diameter in Response to the Emotional Videos

Our next course of action was to ensure that the videos elicited infants' arousal in response to others' emotional displays, in order to validate our paradigm and to replicate prior work (Geangu et al., 2011b). Thus, we conducted a 3 × 2 ANOVA, with emotion (happy, neutral, sad) as a within-subjects factor, and infants' age (12 or 15 months) and stimulus order (happy video first or neutral video first) as between-subjects factors, on infants' pupil diameter during observation of the emotional videos. We found a main effect of emotion, $F(2,90) = 8.20$, $p = 0.001$, $\eta_{p}^{2} = 0.14$, and no other main effects or interactions. Accordingly, we collapsed across infant age and stimulus order in subsequent analyses on infants' pupil dilation during the emotional videos. Planned, paired t-tests to investigate the main effect of emotion confirm that greater pupil dilation was found during observation of the sad video ($M = 0.31$ mm, $SE = 0.05$) relative to the neutral video ($M = 0.08$ mm, $SE = 0.04$), $t(48) = 3.99$, $p < 0.01$, $d = 0.48$, and relative to the happy video ($M = 0.17$ mm, $SE = 0.05$), $t(48) = 2.60$, $p = 0.01$, $d = 0.37$. In addition, pupil dilation during the happy video was marginally greater than pupil dilation during the neutral video, $t(48) = 1.77$, $p = 0.08$, $d = 0.25$ (see **Figure 1**). Importantly, the amount of time that infants spent looking toward the emotional videos was unrelated to their degree of pupil dilation in response to the videos, as assessed by Pearson's correlations between infants' percentage of looking toward each video and their pupil diameter in response to it:

happy, $r(49) = -0.23$, $p = 0.12$; neutral, $r(49) = -0.17$, $p = 0.26$; sad, $r(49) = -0.04$, $p = 0.76$.

Parental Questionnaire Measures: Descriptive Statistics

As a result of skipped questionnaire items and/or un-readable questionnaire responses, scores are missing from: three parents for the empathic concern subscale of the IRI and the self-reported altruism subscale of the PSB, two parents for the perspective taking subscale of the IRI, and four parents for the personal distress subscale of the IRI. Two parents did not complete any of the questionnaire measures. Means, SEs, and ranges for each of the subscales are presented in **Table 1**.

As shown in **Table 1**, scores on the perspective taking subscale were significantly associated with scores on the empathic concern subscale, $r(46) = 0.49$, $p < 0.001$, as well as significantly negatively associated with scores on the personal distress subscale, $r(47) = -0.32$, $p = 0.03$. However, scores on the empathic concern subscale were unrelated to scores on personal distress, $r(46) = -0.14$, $p = 0.37$. Parental scores on the self-reported altruism subscale of the PSB were unrelated to scores on the empathic concern subscale of the IRI, $r(45) = 0.01$, $p = 0.94$, the perspective taking subscale of the IRI, $r(46) = 0.13$, $p = 0.38$, and the personal distress subscale of the IRI, $r(46) = -0.19$, $p = 0.20$.

Relations Between Infants' Pupil Dilation in Response to Others' Emotions and Parental Questionnaire Measures

In order to capture changes in infants' pupil dilation during the sad and happy emotional videos relative to their pupil dilation
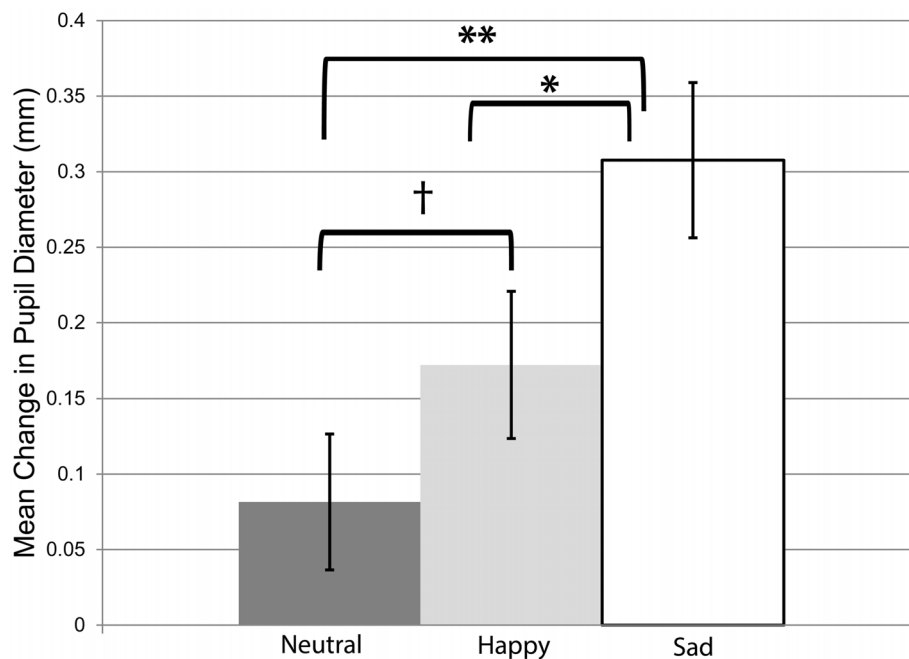

FIGURE 1 | Mean change in infants' pupil diameter during observation of the emotional videos (relative to pupil diameter during observation of the baseline video; in mm). Error bars represent SE. $^{†}p = 0.08$, $^{*}p = 0.01$, $^{**}p < 0.01$.

TABLE 1 | Pearson's correlations, means, SEs, and ranges for the main study variables.

Variable	1	2	3	4	5	6	7
1. Perspective taking (IRI)	—						
2. Empathic concern (IRI)	0.49**	—					
3. Personal distress (IRI)	−0.32*	−0.14	—				
4. Self-reported altruism (PSB)	0.13	0.01	−0.19	—			
5. Composite pupil dilation difference scores[a]	0.34*	−0.02	−0.26[†]	0.32*	—		
6. Happy pupil dilation difference scores[a]	0.28[†]	−0.01	−0.24	0.37*	0.86**	—	
7. Sad pupil dilation difference scores[a]	0.31*	−0.02	−0.21	0.20	0.89**	0.54**	—
Mean	18.96	20.70	9.62	15.24	0.32	0.09	0.23
SE	0.64	0.64	0.75	0.54	0.09	0.05	0.06
Possible range	0–28	0–28	0–28	5–25	–	–	–
Actual range	7–27	6–28	1–21	8–25	−1.01–2.09	−0.61–1.13	−0.56–1.35

[a]*Pearson's correlations are controlling for infants' age.*
***p < 0.001, *p < 0.05, [†]p = 0.08.*

during the neutral video, we created two difference scores by subtracting infants' pupil diameter during the neutral video from their pupil diameter during the sad and happy videos (hereafter referred to as "sad pupil dilation difference scores" and "happy pupil dilation difference scores"). These scores were designed to capture the difference in infants' pupil dilation during the sad (or happy) videos relative to their pupil dilation during the neutral videos. To capitalize on changes in infants' pupil dilation in response to *both* of the happy and sad emotional displays, we computed a composite pupil dilation difference score by summing the sad and happy pupil dilation difference scores together. In the following analyses, we first conducted Pearson's correlations between infants' composite pupil dilation difference scores (controlling for infants' age) and parents' scores on the questionnaire measures (empathic concern, perspective taking, and personal distress subscales of the IRI, and the self-reported altruism subscale of the PSB). If this correlation was significant, we then conducted separate correlations between infants' happy and sad pupil dilation difference scores and the parental variable of interest.

As shown in **Table 1**, parental scores on the perspective taking subscale were significantly associated with infants' composite pupil dilation difference scores, $r(44) = 0.34$, $p = 0.02$, 95% CI [0.05, 0.58], such that higher parental perspective taking predicted greater pupil dilation during the happy and sad videos relative to the neutral video. In addition, parental scores on the personal distress subscale were marginally negatively associated with infants' composite pupil dilation difference scores, $r(44) = -0.26$, $p = 0.08$, indicating that parents who report experiencing less feelings of self-oriented distress during others' misfortunes have infants who exhibit more pupil dilation in response to others' emotional displays. However, no relation was found between parents' scores on the empathic concern subscale and infants' composite pupil dilation difference scores, $r(43) = -0.02$, $p = 0.92$. Lastly, we found a significant relation between parental scores on the self-reported altruism subscale of the PSB and infants' composite pupil dilation difference scores, $r(43) = 0.32$, $p = 0.03$, 95% CI [0.02, 0.57], suggesting that parents who report performing more frequent altruistic behavior have

infants who exhibit greater pupil dilation in response to others' emotions.

In order to further examine the significant relations between parental scores on the perspective taking and self-reported altruism subscales and infants' composite pupil dilation difference scores, we conducted correlations between parental scores on the perspective taking and self-reported altruism subscales and infants' sad and happy pupil dilation difference scores. We found that parental perspective taking was significantly associated with infants' sad pupil dilation difference scores, $r(44) = 0.31$, $p = 0.04$, 95% CI [0.02, 0.56], and marginally associated with infants' happy pupil dilation difference scores, $r(44) = 0.28$, $p = 0.06$, 95% CI [−0.02, 0.53] (see **Figure 2**). In addition, we found that the relation between parents' self-reported altruism and infants' composite pupil dilation difference scores was primarily driven by infants' happy pupil dilation difference scores, $r(43) = 0.37$, $p = 0.01$, 95% CI [0.08, 0.60] (see **Figure 3**), as the relation between parents' self-reported altruism and infants' sad pupil dilation difference scores was non-significant, $r(43) = 0.20$, $p = 0.18$.

Discussion

The present study investigated whether variability in a precursor to mature empathy, namely infants' arousal in response to others' emotions, as indexed by pupillary changes to others' emotional displays, is related to differences in parents' empathic and prosocial dispositions. As an initial step toward this aim, we sought to ensure that our task was a reliable and accurate measure of infants' arousal, by replicating results of previous studies that have found greater arousal toward emotional stimuli (i.e., stimuli with a positive or negative valence) relative to neutral stimuli (Geangu et al., 2011b; see also Partala and Surakka, 2003; Bradley et al., 2008). Consistent with past research, we found that observation of the sad and happy emotions elicited infants' arousal significantly more than the neutral emotion; moreover, observation of the sad emotion elicited significantly more arousal than the happy emotion.

FIGURE 2 | Pearson's correlations (controlling for infants' age) between parental perspective taking scores (IRI) and infants' sad and happy pupil dilation difference scores (in mm).

FIGURE 3 | Pearson's correlations (controlling for infants' age) between parental self-reported altruism scores (PSB) and infants' happy pupil dilation difference scores (in mm).
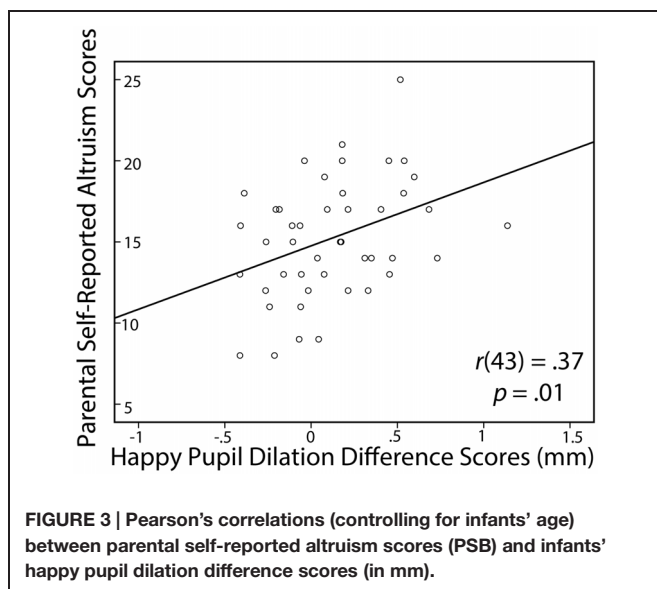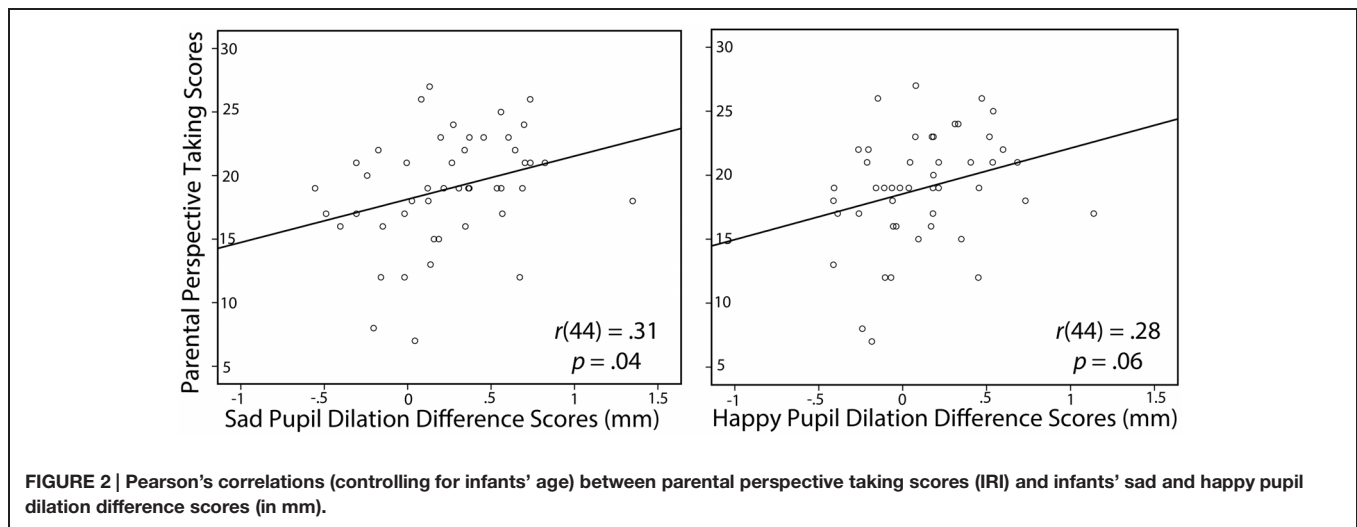
However, the primary aim of the study was to investigate relations between infants' pupillary changes to others' emotional displays and parents' empathic and prosocial dispositions. We found that empathic perspective taking, which reflects parents' tendency to adopt the perspectives of others in social situations, was associated with increases in infants' pupil dilation in response to other infants' happy and sad emotional expressions. In contrast, we found that affective dimensions of parental empathy, specifically empathic concern, were not associated with infants' pupillary responses toward others' emotions. These findings suggest that parents who more frequently take other people's perspectives are more likely to have infants who exhibit greater degrees of arousal in response to others' emotions. In addition, we found that parents' frequency of self-reported altruistic behavior was related to infants' arousal in response to others' emotions, particularly happiness, such that parents who more frequently act altruistically toward others are more likely to have

infants who exhibit greater arousal in response to others' happy emotions.

These findings raise some interesting possibilities regarding why parents' empathic and prosocial dispositions were related to infants' arousal in response to others' emotions. One possibility is that adopting another person's perspective and becoming aroused in response to others' emotions share certain characteristics (e.g., Bengtsson and Arvidsson, 2011; Missana et al., 2014). That is, at baseline, both perspective taking and arousal in response to others' emotions requires one to register that another person is having a different experience than one's own. Thus, the significant relation between parental perspective taking and infants' arousal could reflect shared abilities to recognize the experiences of others, particularly when others' experiences differ from one's own. It is possible that these shared abilities arise because of individual differences in how parents socialize their infants. For example, parents who are high in perspective taking have been found to provide their children with increased opportunities to recognize others' unique perspectives, including their differing emotional states, by highlighting and emphasizing these in their interactions (Soenens et al., 2007; Farrant et al., 2012; van Lissa et al., 2014). Accordingly, it is possible that parents who score high in perspective taking are more likely to mirror their infants' and others' emotional expressions, which could provide infants with the relevant experience needed to register and become aroused by another's emotional state. Indeed, adults who report higher levels of empathic perspective taking are more likely to mimic the behavior and expressions of those around them (e.g., Chartrand and Bargh, 1999; Galinsky et al., 2005; Genschow et al., 2013); in turn, imitation of emotional expressions has been linked to increased empathic perspective taking (Stel and Vonk, 2009). Thus, the relation between parental perspective taking and infants' arousal could also be explained by this disposition influencing how frequently parents engage in activities that serve to identify and reflect other people's behaviors and emotions. Of course, it is also possible that the relation between infants' arousal toward others' emotions and parental perspective

taking represents shared genetic tendencies between parents and infants, and/or an interaction between genetics and socialization. Future work should seek to clarify the exact nature of this relation.

We also found an association between parents' self-reported altruism and infants' arousal in response to others' emotions, and particularly their arousal in response to another's display of happiness. One possible reason for this relation is that helping others, and becoming aroused in response to others' happiness, share similar characteristics. That is, a strong motivator for performing helpful acts is a drive to see others' happiness that results from offering help (Aknin et al., 2013; Barasch et al., 2014; Dunn et al., 2014; see also Telle and Pfister, 2012). For example, being able to see one's recipient, as opposed to giving to others remotely, is associated with increased rates of charitable giving (Aknin et al., 2013). This suggests that seeing another person's happiness, and knowing that one's actions are the impetus for that emotion, is a strong motivator for altruistic behavior. In addition, it is possible that there are individual differences in how aroused adults are toward others' displays of happiness (Heller, 1993). Thus, the selective association between infants' arousal in response to others' happiness and parental altruistic behavior may reflect a shared tendency to become aroused or motivated by another's happiness. Similarly, and in support of this idea, Hepach et al. (2013) found that 2-year-old children's pupil dilation predicted the speed with which children offered to help another person in need, such that greater pupil dilation was associated with quicker offers to help (Hepach et al., 2013; see also Hepach et al., 2012). Importantly, children in this study were not explicitly rewarded or recognized for offering to help the other person, which suggests that children were otherwise motivated, presumably by a desire to elicit and witness another person's happiness. However, another, non-mutually exclusive possibility for this relation, given that helpful acts are likely to elicit happy emotions in both the helper and the helped, is that parents who perform more helpful acts toward others provide their infants with increased experience seeing other people's happiness. As the frequency of observing emotions is linked to one's ability to recognize and process them (Calvo et al., 2014), it is likely that infants with increased experience observing others' happiness are better able to register and become aroused by others' happiness displays. Thus, either a shared sensitivity toward others' happiness, or differences in exposure to others' happy emotional displays, may account for the relation between parents' self-reported altruism and infants' arousal toward others' happiness in the present study.

Intriguingly, we did not find an association between parents' empathic concern and infants' arousal toward others' emotions. This null relation is interesting in part because empathic concern assesses one's other-oriented affective response toward another person's distress, which ostensibly bears similarity to infants' arousal in response to others' emotions. One possibility for this null relation is methodological: self-report measures of empathy, and particularly measures of empathic concern (see Einolf, 2008, for a review), are known to elicit socially desirable responses (Watson and Morris, 1991; Litvack-Miller et al., 1997; Zaki, 2014; see also Eisenberg and Miller, 1987), which would

hamper demonstrating associations with infants' arousal toward others' emotions. Another possibility is that parents exhibit relatively homogenous, and high, levels of empathic concern in their real-life behavior toward their young infants, which is in contrast to what they report exhibiting in everyday life on questionnaire measures. In other words, given that infants are such compellingly helpless and adorable individuals, most anyone would be expected to exhibit high levels of empathic concern toward them, even those who ordinarily demonstrate low levels of empathic concern toward others. If this is the case, questionnaire measures may not accurately assess the level of empathic concern that parents demonstrate toward their infants, which would account for the lack of relation between self-reports of dispositional empathic concern and variability in infants' arousal to others' emotions in the present study. Future work may seek to further explore these possibilities.

Another issue that bears consideration is the meaning of infants' arousal in response to others' emotions, or what infants' arousal in response to others' emotions reflects. One possibility is that pupil dilation in response to others' emotions reflects infants' own feelings of personal distress. However, we believe that this is unlikely for several reasons. First, infants showed arousal in response to others' expressions of both happiness and sadness, the former of which would not be expected to elicit distress. Second, parental personal distress showed a marginal negative relation with infants' arousal toward others' emotions, which indicates that parents with higher levels of personal distress had infants who exhibited less arousal in response to others' emotional displays, which is the opposite of what would be expected if infants' pupil dilation reflected personal distress. Lastly, no infants cried during observation of the videos, even the sad, and crying is commonly operationalized as reflecting personal distress in studies of early empathy (e.g., Roth-Hanania et al., 2011). Altogether, this demonstrates that infants registered and were subsequently aroused by the other infants' emotional displays, without becoming upset by them. Indeed, in contrast to personal distress, we propose that infants' arousal in response to others' emotions reflects infants' emerging sense of emotional attunement with others, or their sense of connectedness and responsivity to others' emotions (see Markova and Legerstee, 2006). Indeed, emotional attunement is thought to be related to empathy (Gallese et al., 2007), which further suggests that infants' arousal in response to others' emotions reflects emotional attunement as opposed to personal distress.

More broadly, this study fits nicely into the literature on the development of empathy and earlier emerging precursors in young children. Specifically, this study compliments past work on infants' arousal in response to others' emotions by confirming that infants exhibit heightened and differential arousal toward others' emotions (i.e., happiness and sadness) by the end of the first year of life (Geangu et al., 2011b). In addition, this study extends upon prior work that has investigated precursors to empathic responding (e.g., Sagi and Hoffman, 1976; Martin and Clark, 1982; Haviland and Lelwica, 1987; Termine and Izard, 1988; Dondi et al., 1999) by demonstrating that variability in infants' arousal toward others' emotions is accounted for by differences in their parents' empathic and prosocial dispositions.

This is important, as it provides further evidence that a precursor to empathic responding is meaningfully connected to mature empathy and theoretically aligned behaviors (Roth-Hanania et al., 2011). Accordingly, the present study encourages continued investigation into relations between variability in precursors to empathy and variability in fully developed empathic responding, in an effort to better understand its developmental trajectory; moreover, the present design provides a methodology for doing so. For example, an interesting question for future research would be to investigate whether variability in infants' arousal in response to others' emotions, as indexed by pupil dilation, is predictive of empathic dispositions in early childhood. In addition, this study calls for more work investigating parental dispositions as a source of variability in children's early empathic responses. For example, future work may seek to investigate how heritability and socialization contribute to the relation between parental dispositions and infants' arousal toward others' emotions.

Conclusion

This study supports investigating individual variability in infants' early empathic responses as a phenomena of interest, rather than treating such variability as noise. Our results suggest that individual variability in infants' arousal toward others' emotions indexes meaningful differences in the manner in which infants are processing the social world. In addition, the present study confirms and extends upon prior work that has found a strong association between parental behaviors and their children's developing empathy, by demonstrating that parents, construed as individuals, and not just in their capacity as parents *per se*, are significant predictors of their infants' automatic responses to another's emotional state. Altogether, the present study highlights the merits of using pupil dilation in response to others' emotions as a measure of children's emerging empathy and encourages directly investigating parental dispositions as a source of variability in these early empathic responses.

References

Aknin, L. B., Dunn, E. W., Sandstrom, G. M., and Norton, M. I. (2013). Does social connection turn good deeds into good feelings? On the value of putting the 'social' in prosocial spending. *Int. J. Happiness Dev.* 1, 155–171. doi: 10.1504/IJHD.2013.055643

Barasch, A., Levine, E. E., Berman, J. Z., and Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *J. Pers. Soc. Psychol.* 107, 393–413. doi: 10.1037/a0037207

Batson, C. D. (1991). *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale: Lawrence Erlbaum Associates, Inc.

Batson, C. D. (2002). "Addressing the altruism question experimentally," in *Altruism and Altruistic Love: Science, Philosophy, and Religion in Dialogue,* eds S. G. Post, L. G. Underwood, J. P. Schloss, and W. B. Hurlbut (New York, NY: Oxford University Press), 89–105. doi: 10.1093/acprof:oso/9780195143584.003.0010

Batson, C. D., Fultz, J., and Schoenrade, P. A. (1987). Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *J. Pers.* 55, 19–39. doi: 10.1111/j.1467-6494.1987.tb00426.x

Beatty, J., and Lucero-Wagoner, B. (2000). "The pupillary system," in *Handbook of Psychophysiology*, 2nd Edn, eds J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson (New York, NY: Cambridge University Press), 142–162.

Bengtsson, H., and Arvidsson, Å. (2011). The Impact of developing social perspective-taking skills on emotionality in middle and late childhood: perspective-taking skills and emotionality. *Soc. Dev.* 20, 353–375. doi: 10.1111/j.1467-9507.2010.00587.x

Beven, J. P., O'Brien-Malone, A., and Hall, G. (2004). Using the interpersonal reactivity index to assess empathy in violent offenders. *Int. J. Forensic Psychol.* 1, 33–41.

Bitsios, P., Szabadi, E., and Bradshaw, C. M. (2004). The fear-inhibited light reflex: importance of the anticipation of an aversive event. *Int. J. Psychophysiol.* 52, 87–95. doi: 10.1016/j.ijpsycho.2003.12.006

Bradley, M. M., Miccoli, L., Escrig, M. A., and Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45, 602–607. doi: 10.1111/j.1469-8986.2008.00654.x

Calvo, M. G., Gutiérrez-García, A., Fernández-Martín, A., and Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *J. Nonverbal. Behav.* 38, 549–567. doi: 10.1007/s10919-014-0191-3

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *J. Pers. Soc. Psychol.* 76, 893–910. doi: 10.1037/0022-3514.76.6.893

Davidov, M., and Grusec, J. E. (2006). Untangling the links of parental responsiveness to distress and warmth to child outcomes. *Child Dev.* 77, 44–58. doi: 10.1111/j.1467-8624.2006.00855.x

Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* 44, 113–126. doi: 10.1037/0022-3514.44.1.113

Davis, M. H. (2005). "Becoming (and remaining) a community volunteer: does personality matter?," in *Processes of Community Change and Social Action*, ed. A. M. Omoto (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 67–82.

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., and Meyer, M. (1999). Empathy, expectations, and situational preferences: personality influences on the decision to participate in volunteer helping behaviors. *J. Pers.* 67, 469–503. doi: 10.1111/1467-6494.00062

de Waal, F. B. M. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300. doi: 10.1146/annurev.psych.59.103006.093625

Dondi, M., Simion, F., and Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Dev. Psychol.* 35, 418–426. doi: 10.1037/0012-1649.35.2.418

Dovidio, J. F., Allen, J. L., and Schroeder, D. A. (1990). Specificity of empathy-induced helping: evidence for altruistic motivation. *J. Pers. Soc. Psychol.* 59, 249–206. doi: 10.1037/0022-3514.59.2.249

Dunn, E. W., Aknin, L. B., and Norton, M. I. (2014). Prosocial spending and happiness: using money to benefit others pays off. *Curr. Dir. Psychol. Sci.* 23, 41–47. doi: 10.1177/0963721413512503

Einolf, C. J. (2008). Empathic concern and prosocial behaviors: A test of experimental results using survey data. *Soc. Sci. Res.* 37, 1267–1279. doi: 10.1016/j.ssresearch.2007.06.003

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annu. Rev. Psychol.* 51, 665–697. doi: 10.1146/annurev.psych.51.1.665

Eisenberg, N., Fabes, R. A., Schaller, M., Carlo, G., and Miller, P. A. (1991). The relations of parental characteristics and practices to children's vicarious emotional responding. *Child Dev.* 62, 1393–1408. doi: 10.2307/1130814

Eisenberg, N., Fabes, R. A., Schaller, M., and Miller, P. A. (1989). Sympathy and personal distress: development, gender differences, and interrelations of indexes. *New Dir. Child Dev.* 107–126. doi: 10.1002/cd.23219894408

Eisenberg, N., and Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychol. Bull.* 101, 91–119. doi: 10.1037/0033-2909.101.1.91

Farrant, B. M., Devine, T. A. J., Maybery, M. T., and Fletcher, J. (2012). Empathy, perspective taking and prosocial behaviour: the importance of parenting practices: empathy, parenting and prosocial behaviour. *Infant Child Dev.* 21, 175–188. doi: 10.1002/icd.740

Galinsky, A. D., Ku, G., and Wang, C. S. (2005). Perspective-taking and self-other overlap: fostering social bonds and facilitating social coordination. *Group Process Intergroup Relat.* 8, 109–124. doi: 10.1177/1368430205051060

Gallese, V., Eagle, M. N., and Migone, P. (2007). Intentional attunement: mirror neurons and the neural underpinnings of interpersonal relations. *J. Am. Psychoanal. Assoc.* 55, 131–175. doi: 10.1177/00030651070550010601

Geangu, E., Benga, O., Stahl, D., and Striano, T. (2011a). Individual differences in infants' emotional resonance to a peer in distress: self-other awareness and emotion regulation: affect sharing in infancy. *Soc. Dev.* 20, 450–470. doi: 10.1111/j.1467-9507.2010.00596.x

Geangu, E., Hauf, P., Bhardwaj, R., and Bentz, W. (2011b). Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE* 6:e27132. doi: 10.1371/journal.pone.0027132

Genschow, O., Florack, A., and Wänke, M. (2013). The power of movement: evidence for context-independent movement imitation. *J. Exp. Psychol. Gen.* 142, 763–773. doi: 10.1037/a0029795

Gredebäck, G., Johnson, S., and von Hofsten, C. (2010). Eye tracking in infancy research. *Dev. Neuropsychol.* 35, 1–19. doi: 10.1080/87565640903325758

Haviland, J. M., and Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Dev. Psychol.* 23, 97–104. doi: 10.1037/0012-1649.23.1.97

Heller, W. (1993). Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychol.* 7, 476–489. doi: 10.1037/0894-4105.7.4.476

Hepach, R., Vaish, A., and Tomasello, M. (2012). Young children are intrinsically motivated to see others helped. *Psychol. Sci.* 23, 967–972. doi: 10.1177/0956797612440571

Hepach, R., Vaish, A., and Tomasello, M. (2013). A new look at children's prosocial motivation. *Infancy* 18, 67–90. doi: 10.1111/j.1532-7078.2012.00130.x

Hess, E. H., and Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science* 132, 132349–132350. doi: 10.1126/science.132.3423.349

Hoffman, M. L. (1982). Empathy and guilt feelings, which develop naturally under ordinary conditions, may be significant affects underlying moral action. *Child. Concept. Spat. Relatsh.* 15, 83.

Hoffman, M. L. (2000). *Empathy and Moral Development: Implications for Caring and Justice.* New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511805851

Hoffman, M. L. (2008). "Empathy and prosocial behavior," in *Handbook of Emotions*, 3rd Edn, eds M. Lewis, J. M. Haviland-Jones, and L. F. Barrett (New York, NY: Guilford Press), 440–455.

Jackson, I., and Sirois, S. (2009). Infant cognition: Going full factorial with pupil dilation. *Dev. Sci.* 12, 670–679. doi: 10.1111/j.1467-7687.2008.00805.x

Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., and Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: genetic and environmental contributions. *Emotion* 8, 737–752. doi: 10.1037/a0014179

Koller, I., and Lamm, C. (2014). Item response model investigation of the (German) interpersonal reactivity index empathy questionnaire: implications for analyses of group differences. *Eur. J. Psychol. Assess.* 1, 1–11. doi: 10.1027/1015-5759/a000227

Laeng, B., Sirois, S., and Gredeback, G. (2012). Pupillometry: a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi: 10.1177/1745691611427305

Litvack-Miller, W., McDougall, D., and Romney, D. M. (1997). The structure of empathy during middle childhood and its relationship to prosocial behavior. *Genet. Soc. Gen. Psychol. Monogr.* 123, 303–324. doi: 10.1111/j.1559-1816.2011.00882.x

Markova, G., and Legerstee, M. (2006). Contingency, imitation, and affect sharing: foundations of infants' social awareness. *Dev. Psychol.* 42, 132–141. doi: 10.1037/0012-1649.42.1.132

Martin, G. B., and Clark, R. D. (1982). Distress crying in neonates: species and peer specificity. *Dev. Psychol.* 18, 3–9. doi: 10.1037/0012-1649.18.1.3

Miller, S. M. (1979). Interrelationships among dependency, empathy, and sharing: a preliminary study. *Moti. Emt.* 3, 183–199. doi: 10.1007/BF01650603

Missana, M., Grigutsch, M., and Grossmann, T. (2014). Developmental and individual differences in the neural processing of dynamic expressions of pain and anger. *PLoS ONE* 9:e93728. doi: 10.1371/journal.pone.0093728

Moreno, A. J., Klute, M. M., and Robinson, J. L. (2008). Relational and individual resources as predictors of empathy in early childhood. *Soc. Dev.* 17, 613–637. doi: 10.1111/j.1467-9507.2007.00441.x

Newton, E. K., Laible, D., Carlo, G., Steele, J. S., and McGinley, M. (2014). Do sensitive parents foster kind children, or vice versa? Bidirectional influences between children's prosocial behavior and parental sensitivity. *Dev. Psychol.* 50, 1808–1816. doi: 10.1037/a0036495

Partala, T., and Surakka, V. (2003). Pupil size variation as an indication of affective processing. *Int. J. Hum. Comput. Stud.* 59, 185–198. doi: 10.1016/S1071-5819(03)00017-X

Penner, L. A., Fritzsche, B. A., Craiger, J. P., and Freifeld, T. S. (1995). "Measuring the prosocial personality," in *Advances in Personality Assessment*, Vol. 10, eds J. N. Butcher and C. D. Spielberger (Hillsdale: Lawrence Erlbaum Associates, Inc), 147–163.

Porter, G., Troscianko, T., and Gilchrist, I. D. (2007). Effort during visual search and counting: insights from pupillometry. *Q. J. Exp. Psychol.* 60, 211–229. doi: 10.1080/17470210600673818

Preston, S. D., and de Waal, F. (2002). Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.

Pulos, S., Elison, J., and Lennon, R. (2004). The hierarchical structure of the interpersonal reactivity index. *Soc. Behav. Pers.* 32, 355–359. doi: 10.2224/sbp.2004.32.4.355

Robinson, J., Zahn-Waxler, C., and Emde, R. N. (2001). "Relationship context as a moderator of sources of individual differences in empathic development," in *Infancy to Early Childhood: Genetic and Environmental Influences on Developmental Change* eds R. N. Emde and J. K. Hewitt (New York, NY: Oxford University Press), 257–268

Roth-Hanania, R., Davidov, M., and Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: early signs of concern for others. *Infant Behav. Dev.* 34, 447–458. doi: 10.1016/j.infbeh.2011.04.007

Sagi, A., and Hoffman, M. L. (1976). Empathic distress in the newborn. *Dev. Psychol.* 12, 175–176. doi: 10.1037/0012-1649.12.2.175

Soenens, B., Duriez, B., Vansteenkiste, M., and Goossens, L. (2007). The intergenerational transmission of empathy-related responding in adolescence: the role of maternal support. *Pers. Soc. Psychol. Bull.* 33, 299–311. doi: 10.1177/0146167206296300

Spinrad, T. L., and Stifter, C. A. (2006). Toddlers' empathy-related responding to distress: predictions from negative emotionality and maternal behavior in infancy. *Infancy* 10, 97–121. doi: 10.1207/s15327078in1002_1

Stel, M., and Vonk, R. (2009). Empathizing via mimicry depends on whether emotional expressions are seen as real. *Eur. Psychol.* 14, 342–350. doi: 10.1027/1016-9040.14.4.342

Strayer, J., and Roberts, W. (2004). Children's anger, emotional expressiveness, and empathy: relations with parents' empathy, emotional expressiveness, and parenting practices. *Soc. Dev.* 13, 229–254. doi: 10.1111/j.1467-9507.2004.000265.x

Taylor, Z. E., Eisenberg, N., Spinrad, T. L., Eggum, N. D., and Sulik, M. J. (2013). The relations of ego-resiliency and emotion socialization to the development of empathy and prosocial behavior across early childhood. *Emotion* 13, 822–831. doi: 10.1037/a0032894

Telle, N. -T., and Pfister, H. -R. (2012). Not only the miserable receive help: empathy promotes prosocial behaviour toward the happy. *Curr. Psychol.* 31, 393–413. doi: 10.1007/s12144-012-9157-y

Termine, N. T., and Izard, C. E. (1988). Infants' responses to their mothers' expressions of joy and sadness. *Dev. Psychol.* 24, 223–229. doi: 10.1037/0012-1649.24.2.223

Vanderhasselt, M. -A., Remue, J., Ng, K. K., and De Raedt, R. (2014). The interplay between the anticipation and subsequent online processing of emotional stimuli as measured by pupillary dilatation: the role of cognitive reappraisal. *Front. Psychol.* 5:207. doi: 10.3389/fpsyg.2014.00207

van Lissa, C. J., Hawk, S. T., de Wied, M., Koot, H. M., van Lier, P., and Meeus, W. (2014). The longitudinal interplay of affective and cognitive empathy within and between adolescents and mothers. *Dev. Psychol.* 50, 1219–1225. doi: 10.1037/a0035050

Volbrecht, M. M., Lemery-Chalfant, K., Aksan, N., Zahn-Waxler, C., and Goldsmith, H. H. (2007). Examining the familial link between positive affect and empathy development in the second year. *J. Genet. Psychol.* 168, 105–130. doi: 10.3200/GNTP.168.2.105-130

Volling, B. L., Kolak, A. M., and Kennedy, D. E. (2008). "Empathy and compassionate love in early childhood: development and family influence," in *The Science of Compassionate Love: Theory, Research, and Applications*, eds B. Fehr, S. Sprecher, and L. G. Underwood (Wiley-Blackwell: Hoboken), 161–200. doi: 10.1002/9781444303070.ch6

Watson, P. J., and Morris, R. J. (1991). Narcissism, empathy and social desirability. *Pers. Individ. Dif.* 12, 575–579. doi: 10.1016/0191-8869(91)90253-8

Young, S. K., Fox, N. A., and Zahn-Waxler, C. (1999). The relations between temperament and empathy in 2-year-olds. *Dev. Psychol.* 35, 1189–1197. doi: 10.1037/0012-1649.35.5.1189

Zahn-Waxler, C., and Radke-Yarrow, M. (1990). The origins of empathic concern. *Mot. Emt.* 14, 107–130. doi: 10.1007/BF00991639

Zahn-Waxler, C., Radke-Yarrow, M., and King, R. A. (1979). Child rearing and children's prosocial initiations toward victims of distress. *Child Dev.* 50, 319–330. doi: 10.2307/1129406

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., and Chapman, M. (1992a). Development of concern for others. *Dev. Psychol.* 28, 126–136. doi: 10.1037/0012-1649.28.1.126

Zahn-Waxler, C., Robinson, J. L., and Emde, R. N. (1992b). The development of empathy in twins. *Dev. Psychol.* 28, 1038–1047. doi: 10.1037/0012-1649.28.6.1038

Zahn-Waxler, C., Schiro, K., Robinson, J. L., Emde, R. N., and Schmitz, S. (2001). "Empathy and prosocial patterns in young MZ and DZ twins: development and genetic and environmental influences," in *Infancy to Early Childhood: Genetic and Environmental Influences on Developmental Change*, eds R. N. Emde and J. K. Hewitt (New York, NY: Oxford University Press), 141–162.

Zaki, J. (2014). Empathy: a motivated account. *Psychol. Bull.* 140, 1608–1647. doi: 10.1037/a0037679

Zhou, Q., Eisenberg, N., Losoya, S. H., Fabes, R. A., Reiser, M., Guthrie, I. K., et al. (2002). The relations of parental warmth and positive expressiveness to children's empathy-related responding and social functioning: a longitudinal study. *Child Dev.* 73, 893–915. doi: 10.1111/1467-8624.00446

Individual differences in toddlers' social understanding and prosocial behavior: disposition or socialization?

Rebekkah L. Gross[1]*, Jesse Drummond[1], Emma Satlof-Bedrick[1], Whitney E. Waugh[1], Margarita Svetlova[2] and Celia A. Brownell[1]

[1] Early Social Development Lab, Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA, [2] Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

We examined how individual differences in social understanding contribute to variability in early-appearing prosocial behavior. Moreover, potential sources of variability in social understanding were explored and examined as additional possible predictors of prosocial behavior. Using a multi-method approach with both observed and parent-report measures, 325 children aged 18–30 months were administered measures of social understanding (e.g., use of emotion words; self-understanding), prosocial behavior (in separate tasks measuring instrumental helping, empathic helping, and sharing, as well as parent-reported prosociality at home), temperament (fearfulness, shyness, and social fear), and parental socialization of prosocial behavior in the family. Individual differences in social understanding predicted variability in empathic helping and parent-reported prosociality, but not instrumental helping or sharing. Parental socialization of prosocial behavior was positively associated with toddlers' social understanding, prosocial behavior at home, and instrumental helping in the lab, and negatively associated with sharing (possibly reflecting parents' increased efforts to encourage children who were less likely to share). Further, socialization moderated the association between social understanding and prosocial behavior, such that social understanding was less predictive of prosocial behavior among children whose parents took a more active role in socializing their prosociality. None of the dimensions of temperament was associated with either social understanding or prosocial behavior. Parental socialization of prosocial behavior is thus an important source of variability in children's early prosociality, acting in concert with early differences in social understanding, with different patterns of influence for different subtypes of prosocial behavior.

Keywords: prosocial behavior, social understanding, temperament, parent socialization, individual differences

Introduction

Remarkably, while still learning how to speak, toddlers in their second and third years of life are attentive to their own and others' internal states and begin engaging in prosocial behaviors such as helping and sharing. Yet, from the outset young children exhibit wide variability in the complexity

and frequency of prosocial action (e.g., Brownell et al., 2006, 2013a; Warneken and Tomasello, 2006; Warneken et al., 2006; Svetlova et al., 2010; Dunfield et al., 2011). To engage in prosocial behavior, one must be able to recognize another's goal, desire, or internal state, and respond to another's needs by intervening to alter their subjective state. Thus, even in its earliest manifestations, prosocial behavior such as helping and sharing depends on and reflects social perception and social understanding (Vaish and Warneken, 2012; Brownell et al., 2013b).

Social understanding – the ability to infer others' internal states such as goals, feelings, and desires – has its origins in the first year of life and gives rise to a variety of other-oriented behaviors at the beginning of the second year (Tomasello et al., 2005). Because social cognition is necessary for prosocial behavior, we might reasonably expect that individual differences in social cognition would relate to individual differences in prosocial behavior. And, indeed, this is known to be the case for older children in whom both social cognition and prosocial behavior are well-developed (Eisenberg et al., 2015). However, there are few studies that examine how variability in very early social understanding relates to emerging prosocial behavior. Thus, their association is not yet clear during the period when prosocial behavior first emerges, and when both systems are undergoing rapid and dramatic change. The current study uses multiple measures of social understanding and prosocial behavior to examine this association in one- and 2-years-old children. The study also includes measures of temperament and parent socialization to explore possible sources of individual differences in toddlers' social understanding and prosocial behavior.

Relations between Social Understanding and Prosocial Behavior in Toddlers

While prosocial behavior, such as handing a blanket to someone who is cold, may seem simple on the surface, such acts require complex understanding and action. The child must recognize that someone else has a problem, which may include understanding emotional facial expressions and their relations to others' internal feelings; understanding that the distressed party is a separate entity with unique desires and goals, while also regulating one's own desires and emotional responses; and understanding the specific type of assistance required and how to intervene, and adapting to any obstacles along the way. A small body of research has demonstrated connections between social understanding and some aspects of prosocial behavior in toddlers. For example, emotion and self-other understanding in 12- to 24-months-old were related to empathic responsiveness to a peer's distress (Nichols et al., 2009). Fifteen-months-old infants who were more sensitive to unfair outcomes were also found to be more willing to share their preferred toy with an adult (Sommerville et al., 2013). Other studies have found associations between empathic concern, self- and other-understanding, and emotion understanding in young children (Bischof-Köhler, 1991; Zahn-Waxler et al., 1992; Ensor and Hughes, 2005; Garner et al., 2008).

Prosocial behavior may be best conceptualized as a multidimensional construct, at least in the early years (Eisenberg and Spinrad, 2014). Indeed, distinct prosocial behaviors are often uncorrelated in toddlers (Dunfield et al., 2011) and have different

neural signatures in the second and third years of life (Paulus et al., 2013). Thus, some prosocial behaviors may rely more heavily on social understanding, such as those that involve inferring an internal state (e.g., empathic helping), in contrast to others that are more strictly goal-oriented (e.g., instrumental helping). Hence, the current study builds on existing research by including measures of several distinct types of prosocial behavior to explore whether social understanding plays a different role for different types of prosocial action.

Individual Differences in Early Social Understanding and Prosocial Behavior

Questions about relations between variability in social understanding and variability in prosocial behavior are fundamentally individual difference questions: to what extent do within-age differences in social understanding account for differences in prosocial behavior? Yet these are also normative developmental phenomena – all typically developing children achieve the basic abilities to help, share, comfort, and cooperate with others, and to represent and act on their own and others' internal subjective states. One source of variability in such normative developments is differences in rates of development. That is, at any given age, some children may be more advanced than others in social cognition and/or social behavior. One potential contributor to such within-age differences in competence may be parents' socialization of their children's responses to others' emotions and behavior (Brownell, 2013; Denham et al., 2015; Eisenberg et al., 2015). Parent socialization, considered broadly, comprises the myriad ways parents help their children to become members of the social group. Socialization is a function of parental beliefs, goals, and values; occurs within many different contexts including play as well as discipline; and includes behavior ranging from the subtle (e.g., praising process vs. outcome) and indirect (e.g., monitoring) to more explicit and didactic (e.g., rewarding; coaching; Eisenberg et al., 1998; Grusec and Hastings, 2007).

A second source of individual differences in social understanding and/or prosocial behavior may be dispositional differences in children's attention to or interest in social and emotional information or social engagement. For example, research has suggested the existence of an "empathic disposition," such that some young children are dispositionally more likely than others to empathize with others in distress, regardless of their early socialization experiences (Nichols et al., 2009). Other dispositional or temperamental differences may affect children's attentiveness to others. These possibilities are considered more fully below.

Parental Socialization

Parental socialization of everyday activities is a known contributor to both prosocial behavior and social understanding throughout childhood (Denham et al., 2015; Hastings et al., 2015). Previous research has revealed that parents utilize a variety of socialization strategies to encourage young children's developing prosocial behaviors, including negotiation (Crockenberg and Litman, 1990), scaffolding (Hammond et al., 2012), and praise (Grusec, 1991) and that socialization approaches vary with the

age of the toddler (Pettygrove et al., 2013; Dahl, 2015; Waugh et al., 2015). Research has also shown that the content and context of parent socialization are associated with young children's prosocial behavior. Toddlers whose mothers scaffolded their everyday helping were more helpful toward an unfamiliar adult (Pettygrove et al., 2013; Hammond and Carpendale, 2015). Those whose parents engaged them in more emotion-related discourse during joint book reading, and who were particularly asked by their parents to attend to and reflect on others' emotions, helped more quickly and more often on emotionally laden helping and sharing tasks with other adults (Garner et al., 2008; Brownell et al., 2013c; Drummond et al., 2014b). Thus, parents socialize early-appearing prosocial behavior in multiple ways, both by scaffolding very early instances of prosocial responding and by drawing their children's attention to the mental and emotional states of those around them.

Parent socialization may similarly influence developing social cognition (Carpendale and Lewis, 2004). Parents' self-reported beliefs about the importance of teaching children about emotions, as well as their self-reported attention to and encouragement of their young children's emotions and their observed responses to children's emotions were associated with emotion understanding longitudinally in 3- and 5-years-old (Denham et al., 1994; Denham and Kochanoff, 2002). A number of researchers have reported associations between parent–child mental state talk and preschool children's understanding of emotions and other psychological states (e.g., Lagattuta and Wellman, 2002; Ruffman et al., 2002; Taumoepeau and Ruffman, 2006; LaBounty et al., 2008), particularly when the discourse occurs within socially connected interchanges (Ensor and Hughes, 2008). Among younger children, those with more sensitive, "mind-minded" parents who respond to their infants "as individuals with minds" tend to later exhibit more advanced mentalizing abilities such as false-belief understanding (Meins and Fernyhough, 1999). During the second year of life, as toddlers begin to label their own and others' emotions, parents also begin to discuss emotions in causal terms (Bretherton et al., 1986). Thus, as with prosocial behavior, parents socialize early-developing social understanding in multiple ways, both direct and indirect.

Building on this empirical base, the current study examined potential associations between parents' self-reported socialization practices and their toddler-aged children's social understanding and prosocial behavior. In particular, we asked parents to report their everyday socialization practices that focused on scaffolding and encouragement of the child's prosocial behavior as well as on the child's attention to and discussion of emotions.

Temperament

Several studies have demonstrated associations between specific dimensions of temperament and social understanding in young children. Using standard false-belief theory of mind (ToM) tasks and parent-reported temperament measures, Wellman et al. (2011) found that a shy-withdrawn temperament at age 3 predicted later ToM understanding at age 5. Shyness in children as young as 18 months has been found to be positively correlated with ToM understanding at 3 years of age (Mink

et al., 2014). Similarly, children who tended to observe their peers rather than play actively with them, possibly denoting shyness or social cautiousness, were more advanced in ToM (Moore et al., 2011). Wellman et al. (2011) reasoned that a shy-observant, possibly more regulated and cautious, approach to social interaction might enhance children's ability to take a reflective stance on others' behavior, thereby contributing to a developing understanding of its causes in underlying mental states. While previous research provides insight into the relationship between temperament and social understanding during the preschool years, to our knowledge no research has tested the concurrent associations between a cautious, shy, possibly more fearful temperament and social understanding in infants; this is one of the aims of the current study.

Associations between temperamental fearfulness and prosocial responding have also been examined, with mixed results. Spinrad and Stifter (2006) reported that fearfulness assessed at 10 months of age predicted greater concern toward a distressed adult at 18 months of age. In contrast, van der Mark et al. (2002) found that fearfulness observed at 16 months of age was associated with reduced empathic concern at 22 months; and Liew et al. (2011) found no association between fearfulness at 18 months and concern for another's distress at 30 months. Moreover, neither Spinrad and Stifter (2006) nor Liew et al. (2011) found any links between temperamental fear and actual prosocial behavior such as comforting or helping. Whether these inconsistent findings reflect variation in how temperament was measured, the particular dispositional constructs assessed (e.g., fearfulness vs. shyness), or different patterns in younger children than in older children is unknown. It thus remains an open question whether or how early temperament relates to early prosocial behavior. The current study adds to this literature by exploring concurrent associations between fearfulness, shyness, and social fearfulness and prosocial behavior in 18- to 30-months-old toddlers.

In sum, the first goal of the current study was to determine whether and how variability in toddlers' social understanding predicted their prosocial behavior. We assessed social understanding using parent report of their children's self-other differentiation and emotion-related vocabulary. We included multiple types of prosocial behavior, both as it occurs in the family environment as reported by parents and as observed in lab tasks of sharing and helping with other adults. We then examined two potential sources of individual variability in social understanding and prosocial behavior; specifically, parent socialization and specific dimensions of temperament known to predict social understanding in preschoolers and sometimes to predict prosocial behavior in toddlers.

Materials and Methods

Participants

Participants were drawn from five previously completed studies of early prosocial behavior. They included 135 18-months-olds ($M = 18.32$, SD $= 0.681$; girls $= 65$; boys $= 70$), 56 24-months-old ($M = 23.36$, SD $= 1.45$; girls $= 27$; boys $= 29$), and 134

30-months-old ($M = 29.18$, SD $= 1.03$; girls $= 59$; boys $= 75$). Families were recruited from a medium-sized city and surrounding suburbs and received a small book or toy for completing the study. All participants were healthy and typically developing. Eighty-two percent were Caucasian, 7% were biracial, 4% were African-American, 3% were Asian, 1% was Hispanic, and 3% did not disclose their race. No participants took part in more than one of the studies. University IRB approval was obtained prior to initiating each study, SRCD ethical guidelines were followed, and parents provided written consent for their own and their children's participation.

General Procedure

Procedures were similar for each study. After a brief warm-up play period in a separate room with an experimenter (E) and an assistant experimenter (AE), the child and a parent were escorted to a playroom where the study procedures were conducted. All sessions were video-recorded from behind a one-way mirror. During the study, the parent remained in the room with the child, filling out questionnaires; parents were asked not to instruct or encourage their children during the session, but otherwise to respond naturally to their children's communications or social bids. Parents completed questionnaires rating children's temperament, social understanding, prosocial behavior, and parental socialization practices (details below). Additionally, children participated in sharing and/or helping tasks with E in each study. Brief periods of free play with a standard set of toys occurred between prosocial tasks; in some cases, children completed additional procedures for a larger study between the prosocial tasks. Within each study, all tasks were administered within-subjects and counterbalanced for order.

The helping and sharing tasks followed a similar format. For each one, E needed an object or objects to which the child had access but which were out of E's reach. The child could alleviate E's need or desire by giving her or him one or more relevant objects. Two types of helping tasks were administered: instrumental helping tasks in which E dropped or misplaced an object that s/he needed to complete a goal-directed action; and empathic helping tasks in which E experienced a negative internal state such as being cold or sad that could be alleviated by a blanket or a favorite toy. In the sharing tasks the child had an abundance of objects (e.g., cars, zoo animals) and E had none. On each of the helping and sharing tasks E delivered a standard series of cues, which became progressively more detailed and specific about E's need or desire and how it could be alleviated. No thanks or praise was provided when children helped or shared with E. Children were given a score of 1 if they helped or shared on any trial and a score of 0 if they did not help or share on any trial. This dichotomous measure of helping is commonly used when evaluating prosocial behavior in infants and toddlers (e.g., Warneken and Tomasello, 2006, 2007; Over and Carpenter, 2009; Dunfield et al., 2011; Sommerville et al., 2013).

Details for each study are provided below. Because some of the measures differed across studies (e.g., temperament; sharing), different subsamples of participants contributed to the analyses. **Table 1** provides a summary of the variables from each study that were used in the current analyses.

TABLE 1 | Ages assessed and measures available for analysis from each study.

Study	Age (in months)	N	Social Und: EWCL+UCLA	Temp: ECBQ (fear)	Temp: ECBQ (shy)	Temp: TBAQ (social fear)	Parent socialization: PBQ	Instrum helping	Emp helping	Sharing	Parent report – child Prosocial: PBQ + SDQ
1	18, 30	66	×				×	×	×	×	×
2	18, 30	58	×				×	×	×		×
3	30	45	×	×	×			×	×		
4	18, 24	82	×			×	×			×	×
5	18, 30	74		×	×		×	×	×		

Studies 1 and 2

Studies 1 and 2 (Waugh et al., 2013) included 71 18-months-old ($M = 18.61$, SD $= 0.843$), and 53 30-months-old ($M = 28.99$, SD $= 0.973$). Both studies included an instrumental helping task and an empathic helping task; Study 1 additionally included two sharing tasks.

In the instrumental helping task (adapted from Over and Carpenter, 2009), while E kneeled to place some things on a small table, she "accidentally" dropped six sticks on the floor on the far side of the table. During the empathic helping task (adapted from Svetlova et al., 2010) E became cold while sitting on one side of the room, having placed her blanket on a table across the room; prior to the empathic helping task, E modeled being cold and demonstrated that her blanket made her warm.

The two sharing tasks (adapted from Brownell et al., 2013a) were administered by E while AE and the child sat next to each other at side-by-side tables. AE served as a playmate and did not direct the child or the activities. The two tasks were identical but for different toys (cars; animals). E first evenly distributed several toys to both AE and the child; after a 60 s free-play period, E removed the toys from AE and the child, placed all of toys in front of the child, and moved to a corner of the room behind the child.

Study 3

Study 3 (Drummond et al., 2014a) included 45 30-months-old ($M = 28.73$, SD $= 1.136$) who were administered two instrumental and two empathic helping tasks. Two of these were identical to the tasks in Studies 1 and 2 (dropped sticks; cold) and two were unique to this study. In the new instrumental helping task, E "accidentally" dropped a stack of papers from a high cabinet onto the floor while reaching into the cabinet. In the new empathic helping task, E became sad after receiving a phone call; he had previously shown the child that his favorite toy made him happy, but it was now on a table out of his reach.

Study 4

Study 4 (Brownell et al., 2013a) included 26 18-months-old ($M = 18.0$ months; SD $= 0.5$) and 56 24-months-old ($M = 23.4$ months; SD $= 1.45$). Six sharing tasks, differing only in the toys to be shared, were administered following the same procedures as in Study 1. See Brownell et al. (2013a) for detailed description of tasks and procedures.

Study 5

Study 5 (Svetlova et al., 2010) included 38 18-months-old ($M = 18.46$ months; SD $= 0.48$) and 36 30-months-old ($M = 30.32$ months; SD $= 0.68$). Three instrumental helping tasks and three empathic helping tasks were administered following the same procedures as in Study 3. The three instrumental helping tasks were unique to this study; two of the empathic helping tasks were identical to those used in Study 3 (cold; sad) and one was unique to this study. In the first instrumental helping task, E dropped a clothespin out of reach while clipping cloths to a clothesline (adapted from Warneken and Tomasello, 2006). In the second instrumental helping task, E ran out of wrappers while wrapping toys and the additional wrappers were out of her reach. In the third instrumental helping task, E needed a toy she had been playing with but it was out of her reach. In the new empathic helping task, E became frustrated with her messy hair hanging in her eyes and needed a hairclip that was out of reach (E had previously demonstrated that the hairclip was used to clip her hair up and that it alleviated the frustration with her messy hair). See Svetlova et al. (2010) for detailed descriptions of tasks and procedures.

Questionnaires

During each study, parents filled out questionnaires about their child including temperament, social understanding (two questionnaires), prosocial behavior (two questionnaires), and parental socialization practices. Cronbach's alphas ranged from 0.79 to 0.97 across the studies. While not all studies contained the same questionnaires, the questionnaires that were used across studies were identical. Please see **Table 1** for a summary of the measures used in each study that are included in analyses.

Social Understanding

Social understanding was measured using a composite score from the UCLA Self-Understanding questionnaire (Stipek et al., 1990) and the Emotion Words Checklist (EWCL; Brownell et al., 2006). The UCLA Self-Understanding questionnaire consists of 24 items rated on a 3-point scale (0 = definitely not; 1 = sometimes; 2 = definitely) that evaluate self-recognition, self-description, and self-evaluation in toddlers and young preschoolers. The EWCL consists of a list of 29 emotion-related words adapted from Bretherton and Beeghly (1982) and Shatz et al. (1983). Parents indicated how often the child had said each emotion word in the past 6 months (0 = Never; 1 = once or twice; 2 = three-five times; 3 = Often). Scores on each instrument were standardized and then summed to create a composite social understanding variable. Raw scores (summed) ranged from 0 to 135 ($M = 47.513$, SD $= 27.985$).

Temperament

Parents rated their children's temperament using the Early Childhood Behavior Questionnaire (ECBQ; Putnam et al., 2006) or the Toddler Behavior Assessment Questionnaire (TBAQ; Goldsmith, 1996). Of interest for the current study were the subscales of fearfulness and shyness from the ECBQ and the social fear subscale from the TBAQ. On the ECBQ parents rated 11 fear-related behaviors and 12 shyness-related behaviors on a 7-point scale according to how often parents had observed them during the previous 2 weeks (1 = Never; 2 = Very rarely; 3 = Less than half the time; 4 = About half the time; 5 = More than half

the time; 6 = Almost always; 7 = Always). Children received an average fearfulness score of 1–7 ($M = 2.545$, SD = 0.87) and an average shyness score of 1–7 ($M = 3.52$, SD = 0.99). We standardized and summed the fearfulness and shyness scores to create a composite shy-fearful score. On the TBAQ parents used the same 7-point scale to rate 11 social fear-related behaviors in their children. Children received an average social fear score of 1–7 ($M = 3.59$, SD = 1.01) which was standardized for analyses.

Parent-reported Prosocial Behavior and Socialization Practices

Parents completed the prosocial behavior subscale of the Goodman Strengths and Difficulties questionnaire (SDQ; Goodman, 1997), which contains five items (e.g., "kind to younger children;" "considerate of other people's feelings") rated on a 3-point scale (0 = Not True; 1 = Sometimes True; 2 = Certainly True); children received a total score of 0–10 ($M = 6.24$, SD = 1.87). Parents also completed a questionnaire (Prosocial Behavior Questionnaire; PBQ) developed to assess their socialization practices related to prosocial behavior (e.g., "Ask my child to help even if I don't really need it, just for the purpose of teaching him/her about helping;" "Praise/thank my child when s/he helps me or someone else;" "Talk about my child's and other people's feelings with my child") as well as the child's demonstration of prosocial behavior at home (e.g., "Tries to help me around the house;" "Willingly shares food or toys with a parent without being asked"). Parents rated 12 socialization items and 14 child prosocial behavior items on a 5-point Likert scale (0 = Not at all; 1 = Once or twice; 2 = Sometimes, a few times a month; 3 = Often, a few times a week; 4 = All the time, everyday), yielding total scores ranging from 0–48 ($M = 36.83$, SD = 8.00) and 0–56 ($M = 36.18$, SD = 7.80), respectively. Scores for children's prosocial behavior from the SDQ and from the PBQ child prosocial behavior subscale were standardized and summed to produce a total parent-reported prosocial behavior variable.

Results

The primary goal of the current study was to examine relations between individual differences in toddlers' social understanding and prosocial behavior, as well as parent socialization and several dimensions of temperament. To examine predictors of parent-reported prosocial behavior, we calculated partial correlations, controlling for age and gender (where appropriate), among social understanding, parent socialization, temperament, and parent-reported prosocial behavior. To examine predictors of the observed measures of prosocial behavior from the lab tasks, we calculated one-way ANCOVAs, controlling for age and gender (where appropriate), with the categorical helping/sharing variable as the independent variable, and temperament, social understanding, and parent socialization as dependent variables. **Table 1** specifies which studies contributed which variables for analyses, and **Table 2** provides descriptive information for the raw (unstandardized) scores.

TABLE 2 | Descriptive statistics (mean and SD) for raw (unstandardized) data.

Study	Age (in months)	N	Social und	Temp (ECBQ): fear	Temp (ECBQ): shy	Temp (TBAQ): social fear	Parent socialization	Lab: instrum helping[a]	Lab: emp helping[a]	Lab: sharing[a]	Parent report: child prosocial[b]
1	18	33	24.78 (16.26)				35.88 (7.85)	20 (61%)	12 (38%)	28 (85%)	3.95 (0.87)
	30	33	71.42 (17.70)				38.42 (5.17)	30 (91%)	26 (79%)	30 (91%)	4.24 (0.62)
2	18	38	22.59 (9.58)				35.54 (5.26)	23 (61%)	17 (45%)		3.92 (0.94)
	30	20	78.65 (17.61)				40.57 (2.82)	18 (95%)	15 (75%)		4.14 (0.72)
3	30	45	71.10 (23.70)	2.45 (0.87)				44 (98%)	43 (96%)		
4	18	26	23.24 (10.00)			30.20 (11.31)	36.72 (6.77)			25 (96%)	3.18 (0.60)
	24	56	46.11 (20.61)			34.55 (12.10)	37.42 (6.60)			56 (100%)	3.52 (0.75)
5	18	38		2.53 (0.92)	3.58 (1.04)		33.71 (6.76)	32 (100%)	31 (97%)		
	30	36		2.67 (0.83)	3.46 (0.98)		38.19 (6.29)	33 (100%)	33 (100%)		

[a]Scores are dichotomous (helped/shared at least once vs. never helped); cell entries are counts (percentages).
[b]Scores are means (SD) of the sum of the average scores on the two parent-report measures, the ESB and PBQ (range = 0 – 6).

Preliminary Analyses

Preliminary analyses revealed, as would be expected, that age was positively correlated with social understanding [$r(229) = 0.78$, $p < 0.001$], parent socialization practices [$r(236) = 0.24$, $p < 0.001$], and parent-reported prosocial behavior [$r(193) = 0.19$, $p < 0.01$]. Furthermore, children who helped were significantly older than those who did not, for both instrumental helping [$M = 25.15$ mos. vs. $M = 19.89$ mos.; $F(1,231) = 30.017$, $p < 0.001$] and empathic helping [$M = 25.53$ mos. vs. $M = 21.06$ mos.; $F(1,231) = 32.841$, $p < 0.001$]. Because the focus of this paper is on individual differences rather than age differences, age is controlled in all analyses.

Analyses revealed few differences by gender. Girls were rated significantly higher than boys on temperamental fear [all scores standardized; $M = 0.33$ vs. $M = -0.34$; $F(1,110) = 12.640$, $p < 0.001$], as well as on the shy-fearful composite score [$M = 0.36$ vs. $M = -0.32$; $F(1,66) = 8.685$, $p < 0.01$]; girls and boys did not differ on the social fear subscale. Girls had significantly higher social understanding scores [$M = 0.22$ vs. $M = -0.18$; $F(1,227) = 9.515$, $p < 0.01$]. Girls also had significantly higher scores on parent-reported prosocial behavior [$M = 0.33$ vs. $M = -0.26$; $F(1,191) = 5.75$, $p < 0.05$] and were marginally more likely to help in empathic helping situations (85 vs. 71%; $\chi^2 = 3.693$, $p < 0.1$, $n = 233$). Because gender differences were not systematic, gender is controlled only when relevant.

Finally, to examine whether different measures of prosocial behavior were associated, we conducted a series of chi-square analyses between each of the observed categorical helping/sharing scores, and point bi-serial correlations between parent-reported prosocial behavior and the observed categorical helping/sharing scores. Results showed that empathic helping was significantly related to instrumental helping ($\chi^2 = 55.98$, $p < 0.001$, $n = 232$), sharing ($\chi^2 = 7.94$, $p < 0.01$, $n = 65$), and parent-reported prosocial behavior ($r_{pb} = 0.19$, $p < 0.05$, $n = 122$). Instrumental helping was significantly related to parent-reported prosocial behavior ($r_{pb} = 0.21$, $p < 0.05$, $n = 122$), but not sharing ($\chi^2 = 0.003$, ns, $n = 66$). Sharing and parent-reported prosocial behavior were unrelated ($r_{pb} = -0.15$, ns, $n = 144$. This variability in associations among various types of prosocial behavior is consistent with other recent findings with toddlers (e.g., Dunfield et al., 2011) and with larger conceptualizations of prosocial behavior as a multifaceted, multidimensional construct, with distinct subtypes that call on unique as well as overlapping skills, understanding, and motivations (Thompson and Newton, 2013; Dunfield, 2014; Eisenberg and Spinrad, 2014).

Associations between Social Understanding and Prosocial Behavior

The first aim was to determine associations between individual differences in social understanding and individual differences in prosocial behavior. For parent-reported prosocial behavior, the partial correlation controlling for age and gender yielded a significant association with parent-reported social understanding [$r(176) = 0.18$, $p < 0.05$]. A series of one-way ANCOVAs controlling for age was then conducted with each of the observed categorical helping/sharing variables as factors and parent-reported social understanding as the outcome. Children who helped in at least one empathic helping task, compared to those who did not help in any, had significantly higher social understanding scores [$M = 0.28$ vs. $M = -0.56$; $F(1,77) = 5.503$, $p < 0.05$]. There were no differences in social understanding for children who helped instrumentally vs. those who did not, or for children who shared vs. those who did not. Thus, early individual differences in social understanding are associated with both parent-reported prosocial behavior and observed empathic helping, but not with sharing or instrumental helping.

Does Parent Socialization or Temperament Predict Social Understanding?

The second aim was to determine whether individual differences in parent socialization or temperament were associated with differences in social understanding. Parent socialization was significantly positively correlated with toddlers' social understanding, controlling for age and gender [$r(156) = 0.28$, $p < 0.001$]. However, no significant associations were found between any of the temperament measures and social understanding, controlling for age and gender.

Does Parent Socialization or Temperament Predict Prosocial Behavior?

The third aim was to determine whether individual differences in parent socialization or in temperament were associated with individual differences in prosocial behavior. For parent-reported prosocial behavior, partial correlations controlling for age yielded significant associations with parent socialization [$r(158) = 0.42$, $p < 0.001$], but not with temperament. A series of one-way ANCOVAs controlling for age was then conducted with each of the observed categorical helping/sharing variables as factors and parent socialization and temperament as outcomes. Results revealed that parent socialization practices did not differ between children who helped or shared at least once on and those who did not; nor were there were differences in shy/fearful temperament between those children who helped or shared and those who did not. Thus, similar to the findings for social understanding, variability in parent socialization with their toddlers was associated with variability in parent-reported prosocial behavior.

Discussion

With the growing number of demonstrations of the remarkable prosociality of children in their second year of life, attention has recently focused on identifying age-related changes and potential developmental mechanisms underlying this capacity (e.g., Warneken and Tomasello, 2009; Brownell, 2013; Barragan and Dweck, 2014; Carpendale et al., 2014; Dunfield, 2014; Paulus, 2014). Much less attention has been paid to the question of individual differences: why are some young children more

likely to behave prosocially than others? One obvious possibility lies in individual differences in empathic concern for others' distress, known from several decades of study to differ among children when it first makes its appearance in the second year (e.g., Bischof-Köhler, 1991; Zahn-Waxler et al., 1992; Roth-Hanania et al., 2011). In the current study, we have looked to another potential source, individual differences in early social understanding.

Because early developments in prosocial behavior and social understanding are universal normative accomplishments, we have conceptualized the sources of individual variability in these capacities as arising from either dispositional differences or differences in rates of development. That is, even as early as the second year of life, some children may be dispositionally or temperamentally more interested in, attentive to, and reflective about others' emotions, needs, and desires than other children. Or some children may be more developmentally advanced in such abilities than others, which can occur for many reasons; here we focused on parental socialization as a potential contributor. There is evidence that differences in both disposition and rate of development relate to social understanding and prosocial behavior in childhood, as previously reviewed, but little evidence for when in development such associations arise.

In the current report, combining data from several previous studies of helping and sharing in 1- and 2-years-old, we found that individual variability in social understanding predicted both parent-reported prosocial behavior at home and whether children were likely to help an adult with emotion-based, empathic helping tasks in the lab. Social understanding did not predict children's instrumental helping in the lab. Examining potential sources of variability in both social understanding and prosocial behavior, we found no evidence that variability in either was associated with the temperamental quality of a fearful, shy or socially cautious stance, as has previously been shown in childhood (e.g., Wellman et al., 2011; Mink et al., 2014). In contrast, we did find that parental socialization of behavior linked with prosocial responding, which is likely to affect rates of development, was associated with individual differences in both social understanding and prosocial behavior. We discuss these findings in greater detail below.

Individual Differences in Social Understanding Predict Prosocial Behavior

Children's social understanding in the current study was assessed by parent report of their self-understanding (e.g., mirror self-recognition; pride) and their understanding and use of emotion and internal state words (e.g., happy, sad, hungry). Controlling for age differences in both social understanding and prosocial behavior, 18- to 30-months-old children with more advanced levels of social understanding were also more likely to help an adult who was cold or sad by bringing a blanket or favorite toy to alleviate the adult's distress, and were reported by their parents to demonstrate prosocial behavior more often in daily life.

These findings are consistent with larger conceptual frameworks in which early-emerging social understanding is hypothesized to contribute to the genesis of prosocial responding in the second year of life (e.g., Vaish et al., 2009; Brownell

et al., 2013b; Paulus, 2014). They are also consistent with the handful of other studies that have directly assessed some aspect of social understanding in toddlers (e.g., fairness; joint attention; personal pronouns; intention understanding) and concurrently assessed some form of prosocial behavior (e.g., sharing, helping, cooperation) and have reported positive associations (e.g., Ensor and Hughes, 2005; Brownell et al., 2006; Nichols et al., 2009; Sommerville et al., 2013; Kartner et al., 2014; Newton et al., 2014).

Unique to this study, we extended the association with prosocial behavior in previous laboratory-based studies to toddlers' everyday prosocial behavior in the family. This is important for demonstrating that the particular opportunities or challenges characteristic of laboratory tasks are not what drives these early-appearing associations between social understanding and prosocial behavior. Rather, variability in social understanding relates to nascent prosocial behavior in the everyday life of the home as well. As previous scholars have argued, family contexts are not only the primary setting in which prosocial behavior arises, but are also distinctively demanding of social understanding (e.g., Dunn and Munn, 1985; Thompson and Newton, 2013).

We also found that variability in social understanding differentially predicted empathic versus instrumental helping in the lab. Notably, in our tasks, the superficial task demands are identical across instrumental and empathic helping: the child needs to bring the adult an object that the child, but not the adult, can reach which will permit the adult to achieve a goal (instrumental helping) or will alleviate the adult's negative internal state such as being cold or sad (empathic helping). What differs between the two types of scenarios is the nature of the social understanding required. For instrumental helping, the problem is immediate, concrete, and requires recognition of another's goals, an ability that infants are known to be capable of in the first year of life (Woodward, 2005). For empathic helping, however, the inferences are more complex and abstract, requiring some understanding of the links among facial and bodily expressions, subjective states, the contextual factors that give rise to particular internal states, and how particular actions can alter another's internal state. It is telling, therefore, that a general measure of social and emotional understanding as reported by parents predicted just the sort of prosocial responding that depends on more complex inferences about others' internal states. This finding also adds to the growing body of evidence that instrumental and empathic helping may derive from different underlying mechanisms (Paulus et al., 2013; Dunfield, 2014).

Does Temperament Predict Individual Variability in Social Understanding or Prosocial Behavior?

One potential source of within-age variability in social understanding and prosocial behavior is dispositional differences in children's interest in others' emotions and mental states or their motivation to intervene in them. However, in contrast to the intriguing, albeit still limited evidence for associations between a shy, socially fearful temperament and advanced social reasoning in preschool children (Wellman et al., 2011;

Mink et al., 2014), in the current study with toddlers we found little evidence for such temperament differences predicting concurrent social understanding. Previous studies have reported modest, but significant, associations with shyness or social fear rather than with more generalized fear such as fear of the dark or of loud noises. In the current data, no relations were found with either type of fearfulness. This may be because the mechanism presumed to underlie the association with ToM in older children, a more reflective stance on others' behavior and mental states (e.g., Wellman et al., 2011), does not yet hold for toddlers. Or the association found in prior studies may be unique to more advanced forms of social reasoning such as those underlying false-belief understanding. Finally, it is also possible that measurement differences could account for the lack of effects in the current study as our measure of social understanding differed not only in substance from the theory-of-mind measures used in previous studies, but it was also parent-reported rather than assessed directly with the child. Additional research will be needed to sort out these possibilities and to determine when in development dispositional differences begin to predict variability in social understanding.

As for prosocial behavior, previous work has been mixed as to the nature of the links between these dimensions of temperament and toddlers' prosocial responses. In line with prior failures to find associations between temperamental fearfulness and prosocial helping and comforting in 1- and 2-years-old (Spinrad and Stifter, 2006; Liew et al., 2011), we found no associations between either social fear or a composite measure of shyness–fearfulness and either parent-reported prosocial behavior or observed helping or sharing in toddlers. Worth noting is that the temperament measures in the current study and one of the measures of prosocial behavior were parent-reported, yet even with the potential for shared method variance, no associations were found. Given that at least three independent studies using a variety of measures for these temperament constructs and for prosocial behavior have failed to find significant associations between them, one might be tempted to conclude that such dispositional variability in very young children is not an important factor in predicting individual differences in emerging prosocial behavior. However, this very variety in measurement may itself limit the conclusions that can be drawn, and points to the need for more systematic and focused research on the question. Additionally, we examined only three dimensions of temperament. It is quite possible that other temperament dimensions (e.g., effortful control; sociability) may relate more strongly to social understanding or prosocial behavior in this age group.

Socialization of Social Understanding and Prosocial Behavior

A second potential source of within-age variability in social understanding and prosocial behavior is parents' efforts to socialize young children's attention to and recognition of others' emotions, needs, and desires and their children's caring for and prosocial action on behalf of others. Here we did find associations between toddlers' social understanding and how often parents reported that they engaged in practices such as asking their toddlers to help them, using facial expressions or gestures when requesting help, and talking about the child's or others' feelings. A growing body of research has found other similar links between socialization and social understanding among toddler-aged children, with some of the studies longitudinal and the effects putatively causal (e.g., Dunn and Munn, 1985; Dunn et al., 1991; Symons et al., 2006; Taumoepeau and Ruffman, 2006; Ensor and Hughes, 2008; Brownell et al., 2013c). Many of these have examined the role of parental talk about emotions and mental states, which is likely to promote young children's perspective taking and consideration of others' needs. Although our measure of parent socialization also included items that refer to parents' talk about feelings, most of the items refer to parents' encouragement of prosocial behavior. Interestingly, then, our findings suggest the possibility that, for very young children at least, engagement in prosocial actions within the family may be another route through which perspective taking and appreciation of others' needs and desires could arise.

We also found associations between parents' socialization of prosocial behavior and their report of children's prosocial behavior in the family. This suggests that parents' routine encouragement of their toddlers' participation in prosocial action in the context of everyday household routines and activities promotes young children's everyday prosocial responding. Recent observations of parents and toddlers at home (Dahl, 2015) have shown that 11- to 25-months-old children are, indeed, frequently encouraged and supported by other family members in everyday helping, and that toddlers participate in a wide range of activities that are especially geared toward assisting others with immediate goals (e.g., cleaning up; sweeping; handing an object that someone else needs). By including the child in routine activities with shared goals, parents may help children figure out how to assist others with their goals and may communicate more general norms and expectations about helpfulness, especially when this occurs in the context of reciprocal, responsive interactions (Ensor and Hughes, 2008; Barragan and Dweck, 2014).

Conclusion

From the current results are subject to several limitations. First, the use of parent-report measures has both advantages and disadvantages. On the one hand, parents observe their children over time and multiple contexts, providing a broader and possibly more valid assessment of their children's competence than lab observations; they are also in the best position to evaluate their own socialization goals and practices. On the other hand, questions about shared method variance and reporter bias, including social desirability, are inevitable. However, we have used multiple measures of prosocial behavior in the current study, including observed behavior, offsetting some of these concerns; and much of the research on parenting, social understanding, and prosocial behavior conducted in childhood uses parent-report (see Eisenberg et al., 2015, for a review) which provides further confidence in the current findings. Second, the cross-sectional and correlational design precludes causal claims. For example, the association between parent socialization and

social understanding could be due to the possibility that parents who are more enthusiastic about encouraging their children to help and share are also more likely to engage in emotion and mental-state talk with their children or to be more sensitive to evidence of early social understanding in their child's behavior. Furthermore, it is worth noting that other factors besides those on which we've focused in the current study undoubtedly contribute to individual differences in social understanding and prosocial behavior. These include child factors such as motivational, attentional, and regulatory capacities; parent characteristics, such

as warmth and reciprocity, and aspects of the parent–child relationship, such as attachment security. Moreover, these are likely to interact with one another and with the constructs examined in this paper to influence early individual differences.

In sum, the current study has shown that individual variability in early social understanding is associated with variability in several different measures of prosocial behavior, and that variability in both constructs is more strongly and consistently associated with differences in parents' socialization of prosocial responding than with dispositional differences among children.

References

Barragan, R., and Dweck, C. (2014). Rethinking natural altruism: simple reciprocal interactions trigger children's benevolence. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17071–17074. doi: 10.1073/pnas.1419408111

Bischof-Köhler, D. (1991). "The development of empathy in infants," in *Infant Development: Perspectives from German-Speaking Countries*, eds M. E. Lamb and H. Keller (Hillsdale, NJ: Erlbaum), 245–273.

Bretherton, I., and Beeghly, M. (1982). Talking about internal states: the acquisition of an explicit theory of mind. *Dev. Psychol.* 18, 906–921. doi: 10.1037/0012-1649.18.6.906

Bretherton, I., Fritz, J., Zahn-Waxler, C., and Ridgeway, D. (1986). Learning to talk about emotions: a functionalist perspective. *Child Dev.* 57, 529–548. doi: 10.2307/1130334

Brownell, C., Ramani, G., and Zerwas, S. (2006). Becoming a social partner with peers: cooperation and social understanding in one- and two-year-olds. *Child Dev.* 77, 804–821. doi: 10.1111/j.1467-8624.2006.t01-1-.x-i1

Brownell, C. A. (2013). Early development of prosocial behavior: current perspectives. *Infancy* 18, 1–9. doi: 10.1111/infa.12004

Brownell, C. A., Iesue, S. S., Nichols, S. R., and Svetlova, M. (2013a). Mine or yours? development of sharing in toddlers in relation to ownership understanding. *Child Dev.* 84, 906–920. doi: 10.1111/cdev.12009

Brownell, C. A., Nichols, S. R., and Svetlova, M. (2013b). "Converging developments in prosocial behavior and self-other understanding in the second year of life: the second social-cognitive revolution," in *Navigating the Social World: What Infants, Children, and Other Species can Teach Us*, eds M. Banaji and S. Gelman (New York, NY: Oxford University Press), 385–390.

Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., and Drummond, J. (2013c). Socialization of early prosocial behavior: parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy* 18, 91–119. doi: 10.1111/j.1532-7078.2012.00125.x

Carpendale, J., Kettner, V., and Audet, K. (2014). On the nature of toddlers' helping: helping or interest in others' activity? *Soc. Dev.* 24, 357–366. doi: 10.1111/sode.12094

Carpendale, J., and Lewis, C. (2004). Constructing an understanding of mind: the development of children's social understanding within social interaction. *Behav. Brain Sci.* 27, 79–151. doi: 10.1017/S0140525X04000032

Crockenberg, S., and Litman, C. (1990). Autonomy as competence in 2-year-olds: maternal correlates of child defiance, compliance, and self-assertion. *Dev. Psychol.* 26, 961–971. doi: 10.1037/0012-1649.26.6.961

Dahl, A. (2015). The developing social context of infant helping in two U.S. samples. *Child Dev.* doi: 10.1111/cdev.12361 [Epub ahead of print].

Denham, S., Bassett, H., and Wyatt, T. (2015). "The socialization of emotional competence," in *Handbook of Socialization: Theory and Research*, 2nd Edn, eds J. Grusec and P. Hastings (New York, NY: Guilford Press), 590–613.

Denham, S., and Kochanoff, A. T. (2002). Parental contributions to preschoolers' understanding of emotion. *Marriage Fam. Rev.* 34, 311–343. doi: 10.1300/J002v34n03_06

Denham, S., Zoller, D., and Couchoud, E. (1994). Socialization of preschoolers' emotion understanding. *Dev. Psychol.* 30, 928–936. doi: 10.1037/0012-1649.30.6.928

Drummond, J., Hammond, S. I., Waugh, W. E., Satlof-Bedrick, E., and Brownell, C. A. (2014a). Helping the one you hurt: toddlers' emotional responses and

prosocial behavior after harming another. *Paper presented at the meeting of the International Conference for Infant Studies*, Berlin.

Drummond, J., Paul, E. F., Waugh, W. E., Hammond, S. I., and Brownell, C. A. (2014b). Here, there and everywhere: emotion and mental state talk in different social contexts predicts empathic helping in toddlers. *Front. Psychol.* 5:361. doi: 10.3389/fpsyg.2014.00361

Dunfield, K. (2014). A construct divided: prosocial behavior as helping, sharing, and comforting subtypes. *Front. Psychol.* 5:958. doi: 10.3389/fpsyg.2014.00958

Dunfield, K., Kuhlmeier, V. A., O'Connell, L., and Kelley, E. (2011). Examining the diversity of prosocial behavior: helping, sharing, and comforting in infancy. *Infancy* 16, 227–247. doi: 10.1111/j.1532-7078.2010.00041.x

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., and Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Dev.* 62, 1352–1356. doi: 10.2307/1130811

Dunn, J., and Munn, P. (1985). Becoming a family member: family conflict and the development of social understanding in the second year. *Child Dev.* 56, 480–492. doi: 10.2307/1129735

Eisenberg, N., Cumberland, A., and Spinrad, T. L. (1998). Parental socialization of emotion. *Psychol. Inquiry* 9, 241–273. doi: 10.1207/s15327965pli0904_1

Eisenberg, N., and Spinrad, T. (2014). "Multidimensionality of prosocial behavior: rethinking the conceptualization and development of prosocial behavior," in *Prosocial Development: A Multidimensional Approach*, eds L. Padilla-Walker and G. Carlo (New York, NY: Oxford University Press), 17–42.

Eisenberg, N., Spinrad, T., and Knafo-Noam, A. (2015). "Prosocial development," in *Handbook of Child Psychology and Developmental Science*, Vol. 3, *Socioemotional Processes*, eds R. Lerner and M. Lamb (Hoboken, NJ: Wiley), 610–656.

Ensor, R., and Hughes, C. (2005). More than talk: relations between emotion understanding and positive behavior in toddlers. *Br. J. Dev. Psychol.* 23, 343–363. doi: 10.1348/026151005X26291

Ensor, R., and Hughes, C. (2008). Content or connectedness? Mother-child talk and early social understanding. *Child Dev.* 79, 201–216. doi: 10.1111/j.1467-8624.2007.01120.x

Garner, P. W., Dunsmore, J. C., and Southam-Gerrow, M. (2008). Mother-child conversations about emotions: linkages to child aggression and prosocial behavior. *Soc. Dev.* 17, 259–277. doi: 10.1111/j.1467-9507.2007.00424.x

Goldsmith, H. H. (1996). Studying temperament via construction of the toddler behavior assessment questionnaire. *Child Dev.* 67, 218–235. doi: 10.2307/1131697

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *J. Child Psychol. Psychiatry* 38, 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Grusec, J. E. (1991). Socializing concern for others in the home. *Dev. Psychol.* 27, 338–342. doi: 10.1037/0012-1649.27.2.338

Grusec, J. E., and Hastings, P. D. (2007). *Handbook of socialization*. New York, NY: Guilford Press.

Hammond, S. I., and Carpendale, J. I. M. (2015). Helping children help: the relation between maternal scaffolding and children's early help. *Soc. Dev.* 24, 367–383. doi: 10.1111/sode.12104

Hammond, S. I., Müller, U., Carpendale, J. I. M., Bibok, M. B., and Liebermann-Finestone, D. (2012). The effects of parental scaffolding on preschoolers' executive function. *Dev. Psychol.* 48, 271–281. doi: 10.1037/a0025519

Hastings, P., Miller, J., and Troxel, N. (2015). "Making good: the socialization of children's prosocial development," in *Handbook of Socialization: Theory and Research*, 2nd Edn, eds J. Grusec and P. Hastings (New York, NY: Guilford Press), 637–666.

Kartner, J., Schuhmacher, N., and Collard, J. (2014). Socio-cognitive influences on the domain-specificity of prosocial behavior in the second year. *Infant Behav. Dev.* 37, 665–675. doi: 10.1016/j.infbeh.2014.08.004

LaBounty, J., Wellman, H. M., Olson, S., Lagattuta, K., and Liu, D. (2008). Mothers' and fathers' use of internal state talk with their young children. *Soc. Dev.* 17, 757–775. doi: 10.1111/j.1467-9507.2007.00450.x

Lagattuta, K. H., and Wellman, H. M. (2002). Differences in early parent-child conversations about negative versus positive emotions: implications for the development of psychological understanding. *Dev. Psychol.* 38, 564–580. doi: 10.1037/0012-1649.38.4.564

Liew, J., Eisenberg, N., Spinrad, T. L., Eggum, N. D., Haugen, R. G., Kupfer, A., et al. (2011). Physiological regulation and fearfulness as predictors of young children's empathy-related reactions. *Soc. Dev.* 20, 111–134. doi: 10.1111/j.1467-9507.2010.00575.x

Meins, E., and Fernyhough, C. (1999). Linguistic acquisitional style and mentalising development: the role of maternal mind-mindedness. *Cogn. Dev.* 14, 363–380. doi: 10.1016/S0885-2014(99)00010-6

Mink, D., Henning, A., and Aschersleben, G. (2014). Infant shy temperament predicts preschoolers theory of mind. *Infant Behav. Dev.* 37, 66–75. doi: 10.1016/j.infbeh.2013.12.001

Moore, C., Bosacki, S. L., and Macgillivray, S. (2011). Theory of mind and social interest in zero-acquaintance play situations. *Child Dev.* 82, 1163–1172. doi: 10.1111/j.1467-8624.2011.01602.x

Newton, E., Goodman, M., and Thompson, R. (2014). Why do some toddlers help a stranger? Origins of individual differences in prosocial behavior. *Infancy* 19, 214–226. doi: 10.1111/infa.12043

Nichols, S. R., Svetlova, M., and Brownell, C. A. (2009). The role of social understanding and empathic disposition in young children's responsiveness to distress in parents and peers. *Cogn. Brain Behav.* 13, 449–478.

Over, H., and Carpenter, M. (2009). Eighteen-month-old infants show increased helping following priming with affiliation. *Psychol. Sci.* 20, 1189–1193. doi: 10.1111/j.1467-9280.2009.02419.x

Paulus, M. (2014). The emergence of prosocial behavior: why do infants and toddlers help, comfort, and share? *Child Dev. Perspect.* 8, 77–81. doi: 10.1111/cdep.12066

Paulus, M., Kühn-Popp, N., Licata, M., Sodian, B., and Meinhardt, J. (2013). Neural correlates of prosocial behavior in infancy: different neurophysiological mechanisms support the emergence of helping and comforting. *Neuroimage* 66, 522–530. doi: 10.1016/j.neuroimage.2012.10.041

Pettygrove, D. M., Hammond, S. I., Karahuta, E. L., Waugh, W. E., and Brownell, C. A. (2013). From cleaning up to helping out: parental socialization and children;s early prosocial behavior. *Infant Behav. Dev.* 36, 843–846. doi: 10.1016/j.infbeh.2013.09.005

Putnam, S. P., Gartstein, M. A., and Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: the early childhood behavior questionnaire. *Infant Behav. Dev.* 29, 386–401. doi: 10.1016/j.infbeh.2006.01.004

Roth-Hanania, R., Davidov, M., and Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: early signs of concern for others. *Infant Behav. Dev.* 34, 447–458. doi: 10.1016/j.infbeh.2011.04.007

Ruffman, T., Slade, L., and Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Dev.* 73, 734–751. doi: 10.1111/1467-8624.00435

Shatz, M., Wellman, H. M., and Silber, S. (1983). The acquisition of mental verbs: a systematic investigation of the first reference to mental state. *Cognition* 14, 301–321. doi: 10.1016/0010-0277(83)90008-2

Sommerville, J. A., Schmidt, M. F. H., Yun, J., and Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy* 18, 40–66. doi: 10.1111/j.1532-7078.2012.00129.x

Spinrad, T. L., and Stifter, C. A. (2006). Toddlers' empathy-related responding to distress: predictions from negative emotionality and maternal behavior in infancy. *Infancy* 10, 97–121. doi: 10.1207/s15327078in1002_1

Stipek, D. J., Gralinski, J. H., and Kopp, C. B. (1990). Self-concept development in the toddler years. *Dev. Psychol.* 26, 972–977. doi: 10.1037/0012-1649.26.6.972

Svetlova, M., Nichols, S. R., and Brownell, C. A. (2010). Toddlers' prosocial behavior: from instrumental to empathic to altruistic helping. *Child Dev.* 81, 1814–1827. doi: 10.1111/j.1467-8624.2010.01512.x

Symons, D., Fossum, K., and Collins, T. (2006). A longitudinal study of belief and desire discourse during mother-child play and later false belief understanding. *Soc. Dev.* 15, 676–691. doi: 10.1111/j.1467-9507.2006.00364.x

Taumoepeau, M., and Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Dev.* 77, 465–481. doi: 10.1111/j.1467-8624.2006.00882.x

Thompson, R. A., and Newton, E. K. (2013). Baby altruists? examining the complexity of prosocial motivation in young children. *Infancy* 18, 120–133. doi: 10.1111/j.1532-7078.2012.00139.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–735. doi: 10.1017/S0140525X05000129

Vaish, A., Carpenter, M., and Tomasello, M. (2009). Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers. *Dev. Psychol.* 45, 534–543. doi: 10.1037/a0014322

Vaish, A., and Warneken, F. (2012). "Social-cognitive contributors to young children's empathic and prosocial behavior," in *Empathy*: *From Bench to Bedside*, ed. J. Decety (Cambridge, MA: MIT Press), 131–146.

van der Mark, I. L., van IJzendoorn, M. H., and Bakermans-Kranenburg, M. (2002). Development of empathy in girls during the second year of life: associations with parenting, attachment, and temperament. *Soc. Dev.* 11, 451–468. doi: 10.1111/1467-9507.00210

Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Dev.* 77, 640–663. doi: 10.1111/j.1467-8624.2006.00895.x

Warneken, F., and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science* 311, 1301–1303. doi: 10.1126/science.1121448

Warneken, F., and Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy* 11, 271–294. doi: 10.1111/j.1532-7078.2007.tb00227.x

Warneken, F., and Tomasello, M. (2009). The roots of human altruism. *Br. J. Psychol.* 100, 455–471. doi: 10.1348/000712608X379061

Waugh, W., Brownell, C., and Pollock, B. (2015). Early socialization of prosocial behavior: how parents encourage toddlers' helping in an everyday household task. *Infant Behav. Dev.* 39C, 1–10. doi: 10.1016/j.infbeh.2014.12.010 [Epub ahead of print].

Waugh, W., Satlof-Bedrick, E., Drummond, J., Marchitelli, A., and Brownell, C. (2013). Emotion priming influences prosocial behavior in toddlers. *Paper presented at the Biennial Meeting of the Society for Research in Child Development*, Seattle, WA.

Wellman, H. M., Lane, J. D., LaBounty, J., and Olson, S. L. (2011). Observant, nonaggressive temperament predicts theory-of-mind development. *Dev. Sci.* 14, 319–326. doi: 10.1111/j.1467-7687.2010.00977.x

Woodward, A. L. (2005). "The infant origins of intentional understanding," in *Advances in Child Development and Behavior*, Vol. 33, ed. R. V. Kail (San Diego, CA: Elsevier Academic Press), 229–262.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., and Chapman, M. (1992). Development of concern for others. *Dev. Psychol.* 28, 126–136. doi: 10.1037/0012-1649.28.1.126

Cumulative biomedical risk and social cognition in the second year of life: prediction and moderation by responsive parenting

Mark Wade[1]*, Sheri Madigan[1], Emis Akbari[2] and Jennifer M. Jenkins[1]

[1] Department of Applied Psychology and Human Development, University of Toronto, Toronto, ON, Canada, [2] Atkinson Centre for Society and Child Development, Fraser Mustard Institute for Human Development, University of Toronto, Toronto, ON, Canada

At 18 months, children show marked variability in their social-cognitive skill development, and the preponderance of past research has focused on constitutional and contextual factors in explaining this variability. Extending this literature, the current study examined whether cumulative biomedical risk represents another source of variability in social cognition at 18 months. Further, we aimed to determine whether responsive parenting moderated the association between biomedical risk and social cognition. A prospective community birth cohort of 501 families was recruited at the time of the child's birth. Cumulative biomedical risk was measured as a count of 10 prenatal/birth complications. Families were followed up at 18 months, at which point social-cognitive data was collected on children's joint attention, empathy, cooperation, and self-recognition using previously validated tasks. Concurrently, responsive maternal behavior was assessed through observational coding of mother–child interactions. After controlling for covariates (e.g., age, gender, child language, socioeconomic variables), both cumulative biomedical risk and maternal responsivity significantly predicted social cognition at 18 months. Above and beyond these main effects, there was also a significant interaction between biomedical risk and maternal responsivity, such that higher biomedical risk was significantly associated with compromised social cognition at 18 months, but only in children who experienced low levels of responsive parenting. For those receiving comparatively high levels of responsive parenting, there was no apparent effect of biomedical risk on social cognition. This study shows that cumulative biomedical risk may be one source of inter-individual variability in social cognition at 18 months. However, positive postnatal experiences, particularly high levels of responsive parenting, may protect children against the deleterious effects of these risks on social cognition.

Keywords: social cognition, biomedical risk, parenting, maternal responsivity, risk-resilience

Introduction

Social cognition is the set of cognitive processes related to social understanding and behavior. The capacity to understand human actions in terms of the psychological states that motivate behavior

is a fundamental component of social cognition. While social cognition is broadly defined and includes a number of cognitive processes, it is generally well accepted that by the second year of life children evince many basic social-cognitive competencies, including an understanding of others' goals (Csibra et al., 2003), intentions (Behne et al., 2005), desires (Repacholi and Gopnik, 1997), emotions (Moses et al., 2001), and perhaps even beliefs (Buttelmann et al., 2009). The ability to understand others' mental states manifests itself in a number of overt behaviors in the second year of life, many of which are used to index early social cognition. For instance, by 18 months children engage in regular bouts of joint attention (Tomasello et al., 2005; Tomasello and Carpenter, 2007), empathy (Roth-Hanania et al., 2011), cooperation (Brownell et al., 2006; Warneken et al., 2006; Warneken and Tomasello, 2007), and self-recognition (Nielsen and Dissanayake, 2004; Brownell et al., 2007). These social-cognitive skills rely on the capacity to differentiate self from other (Asendorpf et al., 1996; Lewis, 2003), and it has been suggested that children's emergent aptitude for understanding intentions may play a critical role in their ability to engage successfully in these behaviors (Moore, 2007; Knoblich and Sebanz, 2008).

Although social cognition develops progressively over childhood (Gergely and Csibra, 2003; San Juan and Astington, 2012; Thoermer et al., 2012), there are important individual differences in early social cognition that have a bearing on later skills such as theory of mind (Legerstee, 2005; Aschersleben et al., 2008; Wellman et al., 2008). This variability in social reasoning can also be observed in adolescence (Moriguchi et al., 2007; Dumontheil et al., 2010). Longitudinal studies show that individual differences in social cognition are quite stable (Pons and Harris, 2005) and are related to multiple developmental outcomes (Frischen et al., 2007; Fiske and Taylor, 2013). For instance, theory of mind ability has been linked to children's academic achievement (Blair and Razza, 2007), behavioral problems (Hughes and Ensor, 2006), and social competence (Razza and Blair, 2009). Accordingly, it is important to identify sources of variability in early social cognition, which may exert downstream effects on multiple domains of functioning.

To date, the preponderance of literature on predictors of social cognition has focused on contextual factors such as family processes and socioeconomic variables. For instance, Dunn et al. (1991) have shown that mothers' mental state discourse and family socioeconomic status (SES) at 33 months are associated with emotion understanding at 40 months. The effect of socioeconomic factors on individual differences in theory of mind has been replicated in numerous investigations (Holmes et al., 1996; Shatz et al., 2003). Moreover, the effect of parenting behavior on social cognition is one of the most robust findings in the literature on social cognition (Pears and Moses, 2003; de Rosnay and Hughes, 2006; Ruffman et al., 2006). Also relevant are child-level factors such as gender, with females demonstrating overall better social cognition than males (Dunn et al., 1991). One of the strongest factors associated with social cognition is language ability (Astington and Jenkins, 1999; Cutting and Dunn, 1999; de Rosnay and Harris, 2002; Pons et al., 2003), which may play both a communicational and representational role in social cognition (see Dunn and Brophy, 2005). Thus, there appears to be a range of known environmental and child-specific factors that contribute to individual differences in social cognition across childhood.

Importantly, much of the existing literature has focused on predictors of social cognition in preschool and school-age children. Relatively less is known about the factors associated with social cognition at earlier stages of development. However, recent studies suggest that, as early as the second year of life, there may be multiple influences on social cognition, such as cumulative social disadvantage, maternal sensitivity, and language ability (Wade et al., 2014c) as well as oxytocin genetic variability (Wade et al., 2014b) and pregnancy hypertension (Wade and Jenkins, 2014). These results are consistent with the manifold biopsychosocial correlates of social cognition observed in preschool children. However, across all studies there remains a substantial proportion of unexplained residual variance, suggesting the presence of currently unspecified influences on social cognition. The goal of the current study was to examine whether early biomedical risk, or the occurrence of combined pre- and perinatal complications, represented another source of variability in social cognition in the second year of life. Further, supposing that such a relationship exists, and consistent with the known effects of contextual factors on social-cognitive development, we aimed to determine whether positive postnatal interpersonal experiences with caregivers (i.e., responsive parenting) protected children against these potentially adverse biomedical risks.

Specific biomedical risk factors for early social cognition have been vastly understudied. In one recent study, Wade and Jenkins (2014) demonstrated that pregnancy hypertension is associated with lower social cognition at 18 months, as well as theory of mind ability in the preschool period. Another recent study showed that birth weight was positively associated with theory of mind at age 4.5 in a typically developing sample (Wade et al., 2014a). Together, these studies provide preliminary evidence that pre- and perinatal factors may be involved in a mechanism through which early fetal stress impinges on healthy brain development that supports social cognition. Aside from these findings, however, little is known about the role of biomedical factors on social cognition in the second year of life.

Indirect evidence for a role of early medical complications on social cognition comes from research showing that such factors are related to the risk of neurodevelopmental and psychiatric disorders characterized by deficits in social cognition. For instance, a comprehensive review by Kolevzon et al. (2007) revealed that the most prominent obstetric complications associated with risk for autism spectrum disorder (ASD) included birth weight, gestational age, as well as intrapartum hypoxia. Obstetrical complications have also been linked to the risk for schizophrenia (Geddes and Lawrie, 1995; Verdoux et al., 1997), eating disorders (Cnattingius et al., 1999), early onset affective disorders (Guth et al., 1993), substance abuse (Sydsjö, 2011), attention-deficit hyperactivity disorder (Milberger et al., 1997; Bhutta et al., 2002), and conduct, oppositional, and internalizing problems (Cohen et al., 1989). In a prospective follow-up study, Buka et al. (1993) suggested that fetal hypoxia was the common underlying mechanism and was the strongest predictor of later cognitive and

psychiatric difficulties. Several maternal pathologies during pregnancy have been linked to perinatal hypoxia–ischemia, such as infections, diabetes, hypertension, and thyroid problems (Shah, 2001; Kurinczuk et al., 2010; Teramo, 2010; Stanek, 2013). Thus, it is conceivable that these biomedical factors increase the risk of hypoxic-ischemic events which compromise development in key social-cognitive domains that typify neurodevelopmental and psychiatric conditions.

Two important points deserve consideration here. The first is that early biomedical complications likely produce a continuum of postnatal biopsychosocial-health variability, rather than just the extremes of problems (Pasamanick and Knobloch, 1961). This means that we should expect to observe individual differences in discrete social, cognitive, and emotional phenotypes that characterize neurodevelopmental and psychiatric conditions as a function of biomedical risk. Second, the existing research is limited in differentiating between the effect of different types of prenatal/birth complications on developmental outcomes (Allen et al., 1998). Indeed, there are a variety of biomedical complications that can occur during the pre-, peri-, and neonatal period, including those related to maternal physical health (e.g., endocrine/inflammatory diseases), intrapartum events (e.g., physical trauma), perinatal problems (e.g., low birth weight, prematurity), and immediate postpartum factors (e.g., anoxia or hematological problems demanding use of specialized care). However, it may be difficult to ascertain the effect of each individual risk on children's outcomes, particularly in epidemiological samples where the prevalence of certain conditions may be too low to provide powerful estimates and the measurement is not sufficiently detailed to effectively partition risks. As a result, one approach that may be useful is the *cumulative risk* approach. The overarching idea behind cumulative risk measures is that, rather than a single and specific risk, it is the aggregation of multiple risks that compromises development (Dong et al., 2004; Flouri and Kallis, 2007; Burchinal et al., 2008). Indeed, it has been repeatedly demonstrated that cumulative risk indices are more stable than individual risk measures (Burchinal et al., 2000), and explain more variance in child outcomes than risks examined in isolation (Deater-Deckard et al., 1998; Atzaba-Poria et al., 2004; Flouri and Kallis, 2007; Evans et al., 2013).

While the cumulative risk approach has been applied widely within the psychosocial domain, its application to prenatal/birth risks is far less common. Nonetheless, existing research indicates that the accumulation of biomedical risks in the pre- and perinatal period is detrimental to children's socioemotional, intellectual, and motor functioning (Laucht et al., 1997), as well as their visual memory (Levy-Shiff et al., 1994) and attentional control (Carmody et al., 2006). However, these studies have generally assessed the effect of medical complications in children born preterm, which represents a group of already at-risk children who may be particularly vulnerable to negative outcomes. The effect of biomedical risk (i.e., prenatal/birth complications) on social cognition in the general community remains unexplored. Further, no study has examined how enriched postnatal experiences may protect against early biomedical risk on social cognition.

Parental inputs are believed to foster social cognition owing to their role in providing children with the linguistic,

representational, and reflective material needed to understand others' minds (Fernyhough, 2008). Further, it has been demonstrated that positive experiences with caregivers exert a *protective* influence on children (Rutter, 1987; Brody et al., 2002; Burchinal et al., 2006). Protective in this regard does not mean avoiding risk, but persevering in the face of it. These 'moderation' models are typically examined by determining whether the association between two variables depends on the level of a third variable, with the risk variable (e.g., biological risk) being less predictive of the outcome when the presumed protective factor is present. Surprisingly, there is little existing research on parenting as a protective factor in regard to the development of social-cognitive capacities, or as a moderator of the association between biological risk and children's outcomes in general. The limited research to date, however, does suggest that certain aspects of parenting may buffer children against early biomedical risk. For example, Laucht et al. (2001) found that responsive parenting moderated the effect of birth weight on school-aged children's hyperkinetic and internalizing problems, and Voigt et al. (2013) showed that the effect of neonatal distress on children's negative affectivity at 12 months depended on the level of parenting stress, with lower levels of stress protecting against neonatal problems. Finally, another interesting study examining children's executive functioning – a neurocognitive skill that is developmentally linked to social cognition – showed that the effect of neurobiological risk (i.e., direct measurement from neonatal medical records, e.g., need for oxygen/ventilation) on executive functioning was most prominent in socioeconomically disadvatanged children (Ford et al., 2011). Thus, to build on this literature, and in line with risk-resiliency models of development (Luthar et al., 2000; Masten et al., 2009; Jenkins et al., in press), the current study aimed to determine whether, given an association between cumulative biomedical risk and social cognition, responsive parenting moderated this association. Specifically, it was hypothesized that higher levels of biomedical risk would be associated with lower social cognition at 18 months; however, if children received high levels of responsive parenting, the effect of biomedical risk on social cognition would be attenuated.

Materials and Methods

Participants

Participants came from the intensive sample of the Kids, Families, Places Study (iKFP; http://kfp.oise.utoronto.ca/). All women giving birth in Toronto and Hamilton, Ontario, between April 2006 and September 2007 were considered for participation. Families were recruited through a program called *Healthy Babies Healthy Children*. Parents of all registered newborns were contacted within several days of the child's birth. Inclusion criteria for the iKFP study included the presence of an English-speaking mother, a newborn >1500 g, at least two children who are <4 years, and families agreeing to be filmed in the home. Of those contacted, 34% of families agreed to take part in the study. Reasons for non-enlistment included refusals and an inability to contact families from public health's information. The University of Toronto Research Ethics Board approved

all procedures for this investigation, including informed consent.

We compared our sample ($N = 501$) with the general population of Toronto and Hamilton using 2006 Census Data, limiting the census to women between 20 and 50 years and having at least one child. Families were compared based upon immigrant status, number of persons in the home, family structure, maternal personal income, and educational level. Based on these comparisons, iKFP was similar to the general population on family size ($M = 4.52$, SD = 1.01 vs. $M = 4.13$, SD = 1.22) and personal income (C$30,000–39,999 vs. census population mean = C$30,504.16, SD = C$37,808.12). Since our sample was recruited shortly after childbirth, there were predictably fewer non-intact families than in the general population (5% vs. 16.8% lone-parent families; 4.3% vs. 10.3% stepfamilies). The ratio of Canadian-born to immigrants was somewhat higher in the iKFP sample (57.7% vs. 47.6%), likely due to the language requirement for participation. Also, more study mothers had earned a bachelor's degree or higher in the iKFP sample (53.3% vs. 30.6%). The sample was ethnically and socio-demographically diverse (see **Table 1**).

At Time 1 (T1; $M_{age} = 2.0$ months; SD = 1.06), 501 families were enlisted in the study. Due to sample attrition, 397 (79.2%) families were followed up at Time 2 (T2; $M_{age} = 1.60$ years; SD = 0.16). Attrition analysis showed that dropout, similar to other longitudinal studies, was related to higher levels of social risk: maternal depression at T1, χ^2 ($df = 1$) = 7.2, $p = 0.01$, being in a non-intact family, χ^2 ($df = 1$) = 11.1, $p = 0.002$, immigrant status, χ^2 ($df = 1$) = 13.5, $p < 0.001$, teenage parenthood, χ^2 ($df = 1$) = 6.7, $p = 0.02$, maternal education <high school, χ^2 ($df = 1$) = 10.5, $p = 0.002$, and family income < $20,000, χ^2 ($df = 1$) = 7.1, $p = 0.01$. Of the 397 children remaining at T2, no social-cognitive data were available for 24 children due to non-compliance, lack of visibility (e.g., child went off camera), parent intrusion (e.g., directing child), non-administration due to family constraints (e.g., time limitations) or tester administration error (e.g., not following the standardized protocol). This resulted in a final sample of 373 children providing data on social cognition.

TABLE 1 | Demographic characteristics of the sample at study entry ($N = 501$).

Measure	N	% of sample
Ethnicity of mothers	501	100.0
European/Caucasian	283	56.5
South Asian	73	14.6
East Asian	60	12.0
Black	46	9.2
Other	39	7.7
Teenage mother	31	6.2
Single parent family	32	6.4
Immigrant family (mother not Canadian-born)	233	46.5
Low income family (<$20,000)	45	9.5
Mother's years of education (<high school)	34	6.2
Mothers scoring in depressed range on CESD	71	14.4

Total sample at wave 1, N = 501.

Procedure

The study design combined the strengths of epidemiological methodology (large and diverse sample, multiple siblings, home visits) with the strength of developmental methodology (tasks developed in the laboratory, detailed microsocial observational data). At each time point, two trained interviewers visited each family's residence for approximately 2 h. Data collection included questionnaires, age-appropriate developmental tasks for target children at T2, and observational measures of mother–child interactions at T2.

Measures
Cumulative Biomedical Risk

At T1, mothers reported on their own pregnancy complications and a variety of infant birth problems. A single item was used to assess the presence/absence (0 = absent; 1 = present) of each of the following: (1) pregnancy diabetes; (2) hypertension; (3) thyroid problems (4) loss of fetal movement; (5) injury to the abdomen; (6) infant need for intensive care after birth; (7) infant need for oxygen/ventilation; and (8) infant need to be transferred to a specialized hospital. Further, two additional continuous perinatal risk factors were dichotomized based on pre-defined cut-points. These were: (9) low birth weight (<2500 g); and (10) short gestation (<37 weeks). A count of these biomedical risks was computed. The distribution of problems in the sample was as follows: 0 problems (68.0%), 1 problem (25.0%), 2 problems (4.4%), 3 problems (1.2%), 4 problems (1.2%), 5 problems (0%), and 6 problems (0.2%). No individuals reported 7–10 problems. Further, as few individuals existed in the upper tail of the distribution, we combined 4–6 problems into a category of '4 or more' problems (1.4% of the sample). Thus, this variable represented a count of the number of biomedical risks/complications on a scale from zero to '4 or more.'

Maternal Responsivity

Observational data were gathered at T2 on mother–child interactions across three 5-min tasks: (1) unstructured free play with no toys; (2) a structured cooperative building task (using Duplo blocks to build a design from a picture); and (3) reading from a wordless picture book. For all three tasks, three domains of responsivity were coded using the Parent–Child Interaction System of global ratings (PARCHISY, Deater-Deckard et al., unpublished) and the Coding of Attachment Related Parenting (CARP, Matias, 2006). *Sensitivity* (from the CARP) measured the degree to which the parent responded to the child's verbal and non-verbal signals, supported the child's autonomy, showed warmth, and demonstrated an ability to see things from the child's perspective. *Mutuality* (from the CARP) is a dyadic code and is compatible with the concept of the 'goal-corrected partnership' (Bowlby, 1982). Mutuality was indexed by reciprocity in conversation (e.g., a conversation that "goes somewhere" and is a genuine dialog), affect sharing, joint engagement in task, and open body posture. Finally, *positive control* (from the PARCHISY) captures the parents' positive means of getting the child to do something that she wanted him or her to do through the use of praise, explanations, and open ended questions. Each of these three domains – sensitivity, mutuality, and positive control – was rated

on a 7-point scale for each of the three tasks. Internal consistency of the measures was high (α = 0.85). Thus, a composite measure of 'maternal responsivity' was created by averaging the sensitivity, mutuality, and positive control scores across all three tasks. Higher scores reflected higher levels of maternal responsivity. Coders were trained to criterion and then 10% of the interactions were double-coded. Reliability was checked throughout the coding period to guard against rater drift. Inter-rater reliability was high (α = 0.94). Coders were blind to the biomedical history of the children.

Social Cognition

This was measured at T2 (18 months) by four independent observational tasks assessing children's joint attention, empathy, cooperation, and self-recognition. Each of these tasks was previously validated and widely used in laboratory studies, and we adapted these for use in the home interviews. A complete description of these tasks can be found in Supplementary Material, as well as Wade et al. (2014c). Briefly, in the joint attention task children were required to respond to an adult interviewer's bids for directing their attention (Mundy et al., 2003); in the empathy task (Kochanska et al., 1994) children were assessed for their ability to respond to the feigned distress of the interviewer; in the cooperation tasks (Warneken et al., 2006) children had to work collaboratively with the interviewer toward a shared goal; and in the self-recognition task we evaluated children's ability to recognize the objectivity of their body using the mirror-rouge paradigm (Amsterdam, 1972). Inter-rater reliabilities across tasks were good: α = 0.94 for joint attention, α = 0.82 for empathy, α = 0.86 for cooperation, and κ = 0.79 for self-recognition. Scores on these measures were submitted to a confirmatory factor analysis (CFA), consistent with their ostensible coherence as indicators of children's latent social cognition (Wade et al., 2014c). Model fit for the *social cognition* factor was excellent in accordance with Hu and Bentler's (1999) recommended cut-offs: root-mean-square-error of approximation (RMSEA) = 0.023, comparative fit index (CFI) = 0.99, and standardized root-mean-square residual (SRMR) = 0.021. Model-estimated loadings were also positive and significant at the $p < 0.001$ level for all indicators. Factor scores were saved and used as the primary outcome variable. The *social cognition* factor was normally distributed with a mean of zero.

Covariates

Based on previous studies demonstrating the association between certain socio-demographic and constitutional factors and social cognition, a number of variables were controlled for: (1) child age in years; (2) child gender (0 = male; 1 = female); (3) annual family income, assessed on a scale from 1 ('no income') to 16 ('$105,000 or more'); (4) maternal education, assessed as the total number of years of formal schooling, not including kindergarten; (5) immigrant status of the mother (i.e., 0 = immigrant; 1 = born in Canada); (6) maternal depression, assessed using the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), a widely used self-report scale that assesses depression in non-clinical populations; and (7) children's language ability, measured concurrent with social cognition (18 months) using

the MacArthur-Bates Communicative Development Inventories (CDIs; Fenson et al., 1994).

Statistical Analysis

First, all predictor and covariate variables were standardized, and the interaction term between cumulative biomedical risk and maternal responsivity was computed by multiplying the z-scores of these two variables (Preacher and Rucker, 2003). We then performed hierarchical multiple regression using MPlus 7.0. To handle variable amounts of missing data, we used full-information maximum likelihood estimation (FIML), which produces unbiased parameter estimates and SEs when data are missing at random (Enders and Bandalos, 2001). The model was fitted using the maximum likelihood with robust SEs estimator (MLR), which gives parameter estimates with SEs and a chi-square that are robust to non-normality (Yuan and Bentler, 2000). In the first step of the multiple regression analysis, the covariates were entered into the model. In the second step, the covariates plus the main effects of cumulative biomedical risk and maternal responsivity were entered into the model. Finally, in the third step, the interaction between biomedical risk and maternal responsivity was added to the variables from all previous steps in order to determine whether the interaction term predicted social cognition above and beyond covariates and main effects.

Results

Preliminary Descriptive Analysis

Table 2 presents the descriptive statistics for all study variables, including bivariate associations. Notable associations in **Table 2** include the positive relationship between social cognition and child age, female gender, family income, language ability, and maternal responsivity, as well as the negative association between social cognition and cumulative biomedical risk. Higher biomedical risk was also associated with lower socioeconomic status (family income and maternal education), as well as higher levels of maternal depression and lower levels of maternal responsivity. Maternal responsivity was associated with nearly all other study variables. A preliminary trend analysis showed that there was a significant linear association between cumulative biomedical risk and social cognition, B (SE) = −0.02 (0.01), $p = 0.047$. Neither the quadratic, B (SE) = 0.01 (0.01), $p = 0.10$, nor the cubic trend, B (SE) = −0.01 (0.01), $p = 0.22$, were significant, suggesting that as cumulative biomedical risk increases, social cognition decreases in a linear fashion (see Supplementary Figure S1 for a plot of this association). Also, Supplementary Table S1 outlines the inter-relations between individual risk variables in the cumulative risk index. This Table shows a combination of independent and inter-dependent risk variables, making the cumulative risk approach suitable (see Evans et al., 2013).

Primary Regression Analysis

We performed hierarchical multiple linear regression to examine the effect of cumulative biomedical risk, maternal responsivity,

TABLE 2 | Descriptive statistics and correlations between study variables.

	1	2	3	4	5	6	7	8	9	*M* or %	SD
(1) Child age	–									1.60	0.16
(2) Female gender	0.00	–								49.3	–
(3) Family income	0.05	0.02	–							11.9	4.06
(4) Maternal education	0.11*	−0.04	0.51***	–						15.3	2.68
(5) Immigrant status	−0.14**	−0.02	0.34***	0.13**	–					46.5	–
(6) Maternal depression	−0.06	−0.02	−0.32***	−0.25***	−0.17***	–				9.46	7.29
(7) Language ability	0.00	0.00	0.04	0.02	0.12*	0.04	–			0.00[a]	1.00
(8) Maternal responsivity	0.07	0.14**	0.34***	0.26***	0.24***	−0.16**	0.17**	–		3.51	0.79
(9) Social cognition factor	0.42***	0.14**	0.13*	0.06	0.02	−0.05	0.27***	0.24***	–	0.00[a]	0.14
(10) Biomedical risk	−0.01	−0.02	−0.18***	−0.17***	0.04	0.10*	−0.02	−0.12*	−0.10*	–[b]	–

$***p < 0.001$, $**p < 0.01$, $*p < 0.05$, $p < 0.10$.
[a]*These are either standardized scores or factor scores with a mean of zero.*
[b]*See in-text for the distribution of this variable.*

TABLE 3 | Model results for the primary multiple regression analysis.

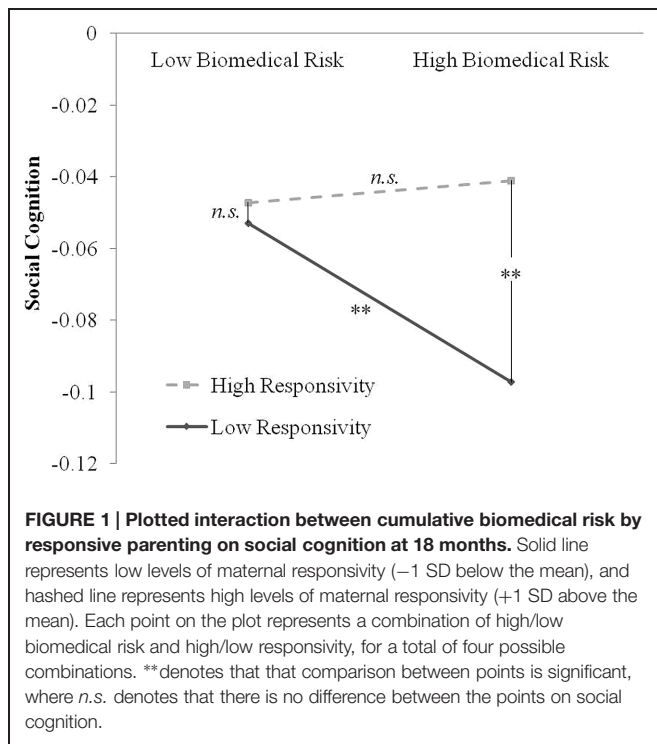|	β	SE	*p*	R^2	*F* change
Step 1:					
Female gender	0.30	0.09	0.001	0.30***	20.49***
Child age	0.44	0.05	<0.001		
Family income	0.09	0.05	0.10		
Maternal education	−0.02	0.06	0.77		
Child language ability	0.27	0.05	<0.001		
Canadian-born	0.10	0.10	0.33		
Maternal depression	−0.01	0.05	0.76		
Step 2:					
Cumulative biomedical risk	−0.12	0.06	0.043	0.32***	3.95*
Maternal responsivity	0.10	0.05	0.051		
Step 3:					
Cumulative biomedical risk*maternal responsivity	0.10	0.05	0.047	0.33***	3.29[†]

$†p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.
R^2 – Cumulative R^2.

and their interaction on social cognition. These results are presented in **Table 3**. In the first step of the model, covariates that were shown to be significant predictors of social cognition at 18 months included age, female gender, and child language ability. Family income was marginally associated with social cognition. None of the other covariates were significant predictors. This step of the model accounted for a significant 30% of the variance in social cognition. In the second step of the model, above and beyond covariates, there was, there was a significant main effect of cumulative biomedical risk and a marginally significant main effect of maternal responsivity on social cognition. This model accounted for an additional 2.1% of the variance in social cognition, or 32.1% overall. Finally, in the third step of the model, over and above covariates and main effects, the interaction between cumulative biomedical risk and maternal responsivity significantly predicted social cognition. The main effects of both biomedical risk and maternal responsivity were reduced to non-significance upon inclusion of the interaction term. This model accounted for a total of 32.8% of the variance in social cognition.

Follow-Up Analysis of Simple Slopes

To explicate the pattern of the interaction between biomedical risk and maternal responsivity, we performed an analysis of simple slopes, which tests the relationship between biomedical risk and social cognition at different levels of the moderator (Aiken and West, 1991). In the case of a continuous moderator (i.e., responsive parenting), the common approach to examine the regression relationship at high (+1 SD) and low (−1 SD) levels of the moderator (Cohen et al., 2013). The pattern of this interaction can be seen in **Figure 1**. This figure shows that, when biomedical risk is low, there was a minimal effect of responsivity on social cognition ($z = 0.38$, $p = 0.71$). Alternatively, at high levels of biomedical risk, responsivity was positively related to social cognition ($z = 2.66$, $p = 0.008$). Examining the converse associations, at low levels of responsivity, biomedical risk was strongly negatively associated with social cognition ($z = −2.70$, $p = 0.002$), while at high levels of responsivity, biomedical risk was not associated with social cognition ($z = 0.38$, $p = 0.70$).

FIGURE 1 | Plotted interaction between cumulative biomedical risk by responsive parenting on social cognition at 18 months. Solid line represents low levels of maternal responsivity (−1 SD below the mean), and hashed line represents high levels of maternal responsivity (+1 SD above the mean). Each point on the plot represents a combination of high/low biomedical risk and high/low responsivity, for a total of four possible combinations. **denotes that that comparison between points is significant, where *n.s.* denotes that there is no difference between the points on social cognition.

Discussion

The aim of the current study was to investigate the association between cumulative biomedical risk and social cognition at 18 months, and whether maternal responsivity moderated this association. It was shown that, above and beyond covariates, both maternal responsivity and cumulative biomedical risk independently predicted social cognition at 18 months. Further, consistent with study hypotheses, maternal responsivity was shown to moderate the association between biomedical risk and social cognition, with the effect of biomedical risk only apparent at low levels of maternal responsivity. Alternatively, at high levels of maternal responsivity, there was no effect of cumulative biomedical risk on social cognition. These results provide the first empirical evidence that accumulating biomedical risk factors may be one source of inter-individual variability in children's social-cognitive skills in the second year of life. Also, and consistent with risk-resiliency models of development, these findings suggest that postnatal socialization factors – specifically responsive caregiving – may protect against the impact of early biomedical risk on child outcomes.

Our finding that responsive parenting acts as a protective factor against early biomedical complications is consistent with intervention studies showing that cognitive and social outcomes of perinatally at-risk children may be fostered through training programs that build parents' cognitive and affective responsiveness (Landry et al., 1997, 2006, 2008, 2012). In general, these studies show that intervention effects on broad cognitive and socio-emotional competence operate through changes in parenting behaviors, and these effects are strongest in the most biologically at-risk children (e.g., very low birth weight, preterm).

Within the context of these intervention studies, the current findings are noteworthy for two reasons: first, they show that, in addition to individual biological insults such as low birth weight, the accumulation of early biomedical risk factors may also compromise children's emerging social-cognitive skill development, operationalized within a framework that posits underlying capacities for self-other differentiation and understanding of intentions (see also Moore, 2007; Wade et al., 2014c); second, they demonstrate that the protective role of responsive maternal behaviors is also present within a normative, epidemiological sample of children with varying degrees of biological risk. Within such a sample, the presence of individual biomedical risks are typically not powerful individual predictors of child outcomes, either because these are low frequency events, or because there are a host of identified or unidentified factors that buffer the effect of isolated risks. Rather, it may be that the accumulation of multiple biomedical risks is what creates meaningful differences in children's social cognition within the general population.

The mechanisms through which biomedical risks influence children's social cognition are presumed to involve changes in infant brain development. However, little research exists to support the idea that prenatal/birth insults specifically impact the neural regions that support social cognition in humans. The postnatal progression following such biomedical risks may shed light on the mechanisms that underlie differences reported here. Infants born with prenatal/perinatal complications are at a higher risk for postnatal complications (e.g., metabolic complications; Lubchenco and Bard, 1971; Hendderson et al., 2006). Experimental evidence from animal models demonstrates that all these factors can stimulate or precipitate neuronal death in the infant brain resulting in volume loss in particular regions within the brain (Bhutta and Anand, 2001). This is supported by findings from Peterson et al. (2000), who examined brain volume differences in 8-year-old children born with birth complications. This study demonstrated smaller volumes in the amygdala, hippocampus, basal ganglia, and cortical regions, all of which were associated with increased risk of ADHD and lower cognitive scores. Some of these regions have also been implicated in social cognition (Adolphs, 2001). Further, in a notable study by Carmody et al. (2006), cumulative medical and environmental risk was shown to be associated with lower cognitive performance in adolescence, as well distinct patterns as brain activation in temporal and parietal cortical regions. This is interesting given that social cognition, including the capacity for self-other differentiation and mental-state inference, is believed to be supported by a distributed neural network that includes temporal and parietal areas (Decety and Sommerville, 2003; Van Overwalle, 2009). By extension, it is plausible that accumulating biomedical risks are associated with social cognition by virtue of their effect on functional brain networks during *in utero* and early postnatal development. Moreover, recent studies suggest the possibility that the strongest associations between pre/perinatal characteristics and brain development may exist within the normal range (Raznahan et al., 2012; Walhovd et al., 2013). The current results show that, indeed, meaningful differences in social cognition may exist as a function of normal variation in

summative biomedical complications. Despite these interesting findings, the exact mechanism(s) connecting biomedical risk, neural development, and social cognition require future research.

Perhaps most interesting to the current study was the finding that responsive parenting moderated the association between cumulative biomedical risk and social cognition. These results are consistent with other observational studies on the protective effect of positive caregiving on children's varied behavioral and mental health outcomes (Raine et al., 1994, 1997; Landry et al., 1997; Laucht et al., 1997, 2001; Voigt et al., 2013). Schore's *regulation theory* suggests that positive parent–child interactions help promote adaptive functioning through regulation of neurobiological processes, including structural and functional neuroanatomy (Schore, 1996, 2001). Moreover, regulation theory posits a maturational process from prenatal to postnatal development, consistent with the notion that there is substantial brain development over the first 2 years of life (Knickmeyer et al., 2008). The developing brain is also very vulnerable to both environmental insult and enrichment, the latter of which may promote some the protective effects of responsive caregiving. Interestingly, recent findings from longitudinal studies show that the provision of early responsive caregiving is associated with enhanced physiological organization and resultant cognitive functioning over the first 10 years of life (Feldman et al., 2014). The precise role of responsive parenting, including the specific forms of care that foster neurobiological development and social cognition, requires further investigation. However, collaborative evidence from the fields of pediatrics, developmental psychology, and social neuroscience point to the importance of early responsive care in ameliorating the long-term sequelae of adverse pre/perinatal events on neurological and cognitive morbidity. Indeed, small variations in biological risk may create momentous gaps in children's social and cognitive development, and these effects may persist across the lifespan in the absence of interventions that target foundational inter-personal transactions with caregivers early in postnatal life (Walhovd et al., 2014).

The results of this study should be considered in light of several strengths and limitations. The strengths included the prospective, multi-method, longitudinal design, large and diverse sample, and use of detailed observational outcome data on 18-month social-cognitive measures. Inclusion of numerous sociodemographic confounding variables also adds to the robustness of the current findings. In regard to limitations, the current Canadian sample was more advantaged than the general population, and participation was restricted to children born >1500 g. These sampling factors may limit the generalizability of the results. Also, each of the 10 biomedical risks was low frequency, measured through maternal report, and typically dichotomous. Agreement between self-report and criterion-standard medical record data has been shown to be high for prenatal complications (Okura et al., 2004) and other pre/perinatal events (Lederman and Paxton, 1998; Tomeo et al., 1999). However, future studies using more comprehensive information from obstetrical records would strengthen these findings. Moreover, additional information on the timing and severity of particular prenatal conditions (e.g., diabetes, hypertension, thyroid problems), as well as the specific reasons neonatal specialized care was needed (e.g., ischemia, anoxia, hematological problems), would improve suggestions about the mechanisms at play. More extensive records of prenatal care – which were not available in the current epidemiological study – would also shed light on the nature of these influences on child outcomes. Also, although significant, the effects documented herein were generally small in magnitude, suggesting that there are additional sources of unexplained variability in social cognition worthy of future investigation. Likewise, biomedical risk and responsive parenting were not completely independent predictors of social cognition, leading to the possibility that heightened biomedical risk may also predict variability in parenting. Possible mechanisms that link early biomedical risk to both parenting and child behavior – for instance, through the use of longitudinal cross-lagged mediation models – may be useful in elucidating these pathways to social cognition. On a related note, the fact that social cognition and maternal responsivity were measured contemporaneously (i.e., both at 18 months) precludes inferences about causality, and additional studies are warranted to determine the directionality of effects. Finally, although cumulative risk indices are powerful measures for examining the extent of risk exposure on developmental outcomes, future studies comparing the utility of these metrics to individual risk factors (measured through client records or direct measurement of risk, e.g., degree of hypoxia, level of hyperglycemia or hypertension, length of time in specialized care, etc.), are warranted based on these preliminary results.

Supplementary Material

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fpsyg.2015.00354/abstract

References

Adolphs, R. (2001). The neurobiology of social cognition. *Curr. Opin. Neurobiol.* 11, 231–239. doi: 10.1016/S0959-4388(00)00202-6

Aiken, L. S., and West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions.* Thousand Oaks: Sage.

Allen, N. B., Lewinsohn, P. M., and Seeley, J. R. (1998). Prenatal and perinatal influences on risk for psychopathology in childhood and adolescence. *Dev. Psychopathol.* 10, 513–529. doi: 10.1017/S0954579498001722

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Dev. Psychobiol.* 5, 297–305. doi: 10.1002/dev.420050403

Aschersleben, G., Hofer, T., and Jovanovic, B. (2008). The link between infant attention to goal-directed action and later theory of mind abilities. *Dev. Sci.* 11, 862–868. doi: 10.1111/j.1467-7687.2008.00736.x

Asendorpf, J. B., Warkentin, V., and Baudonnière, P.-M. (1996). Self-awareness and other-awareness. II: mirror self-recognition, social contingency awareness, and synchronic imitation. *Dev. Psychol.* 32, 313–321. doi: 10.1037/0012-1649.32.2.313

Astington, J. W., and Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory of mind development. *Dev. Psychol.* 35, 1311–1320. doi: 10.1037/0012-1649.35.5.1311

Atzaba-Poria, N., Pike, A., and Deater-Deckard, K. (2004). Do risk factors for problem behaviour act in a cumulative manner? An examination of ethnic minority and majority children through an ecological perspective. *J. Child Psychol. Psychiatry* 45, 707–718. doi: 10.1111/j.1469-7610.2004.00265.x

Behne, T., Carpenter, M., Call, J., and Tomasello, M. (2005). Unwilling versus unable: infants' understanding of intentional action. *Dev. Psychol.* 41, 328–337. doi: 10.1037/0012-1649.41.2.328

Bhutta, A. T., and Anand, K. J. S. (2001). Abnormal cognition and behavior in preterm neonates linked to smaller brain volumes. *Trends Neurosci.* 24, 129–132. doi: 10.1016/S0166-2236(00)01747-1

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M., Anand, K. J. S., and Phil, D. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 288, 728–737. doi: 10.1001/jama.288.6.728

Blair, C., and Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Dev.* 78, 647–663. doi: 10.1111/j.1467-8624.2007.01019.x

Bowlby, J. (1982). Attachment and loss: retrospect and prospect. *Am. J. Orthopsychiatry* 52, 664–678. doi: 10.1111/j.1939-0025.1982.tb01456.x

Brody, G. H., Dorsey, S., Forehand, R., and Armistead, L. (2002). Unique and Protective contributions of parenting and classroom processes to the adjustment of African American children living in single-parent families. *Child Dev.* 73, 274–286. doi: 10.1111/1467-8624.00405

Brownell, C. A., Ramani, G. B., and Zerwas, S. (2006). Becoming a social partner with peers: cooperation and social understanding in one- and two-year-olds. *Child Dev.* 77, 803–821. doi: 10.1111/j.1467-8624.2006.t01-1-.x-i1

Brownell, C. A., Zerwas, S., and Ramani, G. B. (2007). "So Big": the development of body self-awareness in toddlers. *Child Dev.* 78, 1426–1440. doi: 10.1111/j.1467-8624.2007.01075.x

Buka, S. L., Tsuang, M. T., and Lipsitt, L. P. (1993). Pregnancy/delivery complications and psychiatric diagnosis: a prospective study. *Arch. Gen. Psychiatry* 50, 151–156. doi: 10.1001/archpsyc.1993.01820140077009

Burchinal, M. R., Roberts, J. E., Hooper, S., and Zeisel, S. A. (2000). Cumulative risk and early cognitive development: a comparison of statistical risk models. *Dev. Psychol.* 36, 793–807. doi: 10.1037/0012-1649.36.6.793

Burchinal, M., Roberts, J. E., Zeisel, S. A., Hennon, E. A., and Hooper, S. (2006). Social risk and protective child, parenting, and child care factors in early elementary school years. *Parent. Sci. Pract.* 6, 79–113. doi: 10.1207/s15327922par0601_4

Burchinal, M., Vernon-Feagans, L., Cox, M., and Investigators, K. F. L. P. (2008). Cumulative social risk, parenting, and infant development in rural low-income communities. *Parent. Sci. Pract.* 8, 41–69. doi: 10.1080/15295190701830672

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112, 337–342. doi: 10.1016/j.cognition.2009.05.006

Carmody, D. P., Bendersky, M., Dunn, S. M., DeMarco, J. K., Hegyi, T., Hiatt, M., et al. (2006). Early risk, attention, and brain activation in adolescents born preterm. *Child Dev* 77, 384–394. doi: 10.1111/j.1467-8624.2006.00877.x

Cnattingius, S., Hultman, C. M., Dahl, M., and Sparén, P. (1999). Very preterm birth, birth trauma, and the risk of anorexia nervosa among girls. *Arch. Gen. Psychiatry* 56, 634–638. doi: 10.1001/archpsyc.56.7.634

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*: London: Routledge.

Cohen, P., Velez, C. N., Brook, J., and Smith, J. (1989). Mechanisms of the relation between perinatal problems, early childhood illness, and psychopathology in late childhood and adolescence. *Child Dev.* 60, 701–709. doi: 10.2307/1130735

Csibra, G., Bıró, S., Koós, O., and Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cogn. Sci.* 27, 111–133. doi: 10.1207/s15516709cog2701_4

Cutting, A. L., and Dunn, J. (1999). Theory of mind, emotion understanding, language and family background: Individual differences and interrelations. *Child Dev.* 70, 853–865. doi: 10.1111/1467-8624.00061

de Rosnay, M., and Harris, P. L. (2002). Individual differences in children's understanding of emotion: The roles of attachment and language. *Attach Hum. Dev.* 4, 39–54. doi: 10.1080/14616730210123139

de Rosnay, M., and Hughes, C. (2006). Conversation and theory of mind: do children talk their way to socio-cognitive understanding? *Br. J. Dev. Psychol.* 24, 7–37. doi: 10.1348/026151005X82901

Deater-Deckard, K., Dodge, K. A., Bates, J. E., and Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: group and individual differences. *Dev. Psychopathol.* 10, 469–493. doi: 10.1017/S0954579498001709

Decety, J., and Sommerville, J. A. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends Cogn. Sci.* 7, 527–533. doi: 10.1016/j.tics.2003.10.004

Dong, M., Anda, R. F., Felitti, V. J., Dube, S. R., Williamson, D. F., Thompson, T. J., et al. (2004). The interrelatedness of multiple forms of childhood abuse, neglect, and household dysfunction. *Child Abuse Negl.* 28, 771–784. doi: 10.1016/j.chiabu.2004.01.008

Dumontheil, I., Apperly, I. A., and Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Dev. Sci.* 13, 331–338. doi: 10.1111/j.1467-7687.2009.00888.x

Dunn, J., and Brophy, M. (2005). Communication, relationships, and individual differences in children's understanding of mind. *Paper presented at the Why Language Matters for Theory of Mind,* University of Toronto, Toronto, ON. doi: 10.1093/acprof:oso/9780195159912.003.0003

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., and Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: individual differences and their antecedents. *Child Dev.* 62, 1352–1366. doi: 10.2307/1130811

Enders, C. K., and Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct. Equ. Modeling.* 8, 430–457. doi: 10.1207/S15328007SEM0803_5

Evans, G. W., Li, D., and Whipple, S. S. (2013). Cumulative risk and child development. *Psychol. Bull.* 139, 1342–1396. doi: 10.1037/a0031808

Feldman, R., Rosenthal, Z., and Eidelman, A. I. (2014). Maternal-preterm skin-to-skin contact enhances child physiologic organization and cognitive control across the first 10 years of life. *Biol. Psychiatry* 75, 56–64. doi: 10.1016/j.biopsych.2013.08.012

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., et al. (1994). Variability in early communicative development. monographs of the society for research in child development. *Monogr. Soc. Res. Child Dev.* 59, 1–173. doi: 10.2307/1166093

Fernyhough, C. (2008). Getting Vygotskian about theory of mind: mediation, dialogue, and the development of social understanding. *Dev. Rev.* 28, 225–262. doi: 10.1016/j.dr.2007.03.001

Fiske, S. T., and Taylor, S. E. (2013). *Social Cognition: From Brains to Culture.* Thousand Oaks: Sage. doi: 10.4135/9781446286395

Flouri, E., and Kallis, C. (2007). Adverse life events and psychopathology and prosocial behavior in late adolescence: testing the timing, specificity, accumulation, gradient, and moderation of contextual risk. *J. Am. Acad. Child Adolesc. Psychiatry.* 46, 1651–1659. doi: 10.1097/chi.0b013e318156a81a

Ford, R. M., Neulinger, K., O'Callaghan, M., Mohay, H., Gray, P., and Shum, D. (2011). Executive function in 7–9-year-old children born extremely preterm or

with extremely low birth weight: Effects of biomedical history, age at assessment, and socioeconomic status. *Arch. Clin. Neuropsychol.* 26, 632–644. doi: 10.1093/arclin/acr061

Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133, 694–724. doi: 10.1037/0033-2909.133.4.694

Geddes, J. R., and Lawrie, S. M. (1995). Obstetric complications and schizophrenia: a meta-analysis. *Br. J. Psychiatry* 167, 786–793. doi: 10.1192/bjp.167.6.786

Gergely, G., and Csibra, G. (2003). Teleological reasoning in infancy: the naive theory of rational action. *Trends Cogn. Sci.* 7, 287–292. doi: 10.1016/S1364-6613(03)00128-1

Guth, C., Jones, P., and Murray, R. (1993). Familial psychiatric illness and obstetric complications in early-onset affective disorder. A case-control study. *Br. J. Psychiatry* 163, 492–498. doi: 10.1192/bjp.163.4.492

Hendderson, M. M., Weiss, N. S., Sacks, D. A., Pettitt, D. J., Selby, J. V., Quesenberry, C. P., et al. (2006). Pregnancy weight gain and risk of neonatal complications: macrosomia, hypoglycemia, and hyperbilirubinemia. *Obstet. Gynecol.* 108, 1153–1161. doi: 10.1097/01.AOG.0000242568.75785.68

Holmes, H. A., Black, C., and Miller, S. A. (1996). A cross-task comparison of false belief understanding in a Head Start population. *J. Exp. Child Psychol.* 63, 263–285. doi: 10.1006/jecp.1996.0050

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118

Hughes, C., and Ensor, R. (2006). Behavioural problems in 2 year olds: links with individual differences in theory of mind, executive function and harsh parenting. *J. Child Psychol. Psychiatry* 47, 488–497. doi: 10.1111/j.1469-7610.2005.01519.x

Jenkins, J., Madigan, S., and Arsenault, L. (in press). "Psychosocial adversity," in *Rutter's Handbook of Child and Adolescent Psychiatry*, 6th Edn, eds M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, and E. A. Taylor (Oxford: Blackwell).

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., et al. (2008). A structural MRI study of human brain development from birth to 2 years. *J. Neurosci.* 28, 12176–12182. doi: 10.1523/JNEUROSCI.3479-08.2008

Knoblich, G. N., and Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 2021–2031. doi: 10.1098/rstb.2008.0006

Kochanska, G., DeVet, K., Goldman, M., Murray, K., and Putnam, S. P. (1994). Maternal reports of conscience development and temperament in young children. *Child Dev.* 65, 852–868. doi: 10.2307/1131423

Kolevzon, A., Gross, R., and Reichenberg, A. (2007). Prenatal and perinatal risk factors for autism: a review and integration of findings. *Arch. Pediatr. Adolesc. Med.* 161, 326–333. doi: 10.1001/archpedi.161.4.326

Kurinczuk, J. J., White-Koning, M., and Badawi, N. (2010). Epidemiology of neonatal encephalopathy and hypoxic-ischaemic encephalopathy. *Early. Hum. Dev.* 86, 329–338. doi: 10.1016/j.earlhumdev.2010.05.010

Landry, S. H., Smith, K. E., Miller-Loncar, C. L., and Swank, P. R. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Dev. Psychol.* 33, 1040–1053. doi: 10.1037/0012-1649.33.6.1040

Landry, S. H., Smith, K. E., and Swank, P. R. (2006). Responsive parenting: establishing early foundations for social, communication, and independent problem-solving skills. *Dev. Psychol.* 42, 627–642. doi: 10.1037/0012-1649.42.4.627

Landry, S. H., Smith, K. E., Swank, P. R., and Guttentag, C. (2008). A responsive parenting intervention: the optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Dev. Psychol.* 44, 1335–1353. doi: 10.1037/a0013030

Landry, S. H., Smith, K. E., Swank, P. R., Zucker, T., Crawford, A. D., and Solari, E. F. (2012). The effects of a responsive parenting intervention on parent-child interactions during shared book reading. *Dev. Psychol.* 48, 969–986. doi: 10.1037/a0026400

Laucht, M., Esser, G. N., and Schmidt, M. H. (1997). Developmental outcome of infants born with biological and psychosocial risks. *J. Child Psychol. Psychiatry* 38, 843–853. doi: 10.1111/j.1469-7610.1997.tb01602.x

Laucht, M., Esser, G. N., and Schmidt, M. H. (2001). Differential development of infants at risk for psychopathology: the moderating role of early maternal responsivity. *Dev. Med. Child Neurol.* 43, 292–300. doi: 10.1017/S0012162201000561

Lederman, S. A., and Paxton, A. (1998). Maternal reporting of prepregnancy weight and birth outcome: consistency and completeness compared with the clinical record. *Matern. Child Health J.* 2, 123–126. doi: 10.1023/A:1022996924094

Legerstee, M. (2005). *Infants' Sense of People: Precursors to a Theory of Mind.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511489747

Levy-Shiff, R., Einat, G., Mogilner, M. B., Lerman, M., and Krikler, R. (1994). Biological and environmental correlates of developmental outcome of prematurely born infants in early adolescence. *J. Pediatr. Psychol.* 19, 63–78. doi: 10.1093/jpepsy/19.1.63

Lewis, M. (2003). The emergence of consciousness and its role in human development. *Ann. N. Y. Acad. Sci.* 1001, 104–133. doi: 10.1196/annals.1279.007

Lubchenco, L. O., and Bard, H. (1971). Incidence of hypoglycemia in newborn infants classified by birth weight and gestational age. *Pediatrics* 47, 831–838.

Luthar, S. S., Cicchetti, D., and Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Dev.* 71, 543–562. doi: 10.1111/1467-8624.00164

Masten, A. S., Cutuli, J. J., Herbers, J. E., and Gabrielle-Reed, M. J. (2009). "Resilience in development," in *The handbook of positive psychology,* 2nd Edn, eds C. R. Snyder and S. J. Lopez (New York: Oxford University Press), 117–131.

Matias, C. S. F. (2006). *Direct Observation of Parent-Child Interaction based on Attachment Theory*, Ph.D. thesis, King's College London, London.

Milberger, S., Biederman, J., Faraone, S. V., Guite, J., and Tsuang, M. T. (1997). Pregnancy, delivery and infancy complications and attention deficit hyperactivity disorder: issues of gene-environment interaction. *Biol. Psychiatry* 41, 65–75. doi: 10.1016/0006-3223(95)00653-2

Moore, C. (2007). "Understanding self and others in the second year," in *Socioemotional Development in the Toddler Years: Transitions and Transformations*, eds C. A. Brownell and C. B. Koop (New York, NY: Guilford), 43–65.

Moriguchi, Y., Ohnishi, T., Mori, T., Matsuda, H., and Komaki, G. (2007). Changes of brain activity in the neural substrates for theory of mind during childhood and adolescence. *Psychiatry Clin. Neurosci.* 61, 355–363. doi: 10.1111/j.1440-1819.2007.01687.x

Moses, L. J., Baldwin, D. A., Rosicky, J. G., and Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Dev.* 72, 718–735. doi: 10.1111/1467-8624.00311

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., and Seibert, J. (2003). *Early Social Communikation Scales (ESCS).* Coral Gables: University of Miami.

Nielsen, M., and Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: a longitudinal investigation through the second year. *Infant. Behav. Dev.* 27, 342–365. doi: 10.1016/j.infbeh.2003.12.006

Okura, Y., Urban, L. H., Mahoney, D. W., Jacobsen, S. J., and Rodeheffer, R. J. (2004). Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J. Clin. Epidemiol.* 57, 1096–1103. doi: 10.1016/j.jclinepi.2004.04.005

Pasamanick, B., and Knobloch, H. (1961). *Epidemiologic Studies on the Complications of Pregnancy and the Birth Process. Prevention of Mental Disorders in Children*. New York, NY: Basic Books, 74–94.

Pears, K. C., and Moses, L. J. (2003). Demographics, parenting, and theory of mind in preschool children. *Soc. Dev.* 12, 1–20. doi: 10.1111/1467-9507.00219

Peterson, B. S., Vohr, B., Staib, L. H., Dolberg, A., Schneider, K. C., Katz, K. H., et al. (2000). Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. *JAMA* 284, 1939–1947. doi: 10.1001/jama.284.15.1939

Pons, F., and Harris, P. (2005). Longitudinal change and longitudinal stability of individual differences in children's emotion understanding. *Cogn. Emot.* 19, 1158–1174. doi: 10.1080/02699930500282108

Pons, F., Lawson, J., Harris, P. L., and de Rosnay, M. (2003). Individual differences in children's emotion understanding: Effects of age and language. *Scand. J. Psychol.* 44, 347–353. doi: 10.1111/1467-9450.00354

Preacher, K. J., and Rucker, D. (2003). A primer on interaction effects in multiple linear regression. Available at: http://www.quantpsy.org/interact/interactions.htm (accessed November 25, 2014).

Radloff, L. S. (1977). The CES-D Scale. A self reort depression scale for research in the general population. *Appl. Psychol. Meas.* 1, 385–401. doi: 10.1177/014662167700100306

Raine, A., Brennan, P., and Mednick, S. A. (1994). Birth complications combined with maternal rejection at age 1 year predispose to violent crime at age 18 years. *Arch. Gen. Psychiatry* 51, 984–988. doi: 10.1001/archpsyc.1994.03950120056009

Raine, A., Phil, D., Brennan, P., Sarnoff, A., and Mednick, S. A. (1997). Interaction between Birth complications and early maternal rejection in predisposing individuals to adult violence. Specificity to serious, early-onset violence. *Am. J. Psychiatry* 154, 1265–1271. doi: 10.1176/ajp.154.9.1265

Raznahan, A., Greenstein, D., Lee, N. R., Clasen, L. S., and Giedd, J. N. (2012). Prenatal growth in humans and postnatal brain maturation into late adolescence. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11366–11371. doi: 10.1073/pnas.1203350109

Razza, R. A., and Blair, C. (2009). Associations among false-belief understanding, executive function, and social competence: a longitudinal analysis. *J. Appl. Dev. Psychol.* 30, 332–343. doi: 10.1016/j.appdev.2008.12.020

Repacholi, B. M., and Gopnik, A. (1997). Early reasoning about desires: evidence from 14-and 18-month-olds. *Dev. Psychol.* 33, 12–21. doi: 10.1037/0012-1649.33.1.12

Roth-Hanania, R., Davidov, M., and Zahn-Waxler, C. (2011). Empathy development from 8 to 16 months: early signs of concern for others. *Infant Behav. Dev.* 34, 447–458. doi: 10.1016/j.infbeh.2011.04.007

Ruffman, T., Slade, L., Devitt, K., and Crowe, E. (2006). What mothers say and what they do: The relation between parenting, theory of mind, language and conflict/cooperation. *Br. J. Dev. Psychol.* 24, 105–124. doi: 10.1348/026151005X82848

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *Am. J. Ortho. Psychiatry* 57, 316–331. doi: 10.1111/j.1939-0025.1987.tb03541.x

San Juan, V., and Astington, J. W. (2012). Bridging the gap between implicit and explicit understanding: how language development promotes the processing and representation of false belief. *Br. J. Dev. Psychol.* 30, 105–122. doi: 10.1111/j.2044-835X.2011.02051.x

Schore, A. N. (1996). The experience-dependent maturation of a regulatory system in the orbital prefrontal cortex and the origin of developmental psychopathology. *Dev. Psychopathol.* 8, 59–87. doi: 10.1017/S0954579400006970

Schore, A. N. (2001). Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Ment. Health J.* 22, 7–66. doi: 10.1002/1097-0355(200101/04)22:1<7::AID-IMHJ2>3.0.CO;2-N

Shah, D. M. (2001). Perinatal implications of maternal hypertension. *Semin. Pediatr. Neurol.* 8, 108–119.

Shatz, M., Diesendruck, G., Martinez-Beck, I., and Akar, D. (2003). The influence of language and socioeconomic status on children's understanding of false belief. *Dev. Psychol.* 39, 717–729. doi: 10.1037/0012-1649.39.4.717

Stanek, J. (2013). Hypoxic patterns of placental injury: a review. *Arch. Pathol. Lab. Med.* 137, 706–720. doi: 10.5858/arpa.2011-0645-RA

Sydsjö, G. (2011). Long-term consequences of non-optimal birth characteristics. *Am. J. Reprod. Immunol.* 66(Suppl. 1), 81–87. doi: 10.1111/j.1600-0897.2011.01035.x

Teramo, K. A. (2010). Obstetric problems in diabetic pregnancy: the role of fetal hypoxia. *Best Pract. Res. Clin. Endocrinol. Metab.* 24, 663–671. doi: 10.1016/j.beem.2010.05.005

Thoermer, C., Sodian, B., Vuori, M., Perst, H., and Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *Br. J. Dev. Psychol.* 30, 172–187. doi: 10.1111/j.2044-835X.2011.02067.x

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. *Dev. Sci.* 10, 121–125. doi: 10.1111/j.1467-7687.2007.00573.x

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–690. doi: 10.1017/S0140525X05000129

Tomeo, C. A., Rich-Edwards, J. W., Michels, K. B., Berkey, C. S., Hunter, D. J., Frazier, A. L., et al. (1999). Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology* 10, 774–776. doi: 10.1097/00001648-199911000-00022

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. doi: 10.1002/hbm.20547

Verdoux, H., Geddes, J. R., Takei, N., Lawrie, S. M., Bovet, P., Eagles, J. M., et al. (1997). Obstetric complications and age at onset in schizophrenia: an international collaborative meta-analysis of individual patient data. *Am. J. Psychiatry* 154, 1220–1227. doi: 10.1176/ajp.154.9.1220

Voigt, B., Brandl, A., Pietz, J., Pauen, S., Kliegel, M., and Reuner, G. (2013). Negative reactivity in toddlers born prematurely: indirect and moderated pathways considering self-regulation, neonatal distress and parenting stress. *Infant Behav. Dev.* 36, 124–138. doi: 10.1016/j.infbeh.2012.11.002

Wade, M., Browne, D., Madigan, S., Plamondon, A., and Jenkins, J. M. (2014a). Normal birth weight variation and children's neuropsychological functioning: links between language, executive functioning, and theory of mind. *J. Int. Neuropsychol. Soc.* 20, 909–919. doi: 10.1017/S1355617714000745

Wade, M., Hoffmann, T., and Jenkins, J. M. (2014b). Association between the oxytocin receptor (OXTR) gene and children's social cognition at 18 months. *Genes Brain Behav.* 13, 603–610. doi: 10.1111/gbb.12148

Wade, M., Moore, C., Astington, J. W., Frampton, K., and Jenkins, J. M. (2014c). Cumulative contextual risk, maternal responsivity, and social cognition at 18 months. *Dev. Psychopathol.* 27, 189–203. doi: 10.1017/S0954579414000674

Wade, M., and Jenkins, J. M. (2014). Pregnancy hypertension and the risk for neuropsychological difficulties across early development: a brief report. *Child Neuropsychol.* 1–8. doi: 10.1080/ 09297049.09292014.09958070 [Epub ahead of print].

Walhovd, K. B., Fjell, A. M., Brown, T. T., Kuperman, J. M., Chung, Y., Hagler, D. J., et al. (2013). Long-term influence of normal variation in neonatal characteristics on human brain development. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20089–20094. doi: 10.1073/pnas.1208180109

Walhovd, K. B., Tamnes, C. K., and Fjell, A. M. (2014). Brain structural maturation and the foundations of cognitive behavioral development. *Curr. Opin. Neurol.* 27, 176–184. doi: 10.1097/WCO.0000000000000074

Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Dev.* 77, 640–663. doi: 10.1111/j.1467-8624.2006.00895.x

Warneken, F., and Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy* 11, 271–294. doi: 10.1111/j.1532-7078.2007.tb00227.x

Wellman, H. M., Lopez-Duran, S., LaBounty, J., and Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Dev. Psychol.* 44, 618–623. doi: 10.1037/0012-1649.44.2.618

Yuan, K.-H., and Bentler, P. M. (2000). Three Likelihood-Based Methods For Mean and Covariance Structure Analysis With Nonnormal Missing Data. *Sociol. Methodol.* 30, 165–200. doi: 10.1111/0081-1750.00078

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read,
for greatest visibility

COLLABORATIVE PEER-REVIEW
Designed to be rigorous
– yet also collaborative,
fair and constructive

FAST PUBLICATION
Average 85 days from
submission to publication
(across all journals)

COPYRIGHT TO AUTHORS
No limit to article
distribution and re-use

TRANSPARENT
Editors and reviewers
acknowledged by name
on published articles

SUPPORT
By our Swiss-based
editorial team

IMPACT METRICS
Advanced metrics
track your article's impact

GLOBAL SPREAD
5'100'000+ monthly
article views
and downloads

LOOP RESEARCH NETWORK
Our network
increases readership
for your article

Find us on