Developmental Psychology
2011, Vol. 47, No. 5, 1207–1219

0012-1649/11/$12.00 DOI: 10.1037/a0024808

Infants' Joint Attention Skills Predict Toddlers' Emerging Mental State Language

Susanne Kristen, Beate Sodian, Claudia Thoermer, and Hannah Perst
Ludwig Maximilian University Munich

To assess predictive relations between joint attention skills, intention understanding, and mental state vocabulary, 88 children were tested with measures of comprehension of gaze and referential pointing, as well as the production of declarative gestures and the comprehension and production of imperative gestures, at the ages of 7–18 months. Infants' intention-based imitation skills were assessed at 12, 15, and 18 months. At the ages of 24 and 36 months, toddlers' internal state lexicon was evaluated by parents with a German adaptation of the Mental State Language Questionnaire (Olineck & Poulin-Dubois, 2005). Regression analyses revealed that 9-months-olds' comprehension of referential pointing contributed significantly to the prediction of intention-based imitation skills at 15 months, as well as to children's volition and cognition vocabularies at 24 and 36 months, respectively. Moreover, 12-month-olds' comprehension of an imperative motive was shown to selectively predict toddlers' use of volition terms at 24 months. Overall, these results provide empirical evidence for both general and specific developmental relations between preverbal communication skills and mental state language, thus implying developmental continuity within the social domain in the first 3 years of life.

Keywords: joint attention, mental state language, intention reading, theory of mind, longitudinal development

Children's talk about the mind is an important source of evidence for their nascent understanding of the mental domain in the second and third years of life (Bartsch & Wellman, 1995; Bretherton & Beeghly, 1982). Mental state language is assumed to be crucial in working out an explicit theory of mind by providing the terminology to reflect on mental states (Ruffman, Slade, Rowlandson, Rumsey, & Garnham, 2003). The early development of mental state language was studied extensively in a parent diary study by Bretherton and Beeghly (1982) and in an analysis of the CHILDES database by Bartsch and Wellman (1995). Children begin to talk about mental states about half a year before they can give full-fledged mentalistic explanations of action (Wellman & Woolley, 1990). Talk about their own and others' desires developmentally precedes children's talk about epistemic states (Bretherton & Beeghly, 1982). Around their second birthday, children have acquired the terminology to reflect explicitly on people's desires (e.g., *want*), whereas they begin to use cognition terms (e.g., *think*) to refer to epistemic states shortly before their third birthday (Wellman & Woolley, 1990). Because preverbal infants demonstrate an implicit understanding of mental states in their use and understanding of communicative gestures, beginning around the end of the first year of life, one would expect mental state language, as a first sign of an explicit understanding of the mind, to emerge from a preverbal understanding of communicative intentions.

In the present study, we investigate this issue in a longitudinal design, testing for developmental continuity on the level of individual differences and thus addressing the developmental relation between an implicit understanding of the mind in infancy and a later explicit theory of mind. In recent research, the longitudinal approach has been productively used to investigate developmental continuity between social information processing in the first year of life and theory of mind competencies around the age of 4 years (see e.g., Aschersleben, Hofer, & Jovanovic, 2008; Wellman, Lopez-Duran, LaBounty, & Hamilton, 2008; Wellman, Phillips, Dunphy-Lelii, & Lalonde, 2004; Yamaguchi, Kuhlmeier, Wynn, & VanMarle, 2008). However, to date, longitudinal studies have examined developmental end points of first-order theory of mind skills rather than the continuum from the first emergence of an implicit theory of mind in joint attention, through intention reading skills developing in the second year of life, and toward the emergence of an explicit, verbal theory of mind.

Joint attention is characterized by the coordination of attention between the self, the other, and some external object or event (Tomasello, 1995). Infants begin to respond to joint attention (e.g., follow gaze or follow the pointing gesture) around the age of 6–9 months before they initiate joint attention themselves (e.g., by pointing to an object; Bakeman & Adamson, 1984; Butterworth &

Susanne Kristen, Beate Sodian, Claudia Thoermer, and Hannah Perst, Department of Psychology, Ludwig Maximilian University Munich, Munich, Germany.

This study was funded by Grant So 213/27-1 from the German Research Council to Beate Sodian. The authors would like to thank the parents and children who participated for their time and commitment to the longitudinal project. We thank our research assistants for their help with collecting and coding the data. Constructive comments by Diane Poulin-Dubois and Neil Berthier on earlier versions of the article are gratefully acknowledged.

Correspondence concerning this article should be addressed to Susanne Kristen, Department of Psychology, Ludwig Maximilian University Munich, Leopoldstr. 13, D-80802 Munich, Germany. E-mail: susanne.kristen@psy.lmu.de

Jarrett, 1991; Woodward, 2003). A basic communicative function of joint attention is to establish reference, that is, to direct another person's attention toward a specific external referent (an object or event) or to respond to another person's gesture by picking out the intended referent. Once an intended referent is identified, the adult can share affect or interest in an object or event with the infant and communicate about it. Such declarative competencies are designed to convey information about or attitudes toward a target. The mere comprehension of referential intent has sometimes been distinguished from the comprehension of a declarative motive (cf. Camaioni, Perucchini, Bellagamba, & Colonnesi, 2004), indicated by the tendency to share affect or interest with the adult (e.g., by smiling or "commenting" on the target) rather than to just follow the adult's gaze or point (Aureli, Perucchini, & Genco, 2009). Tomasello (1995) proposed that between 9 and 12 months of age infants begin to understand persons as intentional agents and that a general concept of intentional agency underlies the development of specific joint attentional skills. For instance, to understand a declarative pointing gesture, it is necessary that a pointer and a recipient "know some things or are attending to some things together" (Tomasello, Carpenter, & Liszkowski, 2007, p. 706).

This assumption is supported by the finding of developmental relations between the emergence of referential and declarative joint attentional skills and later theory of mind.

A study by Charman et al. (2000) found an association between joint attention at 20 months (gaze switches between an adult and an active toy, and looking at an adult during an ambiguous goal detection task) and theory of mind at 44 months. Also, 12-month-olds who pointed more frequently to share information with an unfamiliar tester and who followed pointing gestures more consistently had higher social competence scores than did their peers at 30 months in a study by Vaughan Van Hecke et al. (2007). Moreover, 12-month-olds' responding to joint attention and their initiation of joint attention at 18 months is longitudinally related to 24-month-olds' language competence (Mundy et al., 2007). This can be seen as indirect evidence that joint attention relates to children's later theory of mind, because language seems crucial in working out an explicit theory of mind (see e.g., Slade & Ruffman, 2005).

Although declarative joint attentional skills serve the purpose of conveying information about objects or events, imperative gestures are produced to request someone to act on something, for instance to give the child an object, that is, to influence others' behavior (see e.g., Bates, Camaioni, & Volterra, 1975). Therefore, lean interpretations of imperative pointing (Baron-Cohen, 1989; Camaioni, 1993; Moore & D'Entremont, 2001) have proposed that imperative competencies in young children involve the use of gestures as social tools, on the basis of a construal of agents in behavioral terms (see also Bates et al., 1975). The meaning of imperative gestures may at first be established in commonly practiced social situations (cf. Montgomery, 2002) and originate from highly routinized back-and-forth games such as handing over toys (Myers-Walls, Myers-Bowman, & Posada, 2006). Recently, Tomasello et al. (2007) have proposed that imperatives may form a developmental continuum from direct requests with individualistic motives to indirect requests with cooperative motives. Whereas indirect requests require the child to consider others' intentional or motivational states, direct imperative requests are designed to make others' behavior match one's own desire or motive. This does not imply mind-blind behaviorism. Rather, given the rich findings on infants' attribution of goals, motives, and dispositions to agents beginning in the first year of life (see Sodian, 2011, for a review), it seems likely that infants represent others as goal-directed agents when initiating or responding to imperative bids. However, the purpose of imperative bids is to produce behavioral effects (cause people to "do" things) rather than to produce epistemic effects (cause people to "know" things). This difference is also reflected in the way reference is typically established in imperative versus declarative joint attentional episodes. In imperative exchanges, the referents of children's gestures are typically unambiguous and located in the shared visual field between child and adult. In declarative communicative interchanges, referents are typically ambiguous and not in the child's direct line of sight (the communicative partner's attention is drawn to something that he or she was not already attending to).

These considerations have led to differential predictions with respect to the developmental course of declarative and imperative joint attentional skills. It has been argued that declarative pointing, but not imperative pointing, reveals a nascent understanding of intentional states in the sense of distinguishing the surface behavior (what people actually do) from a deeper level (what people intend to do), because declarative pointing is aimed at conveying information and thereby affecting mental states, whereas imperative pointing focuses on influencing others' behavior. Consistent with these predictions, a longitudinal study by Camaioni et al. (2004) revealed that 12- and 15-month-olds' production of declarative pointing gestures was linked to understanding another person's intention in an imitation task (Meltzoff, 1995), whereas production and comprehension of imperative pointing was not. Similarly, point following was related to intention-based imitation skills in the longitudinal study of 24 infants' joint attention development by Carpenter, Nagell, and Tomasello (1998), suggesting that the basic ability to share another person's referential focus may be predictive of later intention reading in the sense of distinguishing between intention-based and arbitrary actions. Furthermore, children with autism and Down syndrome are more severely impaired in their ability to comprehend and produce declarative gestures than in their production and comprehension of requestive gestures (see e.g., Camaioni, Perucchini, Muratori, Parrini, & Cesari, 2003; Legerstee & Fisher, 2008). Finally, Olineck and Poulin-Dubois (2005) found evidence for predictive relations between 14- and 18-month-old infants' selective imitation skills, which is an early indicator of intention reading, and mental state talk at 30 months of age.

Taken together, these findings support the hypothesis that individual differences in the production and comprehension of referential and declarative communicative gestures before or around the first birthday should be predictive of intention reading skills in the second year of life and mental state talk in the third year.

A study by Colonnesi, Rieffe, Koops, and Perucchini (2008) supports the idea that the comprehension of imperative communicative intent may also be related to later mental state reasoning—but to reasoning about desires rather than epistemic states. Colonnesi et al. (2008) found a relation between 12-month-olds' comprehension of an imperative pointing gesture and the same children's predominantly desire-based action explanations at the age of 39 months, independently of children's receptive language skills. This finding indicates that an early, preverbal comprehen-

sion of others' behavioral requests, as indicated by the ability to follow requestive gestures, may be a specific developmental precursor to later reasoning about desires. The hypothesis that imperative communicative competence may be specifically related to talk about desires, but not to epistemic state talk, is supported by evidence from autism research. Autistic children, who have been shown to produce fewer declarative gestures, are relatively unimpaired in their use of imperative gestures (Camaioni et al., 2003) and are also specifically impaired in their ability to talk about cognitive state terms, whereas they are better able to produce desire, emotion, and perception terms (Hadwin, Baron-Cohen, Howlin, & Hill, 1997; Tager-Flusberg, 1992).

The findings just presented are consistent with functional conceptualizations of mental state language, suggesting that by talking about mental states, in analogy to joint attention bids, children quite deliberately direct attention to their own and others' inner states, at first primarily to desires, followed by cognitive states and processes (see e.g., Moore, 2006). Similarly, the contextual view of semantic development emphasizes the relevance of early communicative exchanges in establishing the meaning of mental verbs. Accordingly, proto-imperatives are related to the preverbal obtainment of objects (Bates et al., 1975) and are at the center of communicative exchanges from which desire verbs emerge. The subsequent linguistic development involves the expanding of ways of talking about desire within communicative acts. In contrast, proto-declaratives are at the center of communicative exchanges about knowledge and information from which cognition terms (e.g., *know*) emerge. Thus, the subsequent linguistic development involves the expanding of ways of talking about cognition within communicative acts. Therefore, theoretical considerations as well as empirical findings lead one to expect a developmental relation between preverbal imperative communicative competencies and later talk about desires, as well as between preverbal declarative competencies and later talk about cognition.

If, however, imperative and declarative communicative competencies emerge from a common joint attentional frame within which the child and another person direct their attention toward external objects or events, it appears unlikely that there should be two entirely separate developmental pathways from proto-imperatives to desire reasoning at 24 months and from proto-declaratives to epistemic state reasoning at 36 months. Rather, it is likely that basic joint attentional skills, such as the ability to follow into someone's attentional focus, and thus to share reference, should be related to the development of both imperative and declarative competencies.

Previous longitudinal studies have not spanned a large enough age range to test these predictions. For instance, in the study by Camaioni et al. (2004) children had an average age of 11 months at the first measurement point, when there is little individual variation in the basic ability to comprehend another person's referential intent. Therefore, there was no way of addressing the developmental relation between individual differences in the early acquisition of the joint attentional frame and later individual differences on both the declarative and the imperative pathways. If infants' social understanding undergoes a conceptual revolution around the age of 9 months, as Tomasello (1995) proposed, then individual differences in basic joint attentional skills, such as referential point following at the age of 9 months, should be predictive of later intention reading and mental state understanding. Moreover, previous studies have failed to include control measures such as working memory capacity and general language ability. Both joint attention skills and imitation skills have been empirically linked to working memory skills (cf. Nichols, Fox, & Mundy, 2005). Thus intention-based imitation skills and joint attention skills might be linked to each other by general processing abilities rather than conceptual comprehension. Likewise, a developmental link between joint attention skills and mental state language might be based on children's general language ability rather than their conceptual understanding (cf. Mundy & Gomes, 1998).

In the present study, we aimed at a fuller understanding of the developmental relations between joint attention skills, intention understanding, and mental state talk. To test for developmental links between preverbal communicative gestures, mind-reading skills, and mental state vocabulary, we assessed imperative and declarative joint attentional skills at the ages of 7, 9, 12, 15, and 18 months, as well as (a) implicit intention reading skills in an imitation task in the first half of the second year and (b) volition and cognition talk at the ages of 24 and 36 months, respectively, in a relatively large sample.

Finally, we included measures of working memory and general language ability. We expected to find, on the basis of theoretical considerations and previously reported findings by Carpenter et al. (1998), Camaioni et al. (2004), and Colonnesi et al. (2008), both general and specific predictive relations from individual differences in joint attentional skills to later intention reading and desire and cognition language. Further, we expected individual differences in basic referential understanding around the age of 9 months to be related to a wide range of later intention reading abilities and mental state understanding. We also expected specific predictive relations from individual differences in imperative skills, emerging around the age of 12 months, to individual differences in talk about volitional states at the age of 24 months, whereas we did not expect imperative skills to be related to tasks focusing on the sharing of information and verbal references to epistemic states. In contrast, we expected individual differences in these competencies to be predicted by individual differences in the production of declarative communicative acts, beginning around the age of 12 months.

Method

Participants

The sample consisted of 88 healthy, full-term infants (42 female, 46 male). Children's mean age at the first observation of the longitudinal study was 6.65 months ($SD = 0.41$ months), at the second observation 8.71 months ($SD = 0.41$ months), at the third observation 11.66 months ($SD = 0.41$ months), at the fourth observation 14.77 months ($SD = 0.43$ months), at the fifth observation 17.74 months ($SD = 0.41$ months), at the sixth observation 23.67 months ($SD = 0.42$ months), and at the seventh observation 36.09 months ($SD = 0.20$ months). Children came from the lower to upper middle class (as determined by parents' educational level) in an urban area in the southern part of Germany, and 36% of them had older siblings. In Study 2 all of the 76 monolingual children (38 female, 38 male) from Study 1 participated. At 24 months, Mental State Language Questionnaire (MSLQ) data were available for 69 monolingual children, and at 36 months, MSLQ data were

available for 62 monolingual children. However, at both time points, two children had to be excluded, one because of a temporary hearing problem and the other because of a dysgrammatism diagnosis.

Procedure

Each observation started with a warm-up period in which the experimenter (E) interacted with the child until the child seemed acclimated to the surroundings. All tasks were videotaped for later analyses.

Measures

Joint attention tasks. For the pointing and gaze-following tasks, infants were seated in a high chair opposite E at a rectangular table measuring 78 cm in height, 120 cm in width, and 78 cm in depth.

Point and gaze comprehension tasks.

Referential point and gaze comprehension task. At 7, 9, and 12 months, infants completed a task adapted from Mundy, Hogan, and Doehring (1996) and Mundy et al. (2003), in which E takes turns pointing and gazing at a target. Four colored DIN A3 pictures were attached to the left and right walls approximately 2 m distance from and slightly behind the child. Each child received two gaze-following and two point-following trials. E first got the child's attention by tapping the table and then touching her own nose. Once the child's attention was focused on E, in the two gaze-following conditions E would turn her head approximately 90 degrees off midline to look at pictures on the wall while calling the child's name three times in a row within 3 s, and in the two point-following conditions E looked and pointed at the pictures on the wall while calling the child's name three times in a row within 3 s. In both conditions, E did not look back at the child until after stating the child's name for the third time. E then made a statement related to the target to acknowledge that the child had turned and seen the poster or to further engage the child in the task. The point and gaze trials were administered in the following order: point right, gaze right, point left, gaze left. A score of 0 was given if a child failed to show any target behavior in both of the respective trials or looked at the tester. A score of 1 was given if the child looked at the index finger (point trials) or respectively at E's face (gaze trials). Children were given a score of 2, if they followed E's line of regard (gaze trials) or the pointing gesture (point trials) to the object. On the left and right point trials, the children got credit if they were looking in the right direction or respectively beyond the end of their index finger. On gaze trials, the look had to be in the direction of the poster. Scores thus ranged from 0 to 2 for each skill. Twenty-five percent of the children were scored by an independent observer. Cohen's κ for point and gaze comprehension ranged from .82 to 1.0.

Imperative comprehension task. In this task, E showed the child a toy bear that could be pulled apart into two pieces and put back together. One piece (the head of the bear) was given to the child. E then looked at the piece in her own hand and said, "Oops, the bear has no head!" E then pointed at the piece in the child's hand and looked silently at the child for about 15 s (waiting for the child's reaction). E then alternated looks between the child and the stimulus and looked silently at the child for another 15 s (waiting for the child's reaction). To produce the target behavior, the child, before or after the trigger, had to give the stimulus to E, tease E by offering the toy to her, and then draw it back again or refuse to give the stimulus to E and say "yes"/"no" or make a corresponding gesture with the head or leave the stimulus on the table for E while looking at her (score 2). If children just looked at E, they received a score of 1. If they showed no understanding of the situation and ignored E (e.g., by just playing with the toy), they were given a score of 0. Scores could thus range from 0 to 2.

Production tasks. The production task comprised two experimental conditions, designed to elicit declarative and imperative pointing. The two conditions were presented in the following fixed order: (1) imperative point production and (2) declarative point production. At the beginning of each trial, E made eye contact with the child to elicit his/her attention.

Declarative production task. A second assistant (hidden from the child) activated a mobile (pig) hanging from the wall opposite the child and behind E's back. The mobile moved for about 15 s, a period during which E looked silently at the child. When the mobile stopped, E said, "What happened?" without turning back to look at the mobile. E then looked silently at the child for 15 s (waiting for the child's reaction) and finally looked at the mobile and named it. The coding followed a two-step procedure. First, children's pointing gestures were coded as intentionally communicative when looking at E's face occurred within 2 s before or after the pointing gesture or when looks at E and pointing gestures occurred simultaneously. Second, the child was scored as having shown the target behavior (score 2) when she/he produced a pointing gesture before or after the verbal trigger in addition to at least one of the following target behaviors: smiled and/or vocalized toward the stimulus, produced a proto-word or a word relative to the stimulus, or reenacted what the stimulus did. Pointing was defined as an arm and index-finger extension toward the stimulus, with the remaining fingers curled tightly or lightly under the hand (Franco & Butterworth, 1996). If children did not yet point but showed an intentionally communicative whole-hand pointing gesture (e.g., grasping gesture; Leavens & Hopkins, 1999) directed toward the stimulus, in addition to at least one of the target behaviors mentioned, then their behavior was coded as Level 1 competence. If children interacted with E without reaching, grasping, or pointing or did not show any reaction, they received a score of 0. Scores thus ranged from 0 (*no target behavior*) through 1 (*point-like gesture*) to 2 (*pointing with prominent index finger*).

Imperative production task. In this task E activated a windup car, moved it in front of the child for about 10 s, and then stopped it and looked silently at the child for 15 s (waiting for the child's reaction). E then said, "Isn't it nice? Do you like it?" and looked silently at the child for another 15 s (waiting for the child's reaction). After that, the toy was given to the child. Again, the coding followed a two-step procedure. First, children's pointing gestures were coded as intentionally communicative when looking at E's face occurred within 2 s before or after the pointing gesture or when looks at E and pointing gestures occurred simultaneously. Children were scored as having shown the target behavior (score 2) when they pointed at the stimulus while producing request-like gestures (as indicated by a repeated opening of the hand) at the stimulus either before or after the verbal trigger in addition to at least one of the following target behaviors: leaned forward or reached for the stimulus, produced a request vocalization/proto-

word, or whined. As in the declarative point production task, if children did not yet point but showed an intended whole-hand pointing gesture in addition to at least one of the target behaviors mentioned, they were awarded a score of 1. Competence levels ranged from 0 (*no target behavior*) through 1 (*point-like gesture*) to 2 (*pointing with prominent index finger*).

For each joint attention task, a random sample of approximately 25% of the children was scored by an independent observer. Cohen's κ ranged from .87 to 1.0 for imperative comprehension and for the different point production measures.

Intention-based imitation tasks. Children at 12 and 15 months were seated on their accompanying parent's lap on the floor of a university playroom. At 18 months, they were seated in a high chair opposite the E at a rectangular table measuring 78 cm in height, 120 cm in width, and 78 cm in depth.

Intention-based imitation task at 12 and 15 months. In the modified version of the reenactment task[1] at 12 and 15 months, two object sets were used. To allow for some variance, two objects with a medium degree of difficulty were chosen (on the basis of Meltzoff, 1995). The object sets were (1) a plastic Fisher Price dumbbell-like bathing toy that could come apart into two pieces and (2) a loop that could be placed around a horizontally protruding prong. Objects were presented in the following fixed order: (a) dumbbell baseline condition, (b) dumbbell test condition, (c) loop and prong baseline condition, (d) loop and prong test condition. To account for object affordances as an alternative explanation of infants' actions in the reenactment task (see e.g., Huang, Heyes, & Charman, 2002), we included a baseline condition for 12- and 15-month-olds, because data on performance in control conditions are rare in experimental studies using the reenactment paradigm at a comparable age. At 18 months, we conducted the original demonstrate intention condition using the same five object sets as in the original study by Meltzoff (1995) for which sufficient data on control conditions are available. In the baseline condition, E handed the toy to the child and timed a 20-s period for each object.

To ensure that all children had received two clean imitation trials, for the children ($n < 5$ at both ages) who performed the act spontaneously in one of the baseline conditions, both respective test conditions were not analyzed. After each baseline condition, the child was given the demonstration of intention condition: E modeled the intention to perform the target act for the child and was seen to try but fail to perform the target act on each of the two objects, while displaying a concentrated facial expression and providing no linguistic or facial expressions of failure. For the loop and the prong, E tried but failed in an attempt to put the nylon loop over the prong. E picked up the loop, but as the child approached the prong E released it inappropriately so that it "accidentally" dropped to the table surface each time. First, the loop was released slightly too far to the left, then too far to the right, and finally too low, where it fell to the table directly below the prong. The goal state of draping the nylon loop over the prong was not demonstrated. For the dumbbell, E picked it up by the wooden cubes. It appeared that E was trying to pull the ends outward. However, E failed to do so because while pulling, one of E's hands slipped off the end of the object. The direction of slippage alternated from left, to right, to left over the three stimulus presentations (the spatial terms are all referenced from the child's viewpoint). Thus there was no object transformation, and the goal state was never achieved. All that was visible were E's attempts to pull it apart. In both conditions, the child thus never saw the object's end state. E then offered the object set to the child and timed a 20-s period. Following Slaughter and McConnell (2003) and Bellagamba and Tomasello (1999), some infants, who first explored the object with their mouth, were given about 10 s extra.

Children were coded as having no competence (Level 0) when they did not perform the target act on either one of the objects and beginning competence (Level 1) when they performed the target act on one of the objects. They were rated as competent (Level 2) when they performed the target act on both objects. Scores ranged from 0 to 2.

Intention-based imitation task at 18 months. Eighteen-month-olds were presented with wooden duplicates of the five object sets, as used by Meltzoff (1995). They were given the demonstration of intention condition for each of the five objects in the following fixed order: dumbbell, buzzer box, loop and prong, cylinder and beads, and square with hole (for exact procedures for the five objects, see Meltzoff, 1995). At all measurement points, we used the same scoring criteria (see previous section) for children's target acts as those given in the original work by Meltzoff (1995). Infants were coded as either having or not having produced the target behavior. Because there were five objects, scores ranged from 0 to 5. To make the results for intention-based imitation at 18 months comparable to children's scores at 12 and 15 months, on the basis of the competence levels in the study by Bellagamba and Tomasello (1999) and children's median score at 18 months, the raw scores were recoded into three competence levels as follows: 0 (*raw score 0 to 2*), 1 (*raw score 3*) and 2 (*raw score 4 or 5*). At each age, a random sample of approximately 25% of the children was scored by an independent observer. Cohen's κ for all imitation scores ranged from .95 to 1.00.

Working memory skills. Working memory skills were used as a control measure in all partial correlations involving joint attention skills and intention-based imitation skills. The working memory task was modeled after Reznick, Morrow, Goldman, and Snyder (2004), using six instead of 12 trials. Infants sat on their caregiver's lap facing a frame containing two openings, which were side to side, 42 cm apart from center to center. Two curtains were attached to the back of the frame, designed to cover the windows. At the beginning of each trial, E pulled aside two curtains, put her face in one of the windows, and engaged the infant's attention. E then withdrew her face, replaced the two curtains, and wiggled her fingers at the top center of the frame. As soon as the infant looked toward the fingers, E reopened the curtains, and after a 2- to 3-s pause, she reappeared in her previous location. The curtains were then closed again. After a short pause, E reopened the curtains to initiate the next trial. E's location of appearance was counterbalanced between the left and right windows, and the procedure lasted for six trials (for a detailed de-

[1] A recent study by Nielsen (2009) indicated that when conducting the reenactment paradigm using novel objects, designed to enable young infants to readily identify an actor's goal, children as young as 12 months can infer the goal of an intended but failed attempt. In contrast, using the original objects, infants are able to master the task at 15 months (Bellagamba & Tomasello, 1999). Thus, to make the results comparable to those of other studies using the Meltzoff (1995) task (see e.g., Camaioni et al., 2004), although we modified the task procedure, we used objects that closely resembled those in Meltzoff's original version.

scription of the apparatus and procedure, see Reznick et al., 2004). Videotapes were coded by a research assistant, who assessed the direction of the infant's first gaze after the reopening of the curtain. Children were given a score of 1 if they looked toward the cued direction and a score of 0, if their gaze was directed toward some other location. Scores thus ranged from 0 to 6. A random sample of 25% children was coded by a second observer. Cohen's κ was .72.

Mental state language. Twenty-four-month-olds' and 36-month-olds' mental state language was assessed using a parent report measure by Olineck and Poulin-Dubois (2005). The Mental State Language Questionnaire (MSLQ) included 78 internal state words adapted from Bretherton and Beeghly (1982). Based on a linguistic coding scheme for mental state terms comprising additional terms (Klann-Delius, 1998), it was translated into German. The final, adapted version used at both 24 and 36 months consisted of 85 mental state terms from seven different categories.[2] There were 12 cognition terms (e.g., *think),* 39 emotion terms (e.g., *sad),* 3 volition/desire terms (e.g., *want),* 4 perception terms (e.g., *see*), 8 moral/obligation terms (e.g., *must),* 2 ability terms (e.g., *hard to do*), and 17 physiology terms (e.g., *hungry*). In this study, the volition/desire and cognition subscales were used.

Parents were asked to indicate for each word whether their child made active use of the respective mental state term. Children's percentage scores for volition and cognition terms were calculated by dividing the number of inner state terms children actively use by the total number of items included in the subscales.

Language. When children were 18 months of age, mothers filled out the Elternfragebögen für die Früherkennung von Risikokindern (Parent-Report Screening Instrument for At-Risk Children; Grimm & Doil, 2006), a standardized parent report measure on 12- to 24-month-olds' receptive and productive language skills based on the MacArthur Communicative Development Inventory (Fenson et al., 1993) containing 164 items.

Data Analysis

In this study, the hypothesized relations were examined using correlational and regression analyses. For all correlation analyses, *p* values are reported. Due to floor and ceiling effects, in some measures skewness statistics exceeded −0.8 or 0.8 and the kurtosis was higher or lower than 3. In particular, point following at 12 months and declarative point production at 18 months proved to be extremely negatively skewed, and gaze following at 7 months, as well as declarative point production at 9 months, proved to be extremely positively skewed. Thus, these measures were not considered as predictor variables. Also, because of the ordinal nature of the variables and the limited range of the measures, nonparametric data analyses were deemed more appropriate than parametric analyses. Thus, Spearman rank correlations were conducted. To predict intention-reading skills, a binary-logistic regression was conducted. Concerning the prediction of volition and cognition terms, a linear regression model was conducted, using dichotomous joint attention scores as independent dummy variables, taking into account their nonparametric, noncontinuous nature. For this purpose, the raw scores were recoded as follows: Concerning referential gaze and point comprehension, to receive a passing score of 1, children had to follow the tester's gaze and point to the object. If children only looked at E's finger or face or did not interact with E, they received a score of 0. In the case of imperative point comprehension, children had to give the object to E, leave it on the table for E, tease E, or communicate their refusal to respond to E's request by protesting (e.g., headshake) to receive a passing score of 1. For declarative point production, to receive a passing score of 1, children had to display a pointing gesture in addition to at least one of the following target behaviors: smiled and/or vocalized toward the stimulus, produced a proto-word or a word relative to the stimulus, or reenacted what the stimulus did. In the case of imperative pointing, children were awarded a passing score of 1 if they intentionally used a pointing gesture to request an object, in addition to at least one of the following target behaviors: leaned forward or reached for the stimulus, produced a request vocalization/proto-word, or whined.

The specific variables chosen for the regression analyses were those that displayed significant or marginally significant correlations with the outcome scores. Marginally significant variables were included to examine more fully the possible effects of multicollinearity among the predictor variables (cf. Mundy et al., 2007; Singer & Willett, 2003).

Although working memory skills at 7 months were related to point-following skills at 9 months, $\rho(59) = .49$, $p < .001$, they were not related to outcome measures and were thus not included in the regression models reported in this study (Baron & Kenny, 1986; Shrout & Bolger, 2002). To control for possible effects of working memory on the predictor variables, regression models including working memory were also calculated. No effect of working memory on the relations between joint attention and outcomes measures was found.

Results

The descriptive developmental results for the target measures were consistent with the literature and are not reported in detail for space reasons (see Tables 1, 2, and 3, for an overview and Table 4 for correlations between joint attention skills and outcome measures at 24 and 36 months).

Relations Between Joint Attention and Intention-Based Imitation Skills

It was hypothesized that infants' comprehension of referential intent and their production of declarative gestures, but not their comprehension and use of imperative gestures, would predict infants' intention-based imitation skills.

Correlational analyses revealed that the comprehension of referential pointing at 9 months, $\rho(38) = .34$, $p < .05$, was related to intention-based imitation skills at 15 months. Furthermore, production of declarative pointing at 15 months, $\rho(51) = .26$, $p < .10$, was marginally significantly related to children's intention-based imitation skills at 15 months. In contrast, as predicted, no signif-

[2] The term *may* had to be excluded from the questionnaire because it has no translation equivalent in German. Further, to make the German version comparable to the English version, five items that had been rated as polysemous (e.g., *must*, which expresses a feeling of obligation and thus belongs to the emotion category, although it is also a moral term) on the basis of Bretherton and Beeghly (1982) were sorted into one category. The complete list of items can be found in the Appendix.

Table 1
Percentage of Children at Different Competence Levels for Attention Following, as Well as Comprehension and Production of Declarative and Imperative Pointing Tasks

Task and age	0	1	2
Gaze following			
7 months (n = 62)	75.8	22.6	1.6
9 months (n = 69)	59.4	23.2	17.4
12 months (n = 83)	12.0	6.0	82.0
Point following			
7 months (n = 63)	14.3	39.7	46.0
9 months (n = 67)	10.4	15.0	74.6
12 months (n = 83)	3.4	9.8	86.8
Declarative point production			
9 months (n = 82)	93.9	3.7	2.4
12 months (n = 81)	51.9	24.7	23.4
15 months (n = 74)	37.8	17.6	44.6
18 months (n = 57)	7.0	5.3	87.7
Imperative point comprehension			
12 months (n = 63)	20.7	0.0	79.3
Imperative point production			
12 months (n = 86)	26.7	43.0	30.3
15 months (n = 78)	10.3	41.0	48.7
18 months (n = 65)	0.0	21.5	78.5

icant correlations were found between comprehension and production of imperative pointing gestures and intention-understanding skills at any age. To further explore the earlier reported correlations, binary-logistic and hierarchical linear regression analyses were conducted.[3] To explore the specific relations between the comprehension and production of declarative pointing and intention-based imitation skills, a binary-logistic regression analysis was conducted. For this purpose, children's intention-based imitation competence levels were transformed into pass/fail scores (score 0 = fail, score 1 = pass, score 2 = pass) and were used as an outcome measure. The competence levels for point following at 9 months and point production at 15 months were entered into the regression equation as ordinal, independent variables. As can be seen in Table 5, only 9-month-olds' referential point following proved to be a significant predictor for intention-based imitation at 15 months.

Relations Between Joint Attention and Internal State Language

To analyze the differential relations between joint attention skills and mental state language, hierarchical linear regression analyses were conducted. Because differential relations between imperative and declarative pointing gestures and mental state talk were the main focus of this study, the order of the analysis

Table 2
Percentage of Children at Different Competence Levels Across Measurement Points for the Intention-Based Imitation Task

Age	0	1	2
12 months (n = 67)	67.2	29.8	3.0
15 months (n = 53)	41.5	39.6	18.9
18 months (n = 71)	11.3	29.6	59.1

Table 3
Means and Standard Deviations for Language and Working Memory Measures

Measure and age	*M*	*SD*	Range
Mental state language (proportions)			
24 months (n = 67)			
Desire	52.74	38.12	0–100
Cognition	11.94	18.98	0–92
36 months (n = 60)			
Desire	95.00	13.48	33–100
Cognition	61.67	29.32	0–100
General language skills (sum score)			
18 months (n = 60)	141.77	58.28	26–304
Working memory skills (sum score)			
7 months (n = 80)	3.88	1.03	1–6

followed the function of the pointing gesture (referential and declarative vs. imperative) and the hypothesized strength of the relation to mental state language at the respective age. Thus, the strategy was that when predicting 24-month-olds' mental state talk, which was correlated with both imperative and declarative skills, first, imperative skills were entered into the regression model, followed by referential and declarative skills. Further, because children's general language skills were hypothesized to be the most important predictor for mental state language, they were always entered at Step 1 of the regression model.

To test our hypothesis that children's desire psychology was selectively predicted by the comprehension of imperative pointing gestures, a hierarchical linear regression analysis was conducted, with 24-month-olds' volition talk as an outcome variable. At Step 1 of the hierarchical linear regression model, general language skills at 18 months were entered into the model and accounted for 19% of the variance. At Step 2, when children's comprehension and production of an imperative motive at 12 months were entered into the equation, this accounted for 26% of additional variance in children's volition vocabulary.

Although the contribution of general language skills remained significant, children's comprehension of imperative pointing independently and highly significantly predicted 24-month-olds' volition vocabulary. In contrast, 12-month-olds' imperative point production did not make an independent contribution to the prediction of 24-month-olds' volition vocabulary, over and above comprehension skills and language and was therefore excluded from the model. The influence of general language was reduced to marginal significance when, at Step 3 of the regression model, referential and declarative gestures were included, which accounted for 10% of additional variance. Point-following skills at 9 months made an independent contribution to the prediction of desire terms. However, gaze-following skills at 9 months and 12-month-olds' declarative point production skills did not make a significant, independent contribution to the prediction of toddlers' active use of desire terms and were thus excluded from the model (see also Table 6).

[3] In all regression analyses reported in this article, missing values were deleted listwise. In the case of regression analyses, this method is more robust than were casewise deletion methods to violations of the missing at random assumption (Meyers, Gamst, & Guarino, 2006).

Table 4
Correlations Between Predictor Variables (Joint Attention Skills) and Outcome Measures (Mental State Language)

Joint attention skills	Desire at 24 months	Cognition at 36 months
Gaze following		
9 months	.27† (52)	.06 (46)
12 months	.04 (63)	−.12 (56)
Point following		
7 months	.05 (48)	−.03 (45)
9 months	.30* (51)	.48** (44)
Declarative point production		
12 months	.29* (62)	.23† (55)
15 months	.05 (59)	.11 (55)
Imperative point comprehension		
12 months	.29* (50)	−.04 (46)
Imperative point production		
12 months	.25* (66)	.12 (59)
15 months	−.07 (61)	−.05 (58)
18 months	.19 (54)	.18 (55)

Note. Spearman rank correlations. *n* is given in parentheses.
† $p < .10$. * $p < .05$. ** $p < .01$.

Next, the specific relations between joint attention skills and children's use of epistemic state terms at 36 months were assessed. When entering the predictor variables in a hierarchical linear regression analysis, while accounting for general language skills, as shown in Table 7, the equation was significant at Step 1, with general language skills accounting for 48% of the variance, whereas when 9-month-olds' point following and 12-month declarative point production were entered into the equation at Step 2, combined they accounted for 6% of additional variance. However, only 9-month point following proved to be an independent, significant predictor of cognition talk at 36 months (see also Table 7).

In sum, consistent with hypotheses, both children's comprehension of reference and their comprehension of imperative pointing gestures were important predictors of children's use of desire terms at 24 months. Further, children's imperative skills were not significantly related to 36-month-olds' use of cognition terms, whereas children's comprehension of reference was the most important predictor for children's use of cognition terms at 36 months.

Discussion

The present study aimed at detecting developmental relations among different dimensions of joint attention at the ages of 7, 9, 12, 15, and 18 months; intention-based imitation skills in the second year; and volition and cognition talk at the ages of 24 and 36 months, respectively. Evidence was obtained for both general and specific developmental relations between joint attention and later social cognitive development. In particular, individual differences in the ability to comprehend another's referential intent were predictive of both intention reading and later mental state talk, suggesting that the comprehension of referential intent, a fundamental aspect of the joint attentional frame, is, in fact, the basis for both an implicit and an explicit understanding of the mind. Furthermore, the findings support the view that imperative joint attentional skills are specific developmental precursors of later explicit talk about desires.

Table 5
Binary-Logistic Regression Analysis for Variables Predicting 15-Month Intention Reading (N = 37)

Variable	*B*	Wald	Odds ratio
Step 1			
Point following at 9 months	1.719	3.871	5.581*
Declarative point production at 15 months	−0.218	0.433	0.804

* $p < .05$.

Table 6
Hierarchical Linear Regression Analysis for Variables Predicting 24-Month Volition Talk (N = 31)

Variable	β	*t*	*p*
Step 1[a]			
General language at 18 months	.44*	2.69	.01
Step 2[b]			
General language at 18 months	.37*	2.70	.01
Imperative point comprehension at 12 months	.52**	3.74	.00
Imperative point production at 12 months[c]	−.06	−0.43	.67
Step 3[d]			
General language at 18 months	.25†	1.82	.08
Imperative point comprehension at 12 months	.62**	4.66	.00
Imperative point production at 12 months[c]	−.01	−0.05	.96
Gaze following at 9 months[c]	−.03	−0.21	.84
Point following at 9 months	.35*	2.54	.02
Declarative point production at 12 months[c]	.04	0.30	.77

[a] $R^2 = .19$, $\Delta R^2 = .19$; $F(1, 30) = 7.25$, $p = .01$. [b] $R^2 = .45$, $\Delta R^2 = .26$; $F(2, 29) = 12.19$, $p = .00$. [c] Excluded variable. [d] $R^2 = .55$, $\Delta R^2 = .10$; $F(3, 28) = 11.82$, $p = .00$.
† $p < .10$. * $p < .05$. ** $p < .01$.

Relations Between Joint Attention and Intention-Based Imitation Skills

Consistent with Carpenter et al. (1998) and Camaioni et al. (2004), individual differences in infants' developing ability to follow a referential, declarative pointing gesture and to produce a declarative pointing gesture were related to intention-based imitation skills in the present study. A regression analysis revealed that

Table 7
Hierarchical Linear Regression Analysis for Variables Predicting 36-Month Cognition Talk (N = 34)

Variable	β	*t*	*p*
Step 1[a]			
General language at 18 months	.69**	5.50	.00
Step 2[b]			
General language at 18 months	.57**	4.29	.00
Point following at 9 months	.28*	2.11	.04
Declarative point production at 12 months[c]	.10	0.76	.45

[a] $R^2 = .48$, $\Delta R^2 = .48$; $F(1, 33) = 30.24$, $p = .00$. [b] $R^2 = .54$, $\Delta R^2 = .06$; $F(2, 32) = 18.92$, $p = .00$. [c] Excluded variable.
* $p < .05$. ** $p < .01$.

over and above declarative point production, point following at an early age (9 months) predicted intention reading half a year later. Infants' capacity to follow a pointing gesture toward a specific object reflects a sensitivity to the referential function of the pointing gesture.

As was pointed out in the introduction, referential point following is to be seen as distinct from children's communicative understanding of a declarative motive (see Aureli et al., 2009). Hence, the present findings both support and extend Camaioni et al.'s (2004) longitudinal findings by showing that the ability to infer an action goal by identifying a target referent appears to be a powerful predictor of later inferences about unseen action goals, regardless of a full communicative comprehension of declarative motives. Further, consistent with Camaioni et al., the comprehension and production of imperative gestures was not related to intention reading. This is consistent with the view that the comprehension of declarative motives and the preceding comprehension of referential intent (reference being directed to objects outside the shared visual field) pave the way toward inferring information about objects or events, such as a person's unseen action goals, unlike the imperative pathway that is primarily relevant to directing others' behavior.

In contrast to point following, gaze following to distal objects was not related to later intention reading. Because adults look at objects in everyday life without necessarily intending their infant to share an experience with them, one possible explanation is that because pointing is less frequent and thus more exclusively related to caregivers sharing information, infants may come to see it as being a more reliable marker for intentional behavior in others than is gaze shifting alone. Only later may children come to perceive direction of gaze, independently of other verbal and nonverbal signals, as an intentional, communicative act (see also Leavens, 2006). Consistent with this interpretation, other experimental studies and naturalistic observations have shown that infants between 5 and 10 months of age are far more likely to follow an adult's gaze when it is accompanied by a point than when presented without any further communicative cues (Deák, Wakabayashi, Sepeta, & Triesch, 2004; Flom, Deák, Phill, & Pick, 2004). Furthermore, gaze-following behavior can be found in a variety of species (see Gomez, 2005, for a review), whereas declarative pointing seems to be an exclusively human social activity (see e.g., Liszkowski, Schäfer, Carpenter, & Tomasello, 2009).

Relations Between Joint Attention Skills and Internal State Language

Because mental state talk is rare in children's spontaneous speech, the present study assessed children's mental state vocabulary by using an adapted version of the MSLQ (cf. Bretherton & Beeghly, 1982; Olineck & Poulin-Dubois, 2005; Poulin-Dubois, Chiarella, & Polonia, 2009). Harris and Jones (1997) pointed out that parental reports, in particular when parents are presented with a set of possible words their children might use, are quite reliable. In general, even techniques that require active recall instead of mere recognition (e.g., mothers' vocabulary diaries) were found to be reliable and valid, with highly positive relations to direct observations of infant language (see e.g., Dale, Bates, Reznick, & Morisset, 1989; Snyder, Bates, & Bretherton, 1981).

Regarding the longitudinal relations with children's joint attention skills, overall there was evidence for a predictive relation between the early comprehension of referential pointing (at 9 months) and volition and cognition talk at 24 and 36 months, respectively. Furthermore, specific relations were found between the preverbal comprehension of imperative motives and the emergence of talk about desires. Children's comprehension of an imperative motive and their imperative production skills, consistent with our hypothesis, were related to 24-month-olds' talk about desires only. The finding that imperative point comprehension at 12 months predicted desire talk is consistent with findings by Colonnesi et al. (2008), who reported a longitudinal association between imperative point comprehension and psychological action explanation (in desire terms) at 39 months. Grasping imperative motives requires infants to show an interest in the tester's wish and, most important, to grasp the tester's desire to obtain an object from them. Thus, imperative point comprehension can be seen as a precursor to understanding the subjectivity of desires, in the sense of identifying the object that the other person's desire is about independently of one's own desire (e.g., to retain that object). Further, talk about desires also appears to require an understanding of another person's communicative, referential intent. Whereas imperative bids in infancy are typically directed toward concrete objects in the shared visual field, desire reasoning in young children increasingly refers to past and future events and other mentally represented entities (cf. Bartsch & Wellman, 1995). Thus, consistent with our predictions, imperative competencies converge with early referential joint attentional skills in predicting talk about desires.

In sum, talk about desires at the age of 2 years emerges from both a preverbal grasp of referential intent and a preverbal understanding of imperative motives. The link between the comprehension of referential and imperative gestures and children's desire vocabulary at 24 months may reflect the developmental path from infants' joint attentional skills toward children's early nonpropositional desire psychology. By the age of 2 children function with a desire psychology that affords them the ability to predict actions and to make genuine use of desire terms (Wellman & Woolley, 1990). Thus, although 2-year-olds tend to use most other mental state terms (e.g., cognition terms) primarily as a conversational tool rather than as genuine references to mental states, it seems likely that 24-month-olds' use of volition terms can be seen as a valid measure of children's early desire psychology.

This interpretation is also consistent with functional interpretations of inner state expression, assuming that, in analogy to pointing gestures, children acquire internal state terms to communicate the range of psychological relations they experience, which first are directed toward real objects (e.g., wanting things) and persons or perceptible states (e.g., feeling warm) and then become more abstract and complex (referring to propositions) with age (Moore, 2006). Whereas the meaning of such words as *want* may at least sometimes be manifested in the ostensive context (e.g., the speaker might use a requestive gesture when *want* is uttered), it is hard to imagine how the meanings of cognitive terms such as *think* and *know* may be

ostensively available, rather than abstract and removed from purely sensory experience (cf. Gleitman, 1990).

Thirty-six-month-olds' cognition vocabulary proved to be unrelated to children's comprehension of imperative gestures but was selectively predicted by children's comprehension of reference. It is noteworthy that 12-month-olds' declarative point production was marginally significantly related to 36-month-olds' cognition vocabulary, and this relation did not hold up over and above children's general language skills, whereas referential point-following skills at 9 months, independently of general language skills, were the most important predictor of 36-month-olds' use of cognition terms. Thus, infants' comprehension of reference seems to be at the core of the relation between joint attention and mental state talk at the age of 3 years. Bartsch and Wellman (1995) showed that just prior to their third birthday, children begin to use cognitive verbs in contexts that demonstrate their understanding of meaning rather than simply mimicking adults or using words only to regulate interactions. Thus, at 36 months of age, children's mental state vocabulary can be seen as a reflection of children's beginning genuine comprehension of epistemic state terms. However, it is only later, during the preschool years, that children's use of mental state terms is directly linked to children's theory of mind skills and reflects children's deeper level social understanding. Future research should examine predictive relations between joint attention skills and children's belief–desire psychology beyond the age of 3 years.

Further, it seems important to examine how children's comprehension of pointing gestures might aid them in developing a mental state vocabulary. Joint attention episodes do not occur in a vacuum. Rather, infants acquire competence in engaging with others in situations of shared attention where they develop expectations about others' attentiveness or directedness toward aspects of the world (see e.g., Carpendale & Lewis, 2006). Thus, the types of social exchanges involving prelinguistic gestures may serve as a foundation of children's acquisition of mental state language, because they provide the context to extend, refine, and verbally elaborate on prelinguistic concepts (Malcolm, 1991). For instance, recent research (Slaughter, Peterson, & Carpenter, 2009) supports the notion that desire terms such as *want* emerge in a child's lexicon within contexts of obtainment. Within such contexts, mothers were found to comment on volitional states, which may then promote children's acquisition of desire terms. Likewise, caretakers may request an object from the child and talk about their own and the child's desires (e.g., "I can see you want to keep that, but mommy needs it too!"). Similarly, if children attend to caretakers' referential gestures, in such contexts caretakers may refer to epistemic states (e.g., "Look! Do you remember this picture?").

Thus, although highlighting the role of referential comprehension, findings of relations between joint attention skills and mental state language are also consistent with contextual views of mental verb development emphasizing the social nature of language (see e.g., Montgomery, 2002; Tomasello, 1992).

In sum, the results of this study provide the first longitudinal evidence for theoretical arguments that the comprehension of proto-imperatives may serve as a crucial prelinguistic precursor to the development of desire verbs (Montgomery, 2002; Moore, 2006), over and above general language skills. Further, by showing a strong link between infants' comprehension of reference at the age of 9 months and their intention-reading skills at 15 months, as well as between comprehension of reference at 9 months and both children's early (at the age of 24 months) and later (at the age of 36 months) mental state vocabulary, our results emphasize the critical role that infants' ability to establish a joint attentional frame in their object-directed interactions with adults plays for semantic development.

References

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

Aureli, T., Perucchini, P., & Genco, M. (2009). Children's understanding of communicative intentions in the middle of the second year of life. *Cognitive Development, 24,* 1–12. doi:10.1016/j.cogdev.2008.07.003

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

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychology research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. doi:10.1037/0022-3514.51.6.1173

Baron-Cohen, S. (1989). The autistic child's theory of mind: The case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30,* 285–298.

Bartsch, K., & Wellman, H. (1995). *Children talk about the mind.* New York, NY: Oxford University Press.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly, 21,* 205–226.

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12- and 18-month-olds. *Infant Behavior and Development, 22,* 277–282. doi:10.1016/S0163-6383(99)00002-8

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18,* 906–921. doi:10.1037/0012-1649.18.6.906

Butterworth, G., & Jarrett, N. (1991). What minds have in common in space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55–72.

Camaioni, L. (1993). The development of intentional communication: A re-analysis. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicational development.* London, England: Routledge.

Camaioni, L., Perucchini, P., Bellagamba, F., & Colonnesi, C. (2004). The role of declarative pointing in developing a theory of mind. *Infancy, 5,* 291–308.

Camaioni, L., Perucchini, P., Muratori, F., Parrini, B., & Cesari, A. (2003). The communicative use of pointing in autism: Developmental profile and factors related to change. *European Psychiatry, 18,* 6–12.

Carpendale, J. I. M., & Lewis, C. (2006). *How children develop social understanding.* Oxford, England: Blackwell.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4, Serial No. 255).

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development, 15,* 481–498.

Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology, 26,* 561–577.

Dale, P. S., Bates, E., Reznick, J. S., & Morisset, C. (1989). The validity

of a parent report instrument of child language at twenty months. *Journal of Child Language, 16,* 239–249.

Deák, G. O., Wakabayashi, Y., Sepeta, L., & Triesch, J. (2004, May). *Development of attention-sharing from 5 to 10 months of age in naturalistic interactions.* Paper presented at the 14th Biennial International Conference on Infancy Studies, Chicago, IL.

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., . . . Reilly, J. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual.* San Diego, CA: Singular.

Flom, R., Deák, G. O., Phill, C. G., & Pick, A. D. (2004). Nine-month-olds' shared visual attention as a function of gesture and object location. *Infant Behavior and Development, 27,* 181–194.

Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language, 23,* 307–336.

Gleitman, L. (1990). The structural sources of verb meaning. *Language Acquisition, 1,* 3–55.

Gomez, J. C. (2005). Species comparative studies and cognitive development. *Trends in Cognitive Sciences, 9,* 118–125.

Grimm, H., & Doil, H. (2006). *Elternfragebögen für die Früherkennung von Risikokindern (ELFRA-1, ELFRA-2)* [Parent-report screening instrument for at-risk children (ELFRA-1, ELFRA-2)] (2nd ed.). Göttingen, Germany: Hogrefe.

Hadwin, J., Baron-Cohen, S., Howlin, P., & Hill, K. (1997). Does teaching theory of mind have an effect on the ability to develop conversation in children with autism? *Journal of Autism and Developmental Disorders, 27,* 519–537.

Harris, P. L., & Jones, P. H. D. (1997). Commentary on "Talking about feelings" (Aldridge & Wood, 1997). *Child Abuse and Neglect, 21,* 1217–1220.

Huang, C., Heyes, C., & Charman, T. (2002). Infants' behavioral enactment of "failed attempts": Exploring the roles of emulation learning, stimulus enhancement, and understanding of intentions. *Developmental Psychology, 38,* 840–855.

Klann-Delius, G. (1998). *Coding scheme for mental state terms.* Unpublished manuscript, Department of Linguistics, Free University of Berlin, Berlin, Germany.

Leavens, D. A. (2006). It takes time and experience to learn how to interpret gaze in mentalistic terms. *Infant and Child Development, 15,* 187–190.

Leavens, D. A., & Hopkins, W. A. (1999). The whole-hand point: The structure and function of pointing from a comparative perspective. *Journal of Comparative Psychology, 113,* 417–425.

Legerstee, M., & Fisher, T. (2008). Coordinated attention, declarative and imperative pointing in infants with and without Down syndrome: Sharing experiences with adults and peers, *First Language, 28,* 281–311.

Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (2009). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science, 20,* 654–660.

Malcolm, N. (1991). The relation of language to instinctive behaviour. In J. Hyman (Ed.), *Investigating psychology: Sciences of the mind after Wittgenstein* (pp. 27–47). London, England: Routledge.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31,* 1–16.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation.* Thousand Oaks, CA: Sage.

Montgomery, D. E. (2002). Mental verbs and semantic development. *Journal of Cognition and Development, 3,* 357–384.

Moore, C. (2006). *The development of commonsense psychology.* Hillsdale, NJ: Erlbaum.

Moore, C., & D'Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development, 2,* 109–129.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Vaughan Van Hecke, A., & Parlade, V. M. (2007). Individual differences and the development of joint attention in infancy. *Child Development, 78,* 938–954.

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003). *A manual for the abridged Early Social Communication Scales (ESCS).* Available at http://www.ucdmc.ucdavis.edu/mindinstitute/ourteam/faculty_staff/escs.pdf

Mundy, P., & Gomes, A. (1998). Individual differences in joint attention skill development in the second year. *Infant Behavior and Development, 21,* 469–482.

Mundy, P., Hogan, A., & Doehring, P. (1996). *A preliminary manual for the abridged Early Social Communication Scales (ESCS).* Unpublished manuscript, Department of Psychology, University of Miami.

Myers-Walls, J. A., Myers-Bowman, K. S., & Posada, G. (2006). Parenting practices worldwide. In B. Ingoldsby & S. Smith (Eds.), *Families in global and multicultural perspective* (2nd ed.; pp. 147–167). Newbury Park, CA: Sage.

Nichols, K. E., Fox, N., & Mundy, P. (2005). Joint attention, self-recognition, and neurocognitive functioning. *Infancy, 7,* 35–51.

Nielsen, M. (2009). 12-month-olds produce others' intended but unfulfilled acts. *Infancy, 14,* 377–389.

Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' ability to distinguish between intentional and accidental actions and its relation to internal state language. *Infancy, 8,* 91–100.

Poulin-Dubois, D., Chiarella, S., & Polonia, A. (2009). Toddlers' vocabulary about the mind: A cross-linguistic study using the Internal State Questionnaire. *Rivista Di Psicolinguistica Applicata, 9*(3), 35–48.

Reznick, J. S., Morrow, J. D., Goldman, B. D., & Snyder, J. (2004). The onset of working memory in infants. *Infancy, 6,* 145–154.

Ruffman, T., Slade, L., Rowlandson, K., Rumsey, C., & Garnham, A. (2003). How language relates to belief, desire, and emotion understanding. *Cognitive Development, 18,* 139–158.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7,* 422–445.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis.* New York, NY: Oxford University Press.

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of working memory and false belief. *British Journal of Developmental Psychology, 23,* 1–26.

Slaughter, V., & McConnell, D. (2003). Emergence of joint attention: Relationships between gaze following, social referencing, imitation, and naming in infancy. *Journal of Genetic Psychology, 164,* 54–71.

Slaughter, V., Peterson, C., & Carpenter, M. (2009). Maternal mental state talk and infants' early gestural communication. *Journal of Child Language, 36,* 1053–1074.

Snyder, L. S., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565–582.

Sodian, B. (2011). Theory of mind in infancy. *Child Development Perspectives, 5,* 39–43.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development, 63,* 161–172.

Tomasello, M. (1992). The social bases of language acquisition. *Social Development, 1,* 67–87.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78,* 705–722.

Vaughan Van Hecke, A., Mundy, P., Acra, C. F., Block, J., Delgado, C., Parlade, M., . . . Pomares, Y. B. (2007). Infant joint attention, temperament, and social competence in preschool children. *Child Development, 78,* 53–69.

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology, 44,* 618–623.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & Lalonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science, 7,* 283–288.

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition, 35,* 245–275.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6,* 297–311.

Yamaguchi, M., Kuhlmeier, V. A., Wynn, K., & VanMarle, K. (2009). Continuity in social cognition from infancy to childhood. *Developmental Science, 12,* 746–752.

Appendix

Items for the German Version of the Mental State Language Questionnaire

Table A1

Item	*Fähigkeiten*	Abilities	*Physiologie*	Physiology	*Volition*	Volition	*Perzeption*	Perception
1	*schwer*	hard (to do)	*schmutzig*	dirty	*wollen*	want to	*anschauen*	look (at)
2	*können*	can	*weh tun*	hurt	*brauchen*	need	*sehen*	see
3			*heiß*	hot	*möchte*	like to	*hören*	hear
4			*müde*	tired			*beobachten*	observe
5			*schlafen/am Schlafen sein*	sleep/asleep				
6			*hungrig*	hungry				
7			*durstig*	thirsty				
8			*kalt*	cold				
9			*wach/aufgewacht*	wake/awake				
10			*warm*	warm				
11			*schmecken*	taste				
12			*sich anfühlen*	feel				
13			*riechen*	smell				
14			*schlecht sein*	sick				
15			*unordentlich*	messy				
16			*unwohl fühlen*	unwell				
17			*eklig fühlen*	yucky				

(*Appendix continues*)

Table A2

Item	*Emotion*	Emotion	*Moral*	Moral judgment/ obligation	*Kognition*	Cognition
1	*weinen*	cry	*können*	can	*wissen*	know
2	*küssen*	kiss	*dürfen*	may	*vergessen*	forget
3	*nicht mögen*	like not	*lassen*	allow	*verstehen*	comprehend
4	*gut gehen*	feel good	*müssen*	must	*vielleicht*	maybe
5	*Kuss*	kiss	*gut/brav*	good	*so tun als ob*	pretend
6	*lachen*	laugh	*sollen*	supposed to	*echt*	real
7	*lieb*	be good	*schlecht/böse*	bad	*glauben*	believe
8	*ernst*	grave	*gemein*	mean	*träumen*	dream
9	*schön*	beautiful			*meinen*	mean
10	*Angst haben*	scared			*raten*	guess
11	*mögen*	like			*denken*	think
12	*toll*	great			*erinnern*	remember
13	*umarmen*	hug				
14	*ok*	o.k.				
15	*besser*	better				
16	*traurig*	sad				
17	*lieben*	love				
18	*lustig*	funny				
19	*lieber*	like better				
20	*blöd*	stupid				
21	*schade*	a pity				
22	*gern*	like to do				
23	*Spaß haben*	have fun				
24	*eklig*	disgusting				
25	*wütend*	angry				
26	*komisch*	strange				
27	*Schätzchen*	sweetie				
28	*sich fühlen*	to feel				
29	*streiten*	fight				
30	*schlecht fühlen*	feel bad				
31	*lächeln*	smile				
32	*nett*	nice				
33	*überrascht*	surprised				
34	*in Ordnung sein*	be alright				
35	*glücklich*	happy				
36	*sich amüsieren*	have a good time				
37	*stolz*	proud				
38	*verrückt*	crazy				
39	*gruselig*	scary				

Received January 20, 2010
Revision received June 21, 2011
Accepted June 21, 2011 ■