

One-Year-Old Infants Follow Others' Voice Direction

Federico Rossano, Malinda Carpenter, and Michael Tomasello

Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Psychological Science
 23(11) 1298–1302
 © The Author(s) 2012
 Reprints and permission:
sagepub.com/journalsPermissions.nav
 DOI: 10.1177/0956797612450032
<http://pss.sagepub.com>



Abstract

We investigated 1-year-old infants' ability to infer an adult's focus of attention solely on the basis of her voice direction. In Studies 1 and 2, 12- and 16-month-olds watched an adult go behind a barrier and then heard her verbally express excitement about a toy hidden in one of two boxes at either end of the barrier. Even though they could not see the adult, infants of both ages followed her voice direction to the box containing the toy. Study 2 showed that infants could do this even when the adult was positioned closer to the incorrect box while she vocalized toward the correct one (and thus ruled out the possibility that infants were merely approaching the source of the sound). In Study 3, using the same methods as in Study 2, we found that chimpanzees performed the task at chance level. Our results show that infants can determine the focus of another person's attention through auditory information alone—a useful skill for establishing joint attention.

Keywords

social cognition, infant development, attention, inference, joint attention, voice direction, gaze following, referential communication

Received 10/4/11; Revision accepted 4/13/12

One of the key skills acquired during young children's cognitive and social-cognitive development is joint attention. Joint attention is foundational for children's earliest intersubjective social interactions, as well as for intentional communication, including linguistic communication (e.g., Carpenter, Nagell, & Tomasello, 1998). Virtually all of the research on this topic, however, has concerned children's joint visual attention, showing that children can follow and direct others' visual attention to external entities. But the fact that blind children find ways of entering into joint attention with others in order to learn language (Perez-Pereira & Conti-Ramsden, 1999) suggests that there may be routes into joint attention via other perceptual modalities.

One possible route is the auditory modality. Much research has shown that human infants and even many nonhuman primates follow the direction of others' gaze to external targets (e.g., Flom, Lee, & Muir, 2006; Tomasello, Call, & Hare, 1998). Although even newborn human infants are capable of localizing the sources of sounds (e.g., Alegria & Noirot, 1978; Muir & Field, 1979), we are aware of no research investigating infants' (or adults') ability to follow the direction of someone's voice to an external entity. This skill would be highly useful for humans because it would enable individuals engaging in collaborative activities to detect the direction of an

interactant's attention without looking up from what they themselves are doing. Therefore, we began by investigating the ability of 1-year-old infants to follow the direction of an adult's voice to an external target when they could not see the adult.

Study 1

In Study 1, we investigated whether infants could determine where an unseen adult was looking just by hearing her voice. An experimenter positioned herself behind a tall barrier and opened two boxes on the floor at either side of it; one of the boxes contained a toy (the infants did not know which one). Then, from behind the barrier (and therefore out of sight of the infants), the adult verbally expressed excitement while looking (with her head turned) toward one of the two boxes. We tested whether infants subsequently went to that box to get the toy.

Corresponding Author:

Federico Rossano, Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
 E-mail: federico_rossano@eva.mpg.de

Method

Participants. Thirty-two 16-month-olds (mean age = 16 months 0 days, range = 15 months 15 days to 16 months 15 days; 16 boys, 16 girls) participated. An additional 12 infants were tested but were excluded from the final sample because they were fussy ($n = 7$), they were distracted by their parent during the test ($n = 1$), or they did not move from their parent during one or more trials ($n = 4$). (Testing was stopped for infants who did not move on more than one trial.) Only infants whose parents said they could crawl or walk around a room independently participated in the study. Infants were recruited from a database of parents in a midsized German city who had volunteered to participate in child-development studies. The infants' parents came from various socioeconomic backgrounds.

Materials and design. The materials were a large wooden barrier (160 cm \times 0.9 cm \times 122 cm; see Fig. 1) and four pairs of matching cardboard boxes (each 34 cm \times 24 cm \times 17 cm). To reduce perseveration across trials, we decorated each pair of boxes differently from the others and used a different pair of boxes for each trial. On each trial, only one of the two boxes contained a toy (a ball, a car, a duck, or a fish); this was the box (hereafter referred to as the *correct box*) that the experimenter vocalized toward at test. To ensure that infants could not use visual cues or sound cues other than the experimenter's voice to find the toy, we placed it inside the box in such a way that it would not be visible until infants were quite close to the box, and immobilized it with small sponge walls so it could not make noise when the box was moved. Each infant was tested in four consecutive trials, with the correct box twice on the right and twice on the left, in fully counterbalanced order.

Procedure. During a warm-up play period, the experimenter played with each infant (with toys different from those used during the test) and encouraged the infant to move around the

room, including behind the barrier, to make sure he or she was not scared of the barrier or what might be behind it.

During the subsequent testing procedure, the infant sat or stood with his or her parent (who was sitting on a chair) 2 m from the center of the barrier. On each of the four trials, the experimenter first walked behind the barrier and showed the infant a pair of boxes by holding them above the center of the barrier. The experimenter next squatted behind the barrier (still facing forward), simultaneously placed one box on the floor on each side of the barrier, and then simultaneously opened the boxes. (The boxes were open so that, during the test, the experimenter could react as if she had just discovered the toy.) For the test, she stood up and called to the infant from above the center of the barrier ("I opened the boxes. Now pay attention!"). Then, while remaining at the center of the barrier and facing the infant, she again squatted behind it. While blocked by the barrier, she turned her head toward the correct box (see Fig. 2) and said excitedly, "Oh, wow! This is so nice! [Infant's name]! Come! Come here!" The infant could thus hear the experimenter's voice, which was directed toward one of the boxes, but could not see her at all (and the last time the infant had seen her, she had been facing the infant and at the center of the barrier). At this point, the infant was allowed to move freely, in whichever direction he or she chose.

If the infant did not move toward either box, the experimenter repeated her excited call toward the correct box. If the infant still had not moved after the experimenter had repeated the call three times (with a pause after each call), the trial was aborted. If the infant moved toward the correct box and tried to pick up the object, the experimenter gave it to the infant to play with briefly. If the infant went to the empty box, the experimenter showed the infant the toy in the other box without giving it to him or her.

Coding and reliability. All sessions were videotaped, and a coder who was blind to the location of the toy coded which box (left or right) each infant locomoted to during each trial. Infants were coded as locomoting to a box if they covered

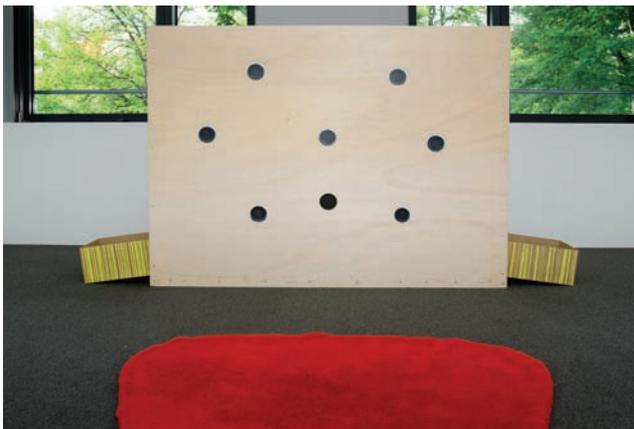


Fig. 1. Photograph of the experimental setup from the infants' perspective. Infants faced a barrier with one box on either side of it.



Fig. 2. Photograph showing the experimenter's position behind the barrier during the test phase in Study 1.

more than half the distance between their start position and the box (once infants were closer to the box, they could potentially see the toy in it, so any change in the direction of locomotion at that point might have been caused by the infant having spotted the toy). To assess interrater reliability, we had a second coder independently code data from a random sample of 25% of the infants. There was perfect agreement.

Results and discussion

Infants' level of performance was above chance even on the first trial (sign test, $p < .001$, $g = .32$), with 81% of the infants going to the correct box on that trial, despite not having any previous exposure to the task. Overall, across the four trials, the infants' level of performance was still significantly better than chance (mean number of correct trials = 2.56, $SD = 0.62$; Wilcoxon signed-rank exact test: $T^+ = 136$, $N = 16^1$, $p < .001$; $r = .68$), although some infants started perseverating to one side after the first trial. Eleven of the 13 infants who developed a side bias (i.e., infants who went to the same side on all four trials) had gone to the correct box on the first trial.

Our results suggested that 16-month-olds can follow another person's voice direction and can use it to infer what that person is attending to even if he or she cannot be seen. However, another explanation of the results is possible. As Figure 2 illustrates, while the experimenter vocalized toward the toy, her mouth was not at the center of the barrier, but rather was on the side of the barrier where the correct box was located. In addition, the experimenter had said, "Come here," which might have led the infants to move toward *her* rather than toward one of the boxes. This raises the possibility that the infants used a different (though still sound-related) cue to solve the task: the source of the sound, rather than the experimenter's voice direction. That is, the infants might simply have approached the source of sound, the experimenter, and come across the box with the toy on their way to her instead of following her voice direction. Study 2 was designed to address this possibility.

Study 2

In Study 2, we pitted the source-of-sound explanation against the voice-direction explanation. While the experimenter vocalized toward the correct box, her head was positioned closer to the box without the toy. Therefore, if infants were heading toward the source of the sound, we expected them to go to the box without the toy, whereas if they were following voice direction, we expected them to go to the correct box. We tested a group of 12-month-old infants, in addition to another group of 16-month-olds.

Method

Participants. Infants were recruited as they had been in Study 1. Thirty-six 12-month-olds (mean age = 12 months 16 days; range = 12 months 0 days to 12 months 29 days; 18 boys, 18

girls) and thirty-six 16-month-olds (mean age = 16 months 0 days; range = 15 months 15 days to 16 months 15 days; 18 boys, 18 girls) participated. An additional 19 infants (12-month-olds: $n = 8$; 16-month-olds: $n = 11$) were tested but excluded from analyses because they were fussy (12-month-olds: $n = 3$; 16-month-olds: $n = 2$), they were distracted by their parent during the test (12-month-olds: $n = 1$; 16-month-olds: $n = 2$), or they did not move from their parent on one or more trials (12-month-olds: $n = 4$; 16-month-olds: $n = 7$).

Procedure. The procedure for Study 2 was identical to that for Study 1 apart from one small but crucial modification. During the test, after squatting behind the barrier, the experimenter positioned herself on the side opposite the correct box (see Fig. 3) while looking toward the correct box and vocalizing excitedly. Thus, if infants were inclined simply to move toward the source of the sound (i.e., the experimenter) or to somehow use its location to locate the box with the toy, we expected their performance to be poor.

Coding and reliability. The coding procedure for Study 2 was identical to that for Study 1. To assess interrater reliability, we had a second coder code data from a random sample of 25% of the 12-month-olds and 25% of the 16-month-olds. There was perfect agreement.

Results and discussion

Again, the 16-month-olds' performance was significantly above chance, even on their first trial (sign test, $p < .01$, $g = .28$), with 78% of the infants going to the correct box. Overall, across the four trials, performance was also better than chance (mean number of correct trials = 2.53, $SD = 0.70$; Wilcoxon signed-rank exact test: $T^+ = 120$, $N = 15$, $p < .001$; $r = .60$), although again some infants developed a side bias. Thirteen of the 19 infants who showed a side bias had gone to the correct box on their first trial and then persisted in going to the same



Fig. 3. Photograph showing the experimenter's position behind the barrier during the test phase in Study 2.

Table 1. Distribution of Infants Who Went to the Correct Box on Each Trial in Studies 1 and 2

Trial	Study 1: 16-month-olds	Study 2: 16-month-olds	Study 2: 12-month-olds
Trial 1	81%	78%	61%
Trial 2	59%	67%	61%
Trial 3	66%	64%	58%
Trial 4	50%	44%	64%

side of the barrier. Our results thus show that 16-month-olds did not simply use the source of the sound to solve the task; instead, they followed the experimenter’s voice direction.

On their first trial, 61% of the 12-month-olds went to the correct box; this level of performance was not significantly above chance (sign test, $p = .24$, $g = .11$). Again, some infants perseverated to one side after the first trial, but only 7 of these 17 infants went to the correct box on the first trial. Still, the 12-month-olds’ overall level of performance across the four trials was significantly better than chance (mean number of correct trials = 2.44, $SD = 0.77$; Wilcoxon signed-rank exact test: $T^+ = 123$, $N = 16$, $p < .01$; $r = .50$). Thus, by 12 months of age, infants are at least beginning to be able to follow others’ voice direction.

Tables 1 and 2 present the percentages of infants who went to the correct box on each trial in Studies 1 and 2 and the distributions of correct responses for both studies. Additional analyses concerning the performance of infants who did not complete all four trials, measurements of the sound intensity of the experimenter’s voice, and the effects of trial number, counterbalancing order, and the number of calls infants listened to before they moved are presented in the Supplemental Material available online.

Study 3

In an attempt to determine whether humans’ closest primate relatives also have the ability to follow voice direction, we next tested chimpanzees.

Table 2. Distribution of Infants by Number of Correct Responses in Studies 1 and 2

Number of correct trials	Study 1: 16-month-olds	Study 2: 16-month-olds	Study 2: 12-month-olds
0	0%	0%	0%
1	0%	0%	6%
2	50%	58%	56%
3	44%	31%	28%
4	6%	11%	11%

Subjects were 16 chimpanzees (*Pan troglodytes*) of various ages (7 males, 9 females). These chimpanzees were used to having humans verbally encourage or direct them (e.g., to come to food or change cages) daily. The procedure for Study 3 was similar to that for Study 2. After a warm-up period in which each chimpanzee had to find food that had been visibly placed for him or her in one of the two boxes at the sides of the barrier, a human experimenter sat behind the barrier and repeatedly called the chimpanzee (“[Name], come here! Come and look!”) while facing the correct box from the opposite side of the barrier (as shown in Fig. 3). Each chimpanzee completed 12 trials. The placement of the food was randomized across trials (but the food was never in the same box more than twice in a row). Interrater agreement on data for a random sample of 25% of the chimpanzees was perfect.

The chimpanzees’ performance on the first trial was at chance level (sign test, $p = .80$), with 7 of the 16 chimpanzees going to the correct box. Overall performance across the 12 trials was also at chance level (mean number of correct trials = 6.19, $SD = 1.64$; Wilcoxon signed-rank exact test: $T^+ = 38$, $N = 11$, $p = .65$; $r = .11$; see Table 3). Interestingly, one chimpanzee went to the correct box on 10 of the 12 trials (binomial probability, $p < .05$). This 10-year-old male chimpanzee had been raised in a human home after birth and cared for by humans from age 1 to age 3. The only other chimpanzee that had some human rearing, a 12-year-old female who had been reared in a nursery, performed at chance level.

General Discussion

Our results for 16-month-old infants were clear: When the experimenter was out of sight behind the barrier and talked excitedly about the toy, they were able to determine which box the toy was in, even on their very first trial. Study 2 showed that infants were not just approaching the source of the sound and chancing upon the toy; rather, they were following the

Table 3. Distribution of Chimpanzees by Number of Correct Responses in Study 3

Number of correct trials (of 12)	Chimpanzees
0–3	0%
4	13%
5	25%
6	31%
7	6%
8	19%
9	0%
10	6%
11–12	0%

experimenter's voice direction to the toy. The performance of 12-month-olds was somewhat less robust, but it too was above chance level overall, even though these infants participated only in the more demanding Study 2, in which the experimenter's head (i.e., the source of sound) was closer to the incorrect box than to the correct box.

Thus, 1-year-old infants not only can discern the directionality of a voice without seeing the speaker, but also appear to be able to use voice direction to infer the speaker's focus of (visual) attention and, therefore, what he or she is referring to. Typically, voice direction and gaze direction coincide, so together they represent redundant pieces of information. However, when infants cannot see the face of a speaker—for example, when they are looking down playing with toys or have their back turned—the ability to follow voice direction helps them determine whether the speaker is talking to them and what he or she is referring to or focused on. As a group, chimpanzees showed no ability to follow voice direction, perhaps because they do not normally use vocalizations to establish joint attention with others in their natural environments. However, future research is needed to determine whether chimpanzees are able to follow voice direction in other contexts or with conspecifics.

The vast majority of research on infants' understanding of others' attention has focused on visual attention, documenting a quite sophisticated, referential understanding of others' gaze (e.g., Butler, Caron, & Brooks, 2000; Csibra & Volein, 2008; Moll, Koring, Carpenter, & Tomasello, 2006; Moll & Tomasello, 2004). The studies we report here demonstrate that infants may rely on other sources of information along with gaze direction to discern what other people are attending and referring to, and thus raise interesting questions about the representation of others' attention and the understanding of reference in infancy.

Acknowledgments

We thank Elena Rossi, Katja Buschmann, Johannes Grossmann, Roger Mundry, Marco Schmidt, Margarita Svetlova, Sven Grawunder, and Leonardo Lancia for help with the study. F. R. thanks the Freiburg Institute for Advanced Studies for support while he wrote the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information may be found at <http://pss.sagepub.com/content/by/supplemental-data>

Note

1. For this and all subsequent Wilcoxon signed-rank test results, the sample size is reduced because of tied observations.

References

- Alegria, J., & Noirot, E. (1978). Neonate orientation behaviour towards human voice. *International Journal of Behavioral Development, 1*, 291–312.
- Butler, S. C., Caron, A. J., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development, 1*, 359–377.
- 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), 1–143.
- Csibra, G., & Volein, A. (2008). Infants can infer the presence of hidden objects from referential gaze information. *British Journal of Developmental Psychology, 26*, 1–11.
- Flom, R., Lee, K., & Muir, D. (Eds.). (2006). *Gaze following: Its development and significance*. Mahwah, NJ: Erlbaum.
- Moll, H., Koring, C., Carpenter, M., & Tomasello, M. (2006). Infants determine others' focus of attention by pragmatics and exclusion. *Journal of Cognition and Development, 7*, 411–430.
- Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science, 7*, F1–F9.
- Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child Development, 50*, 431–436.
- Perez-Pereira, M., & Conti-Ramsden, G. (1999). *Language development and social interaction in blind children*. Hove, England: Psychology Press.
- Tomasello, M., Call, J., & Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour, 55*, 1063–1069.