Developmental Science 1:2 pp 233–238

REPORT

Whose gaze will infants follow? The elicitation of gaze following in 12-month-olds

Susan Johnson,[1] Virginia Slaughter[2] and Susan Carey[3]

1. *The University of Pittsburgh, USA*
2. *The University of Queensland, Australia*
3. *New York University, USA*

Abstract

Eighty-three 12-month-old infants faced a noisy, active, object for one minute, after which the object turned 45 degrees to the left or the right. Five conditions explored what object features elicited gaze-following behavior in the infants. In one condition, the object was an adult stranger. The other four conditions used a soft, brown, dog-sized, amorphously-shaped, asymmetrical novel object that varied along two dimensions theorized as central to the identification of intentional beings: facial features and contingently interactive behavior. Infants shifted their own attentional direction to match the orientation of the actor or object in every condition except the one in which the object lacked both a face and contingently interactive behavior. Infants' 'gaze'-following behavior in general, therefore, appears to have been driven selectively by a particular configuration of behavioral and morphological characteristics, specifically those theorized as underlying attributions of intentionality rather than attributions of person per se.

Most researchers agree that the human propensity to follow the gaze of others (*i.e.*, look where someone else is looking) is intimately related to the adult ability to attribute intentions to the gazer. However, since Scaife and Bruner (1975) first provided evidence that infants in their first year are able to follow the gaze of others, researchers have been unable to agree on whether this ability involves the same attributions by infants.

Infants as young as 3 months have now been found to follow the general direction of eye gaze under simplified conditions (Hood, Willen, & Driver, in press). Their ability to accurately locate the target of an adult's gaze develops slowly over the first two years of life (Butterworth, 1991). In addition, Corkum and Moore (1995) have demonstrated that infants are more likely to follow adult head turns (with closed eyes) than eye turns (with no head movement), though both cues together are better elicitors than either alone.

Despite the increasing information available on infants' ability to follow gaze, the fundamental question of its meaning remains unresolved. Some researchers argue that gaze-following implies the same attributions in infants as it does in adults, *i.e.*, the implicit attribution of a mind to the gazer. Several researchers have reported a relationship between referential abilities and gaze-following in older infants. Molares, Mundy, and Rojas (1997) showed a correlation between gaze-following at 6 and 8 months and receptive and expressive language between 12 and 18 months. Caron, Krakowski, Liu, and Brooks (1996) reported that 14-month-olds are sensitive to the presence or absence of potential attentional targets. Finally, Baldwin (1995) has shown that 18-month-olds use gaze to select the referents of novel words.

Other researchers, reluctant to accept this interpretation, have offered alternative explanations. Some have posited the existence of signal releasers (*e.g.*, directional movement of the head or eyes) that allow infants to share important information about the environment with caretakers without attributing intentionality to the gazer. Butterworth's ecological mechanism of early gaze-following depends heavily on infants comprehending the signal function of the adult's direction of gaze (Butterworth, 1991). Based on work with chimps,

Address for correspondence: Dr Susan Johnson, Department of Psychology, 405 Langly Hall, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Povenelli and Eddy (1996) also argue for an evolutionarily shaped ability to follow eye gaze in the absence of attributions of intentionality.

Others propose that extensive interactions with caretakers condition infants to anticipate interesting events in the direction of the caretakers' headturns (Moore & Corkum, 1995; Perner, 1991). In support of this position, Corkum and Moore (1995) demonstrated that gaze-following can be partially shaped by conditioning in 8- to 9-month-old infants who otherwise fail to follow gaze spontaneously.

Arguments and empirical work on both sides of the debate have one thing in common. They presuppose the role of people in the elicitation of gaze-following. However, insofar as gaze-following may reflect the attribution of a mind to the gazer, there is *no a priori* reason to believe the behavior should be restricted to interactions with people. There exist proposals of properties that are generally characteristic of entities with minds – people and animals alike – to which infants may be sensitive, *e.g.*, the presence of eyes, asymmetric body shape, contingent or goal-directed behavior, and self-generated behavior or movement (Baron-Cohen, 1995; Gergely, Nadasdy, Csibra, & Biro, 1995; Leslie, 1995; Premack, 1990). If infants follow the gaze of adults because they assume on the basis of some or all of the above characteristics that adults have minds, then there is no reason to presume they would not also follow the gaze of other entities evincing similar properties.

The discovery that people have no privileged status in the elicitation of infants' gaze-following would not immediately resolve the debate over its meaning to the infant. However, both the signal releaser and conditioning accounts would need revision. The plausibility of such revisions would depend in large part on the particular features found to elicit gaze-following.

The current study is therefore designed to explore the possibility that gaze-following can be spontaneously elicited in 12-month-olds by non-person entities which nonetheless have some of the qualities of intentional beings.[1] One person and four object conditions were run. The object conditions consisted of a single object varying on two dimensions; contingent behavior with or without a face (+C + F; +C − F) and non-contingent behavior with or without a face (−C + F; −C − F). In the person condition the face of a contingently-reacting actor was visible.

[1] These studies were inspired in part by a similar study by Movellan and Watson (1986).

Participants

Eighty-three full-term 12-month-olds were tested; fourteen in the +C + F condition (8 male, 6 female, *M* = 12-12; range 12-3 to 13-0) and in the −C + F condition (6 male, 8 female, *M* = 12-18; range 12-4 to 12-28). Twenty were tested in the + C-F condition (9 male, 11 female, *M* = 12-10; range 11-18 to 12-26) and in the −C − F condition (9 male, 11 female, *M* = 12-16; range 12-5 to 13-2). Fifteen were tested in the person condition (8 male, 7 female, *M* = 12-09; range from 11-25 to 12-29).

Method

Apparatus

The infant was seated on a caretaker's lap facing the actor/object across a diamond-shaped setup. A target was mounted in each of the two remaining corners of the diamond at the eye-level of the actor/object (see Figure 1). In the non-person conditions the object was mounted on a small table using a hidden turntable. A curtainwall behind the table hid the experimenters and camera from view. The experimental room had bare walls and was otherwise empty except for a covered visual cliff setup.

The object, covered in fake brown fur, was 14″ at its widest point, 12″ tall, and 18″ long. It had the approximate shape and orientation of a quadruped with no

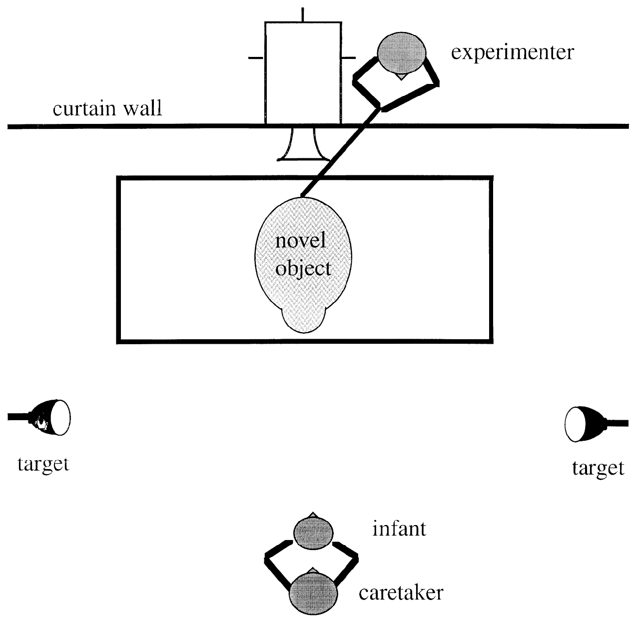

Figure 1 *Schematic of the experimental setup.*

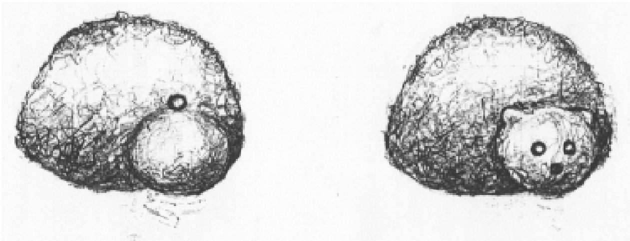

Figure 2 *Schematics of the novel object with the face (right) and without the face (left).*

distinct body parts although the face was set on a bulge at one end. A remotely controlled, battery-operated light and beeper were hidden inside. For the face conditions, two eyes (made of holes outlined in black felt), a black felt nose, and two rounded ears, were added. In the no-face conditions these features were absent and a single, front-facing, black-rimmed hole was instead placed on the object's 'body' directly above the now-faceless 'head' bulge (see Figure 2). The light shone through the holes when flashed. A hidden handle attached to the turntable allowed the experimenters to rotate the object to prefixed stops in the direction of the targets.

The targets were two black clip-on lamps with 60 watt incandescent bulbs.

Procedure

After seating the infant on the caretaker's lap, an assistant directed the infant's attention to each target by switching it on and off while calling for the infant's attention.[2] The lights were then left off for the remainder of the procedure. The assistant next turned to the actor/object and modelled a brief interaction that varied with condition. In all cases the assistant spoke as follows: 'Hi. How are you? I'm fine thanks. Bye bye,' followed by a goodbye wave. In the Person condition, the actor spoke and waved in response. In the contingent object conditions the object turned to the assistant, beeped in response to the speech, and flashed its lights in response to the wave. In the non-contingent conditions the object remained silent and unmoving. The assistant then left the room.

The experimental sequence consisted of a 60 second familiarization period, 2 looking trials, a 15 second familiarization period, and 2 more trials. During each familiarization period the still object sat 'facing' the infant. In the contingent object conditions (+C + F and +C − F) the object beeped in response to the infant's vocalizations and flashed its lights in response to the infant's movements. In the person condition, the beeps were replaced with naturalistic responses like 'um hum' or 'really' and the lights were replaced with winks. Individual infants in the non-contingent object conditions (−C + F and −C − F) were yoked to individual infants in the contingent object conditions (+C + F and +C − F) so that infants in each yoked pair experienced the same total amount and rate of the object's activity independent of their own behavior.

A looking trial consisted of the actor/object turning toward one of the two targets. Each infant saw four trials counterbalanced for order and direction. In all conditions, trials began with a long, attention-grabbing beep, followed by a smooth 45 degree turn towards the target. The actor/object remained silently oriented toward the target for approximately 8 seconds.

Coding of looking behavior

The looking behavior of each infant was videotaped and subsequently scored by a primary coder who was blind to the conditions and movement of the object. Secondary coders recoded 83% of the infants, achieving an average Spearman-Brown reliability coefficient of 0.92. In cases of disagreement, the primary coder's judgement was used.

All looks occurring within the trial window were coded. To code the direction of looks, an imaginary vertical midline was centered on the actor/object. Looks moving to the right of the midline anywhere in the vertical plane were coded as right looks. Similarly for looks to the left. All looks away were thus classified as either left or right looks with two exceptions. Smooth, uninterrupted looks moving directly from the actor/object to the caretaker were coded separately as were looks directly downwards along the midline.

Coding of contingency information

To ensure that the contingent conditions were sufficiently contingent, and to control for the possibility that infants assigned to a non-contingent condition nonetheless achieved some degree of contingency with the object by chance, the familiarization periods were coded for contingency between the infant and object. An object behavior which fell less than 1 second after an infant behavior was operationalized as contingent for these purposes. To satisfy the contingent condition, at

[2] Caretakers wore dark sunglasses to minimize their own view and to hide their eyes from the infants. Caretakers were instructed to avoid interacting with the infants.

least seventy-five percent of the infant's total behaviors had to be followed by a contingent object reaction. To be included in a non-contingent condition less than half of all the object's behaviors could fall less than 1 second after a behavior by the infant. Two of the eighty-three infants were excluded based on these criteria. On average 90% of the behaviors of the infants in the contingent conditions, and 14% of those of the infants in the non-contingent conditions received contingent object reactions.

Results

Comparisons among conditions were designed to answer three specific questions. First, were people the only objects in the study capable of eliciting gaze-following? Second, if infants followed gaze in any non-person condition, what were the relative contributions of the cues of contingency and facial features? And finally, if infants followed the gaze of the object most like a person, the contingent object with a face, how strong was the behavior relative to the person condition?

Preliminary analyses revealed that infants made an average of four looks away from the actor/object in either direction following each turn of the actor/object (*i.e.*, per looking trial), with 3.8, 4.4, 4.1, 3.5, and 3.9 looks in the person, +C + F, +C − F, −C + F, −C − F conditions respectively. A 2×2 analysis of variance (ANOVA) revealed no differences among the object conditions. One additional infant was eliminated for producing an average greater than three standard deviations above the overall mean, leaving a total of eighty infants for further analyses.

Infants received a difference score for each trial calculated by subtracting the total number of looks made in the unpredicted direction from those in the predicted direction. Looks were considered in the predicted direction if they followed the direction in which the actor/object turned and unpredicted otherwise. Analyses of variance showed no effects of looking trial so these were collapsed, generating a single difference score for each subject.

Two-tailed *t*-tests compared the mean scores for each condition to zero, the score expected by chance alone (Table 1). Infants were significantly more likely to look in the predicted direction when the gazer was (1) a person behaving contingently, $M = 1.62$, $S.D. = 2.06$, $t(14) = 3.04$, $p < 0.01$; (2) an object with a face, behaving contingently (+C + F), $M = 1.54$, $S.D. = 1.45$, $t(12) = 3.82$, $p < 0.005$; (3) an object without a face, behaving contingently (+C − F), $M = 1.35$, $S.D. = 1.76$, $t(19) = 3.33$, $p < 0.005$; or (4) an object with a face, behaving non-contingently (−C + F), $M = 1.18$, $S.D. = 1.07$, $t(12) = 3.97$, $p < 0.005$. Infants did not look preferentially in the predicted direction when the object lacked both a face and contingent behavior (−C − F), $M = -0.46$, $S.D. = 1.72$, $t(18) = -1.16$, ns.

Table 1 *Mean difference scores (Total looks in the predicted direction minus total looks in the unpredicted direction).*

Condition	Mean difference score	(S.D.)
person: +C + F ($n = 15$)	1.62**	(2.06)
novel object: +C + F ($n = 13$)	1.54***	(1.45)
novel object: +C − F ($n = 20$)	1.35***	(1.76)
novel object: −C + F ($n = 13$)	1.18***	(1.07)
novel object: −C − F ($n = 19$)	−0.46	(1.72)

*$p < 0.05$
**$p < 0.01$
***$p < 0.005$

Sign tests were used to examine the extent to which individual performances reflected the group data. Infants were categorized into three groups; those looking more in the predicted than unpredicted direction (difference score > 0), those looking equally in both directions (difference score = 0), and those with more looks in the unpredicted than predicted direction (difference score < 0) (see Table 2). These analyses paralleled the group data exactly. In all but the non-contingent, faceless object (−C − F) condition, more babies looked preferentially in the predicted direction than expected by chance, $p < 0.05$.

We next addressed the extent to which either contingency or the presence of a face alone contributed to the looking behavior. We performed a 2 (contingent vs. non-contingent) × 2 (face vs. no-face) ANOVA on difference scores in the four object conditions. Main effects were found for both variables; contingency, $F(1, 61) = 7.42$, $p < 0.01$, and the presence of facial features, $F(1, 61) = 5.24$, $p < 05$. This finding confirms that each feature is independently capable of eliciting gaze-

Table 2 *Total number of subjects in each looking group based on total looks.*

Condition	Diff Score > 0	Diff Score = 0	Diff Score < 0	p-value
person	10	3	2	0.05
+C + F	10	2	1	0.05
+C − F	14	2	4	0.05
−C + F	8	5	0	0.01
−C − F	6	3	10	ns

following behaviors in 12-month-olds. A marginally significant interaction between the variables, $F(1, 61) = 3.30$, $p < 0.08$, reflects the fact that the effects of the two variables were not additive; the presence of a face and contingent behavior together did not elicit more gaze-following than either individually.

A final analysis directly addressed whether the eliciting effect of these two features is facilitated further when embodied in a person rather than a novel object. An unpaired *t*-test revealed no reliable difference between the mean difference scores of the person condition and those of the +C+F (contingent object with face) condition, $t(26) = 0.22$. Apparently the additional cues provided by the familiar category of person gave the infants no discernable advantage in gaze-following.

Discussion

In four of the five conditions tested infants appear to have guided their own looking behavior on the basis of another actor/object's changes in orientation. In three of those four conditions, that object was not a person, nor even an animal. Not only are people not the only objects whose 'gaze' infants will follow, 12-month-olds' tendency to follow the gaze of a person was no greater than their tendency to follow the 'gaze' of a non-human object who possessed certain comparable features. Comparability seems to be predicted by independent analyses of the properties characteristic of intentional beings; in this case these were the presence of facial features and contingent interactivity.

The current findings do not resolve the debate over the nature of the representations supporting gaze-following in 12-month-olds. They do however lend support to the claim that, by 12 months, it is an entity's abstract quality of intentionality that drives infants to follow its 'gaze'. Importantly, the one object configuration in which 12-month-olds failed to be influenced (−C−F) nonetheless shared overall shape, texture, and movement with the successful configurations.[3] This finding is inconsistent with the predictions of the signal releasing theory, at least versions that hold that 12-month-olds' gaze-following is driven by a shape- or motion-based directional cueing mechanism.

Neither can a learning theory based on conditioning explain these results without characterizing the dimensions of similarity leading to this particular pattern of generalization from human headturns. Neither shape nor movement will suffice for the same reasons noted above. Nor will the perceptual category of 'things with faces', given the condition in which infants followed the gaze of a contingent, yet faceless object. One possibility might invoke the abstract category of 'contingent' entity and leave it at that. While this is a possible description of the data presented here, the interpretation based on attribution of intentional states offers an explanation for why contingency might be a useful basis of generalization.

Recent work using the same novel object with adults provides additional evidence that contingency and facial features serve as cues to intentionality (Johnson, 1997). Adults observed an experimenter interacting with the object under the same conditions as the present study, and were then asked to describe and explain what they saw. In just those conditions in which 12-month-old infants followed the object's 'gaze' (+C+F, −C+F, +C−F),[4] adults explained the object's turning behavior in mentalistic terms, *e.g.*, 'It turned because it was looking around the room,' 'Maybe it wanted me to talk to it.' In contrast, the condition that elicited no selective gaze-following in 12-month-olds (−C−F) also failed to elicit any mentalistic descriptors from adults, the vast majority of whom chose instead to describe it as a brown thing that made noise and turned.

Despite the correspondence between 'intentionality-markers' and the successful features in this study, it should be noted that many other 'intentionality-markers' have been proposed which may not prove as effective at eliciting gaze-following. Indeed at least three were not effective in the current studies. Assymetric form as well as self-generated behavior and movement (Baron-Cohen, 1995; Premack, 1990) were all represented, but apparently insufficient, in the failed noncontingent, faceless condition. Nonetheless, these results suggest that an early concept of intentional being may play as central a role in conceptual development as the more commonly considered category of person.

Acknowledgements

This work was supported by NSF grant SBR-9709527 to Susan Johnson and an internal MIT grant to Susan Carey. The authors thank Kimberly Collins, Gergely

[3] Vocalization rates during familiarization and the overall tendency of infants in this condition to look away from the object during trials were in the middle of the range for all the groups, suggesting no difference in overall rates of attention for this group.

[4] Johnson (1997) used a confederate to model the contingent interaction rather than depend on adults to spontaneously talk to the object. Whether infants would treat first- and second-hand information about contingency equivalently is unknown.

Csibra, Uta Frith, György Gergely, Kirsten O'Hearn, Javier Movellan, Jeannie Tyan, John Watson, and two anonymous reviewers.

References

Baldwin, D.A. (1995). Understanding the link between joint attention and language. Development of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origin and role in development*, Hillsdale, NJ: Erlbaum.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development, & simulation of everyday mindreading*. Oxford, England: Blackwell.

Caron, A., Krakowski, O., Liu, A., & Brooks, R. (1996). Infant joint attention: Cued orienting or implicit theory of mind? Presented at the 10th International Conference on Infant Studies, Providence, RI.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origin and role in development*, Hillsdale, NJ: Erlbaum.

Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, **56**, 165–193.

Hood, B.M., Willen, J.D. & Driver, J. (In press). Adults' eyes trigger shifts of visual attention in human infants. *Psychological Science*.

Johnson, S.C. (1997). The attribution of intentionality to novel objects. *Proceedings of the 19th Annual Conference of the Cognitive Science Meeting*, Lawrence Erlbaum: Mahwah, NJ.

Leslie, A.M. (1995). A theory of Agency. In D. Sperber, D. Premack, & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon Press.

Molares, M., Mundy, P., & Rojas, J. (1997). Gaze following and language development in 6- and 8-month-olds. Presented at the Biannual meeting of the Society for Research in Child Development, Washington, DC.

Moore, C., & Corkum, V. (1995). Social understanding at the end of the first year of life. *Developmental Review*, **14**, 349–372.

Movellan, J.R., & Watson, J.S. (1986). Perception of directional attention. Presented at the International Conference on Infant Studies.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Povenelli, D.J., & Eddy, T.J. (1996). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development*, **61**, 3.

Premack, J. (1990). The infant's theory of self-propelled objects. *Cognition*, **36**, 1–16.

Scaife, J.F., & Bruner, J.S. (1975). The capacity for joint visual attention in the infant. *Nature*, **253**, 265–266.

Received: 1 October 1997
Accepted: 4 February 1998