Developmental Psychology
2011, Vol. 47, No. 6, 1499–1503

0012-1649/11/$12.00 DOI: 10.1037/a0025659

BRIEF REPORT

Automated Gaze-Contingent Objects Elicit Orientation Following in 8-Month-Old Infants

Fani Deligianni
Birkbeck, University of London and Imperial College London

Atsushi Senju
Birkbeck, University of London

György Gergely
Central European University

Gergely Csibra
Birkbeck, University of London and Central European University

The current study tested whether the purely amodal cue of contingency elicits orientation following behavior in 8-month-old infants. We presented 8-month-old infants with automated objects without human features that did or did not react contingently to the infants' fixations recorded by an eye tracker. We found that an object's occasional orientation toward peripheral targets was reciprocated by a congruent visual orientation following response by infants only when it had displayed gaze-contingent interactivity. Our finding demonstrates that infants' gaze-following behavior does not depend on the presence of a human being. The results are consistent with the idea that, in 8-month-old infants, the detection of contingent reactivity, like other communicative signals, can itself elicit the illusion of being addressed.

Keywords: infants, eye tracking, contingency, gaze following, communicative intentions

Gaze following is an extensively studied phenomenon of human infant behavior. In naturalistic observations, infants spontaneously shift their gaze to the same object that their interactive partner is looking at from about 10 months of age (Carpenter, Nagell, & Tomasello, 1998). Such a behavior has been suggested to indicate that infants (a) identify the other as an intentional agent (e.g., Johnson, 2003), (b) understand visual perception (e.g., Meltzoff & Brooks, 2008), (c) wish to share attention with the adult (e.g., Carpenter et al., 1998), (d) have learned that gaze shifts predict interesting events (Moore & Corkum, 1994), or (e) interpret gaze as a communicative–referential signal (e.g., Csibra & Gergely, 2009). These explanations of gaze-following behavior are not necessarily mutually exclusive. However, their perspective on its functional significance in early social cognitive development differs considerably.

To understand this phenomenon better, one has to study the early development of gaze following and the exact conditions that elicit such behavior. Gaze following has been shown to occur in 3- to 6-month-old infants in real-life-like situations (D'Entremont, Hains, & Muir, 1997), can be demonstrated in the laboratory at 7 and 8 months of age (Csibra & Volein, 2008; Woodward, 2003), by eye-tracking techniques at 6 months (Gredebäck, Theuring, Hauf, & Kenward, 2008; Senju & Csibra, 2008), and is even detectable in newborns' response tendencies (Farroni, Massaccesi, Pividori, Simion, & Johnson, 2004). It has been shown that gaze following is facilitated by verbal and manual cues that accompany gaze shifts (Flom, Deák, Phill, & Pick, 2004), but there is no evidence that it would be influenced by emotional signals (Flom & Pick, 2005). Furthermore, young infants do not seem to follow gaze in the absence of preceding communicative cues, like eye contact or infant-directed speech (Senju & Csibra, 2008), which are normally employed in infants' interactions with a live partner.

The functional significance of this behavior may be best exposed by situations in which gaze following occurs in the absence of human "gaze," that is, without the presence of a human face or its realistic or schematic representation in dynamic or static form. In particular, temporal contingency, which is not tied to any sensory modality and hence could be considered as an amodal feature of the stimulus environment, has been shown to facilitate gaze following. Ten- and 12-month-old infants tend to follow the orientation of a robot or a nonhuman faceless puppet after it had

This article was published Online First September 26, 2011.

Fani Deligianni, Centre for Brain and Cognitive Development, Birkbeck, University of London, and Department of Computing, Imperial College London; Atsushi Senju, Centre for Brain and Cognitive Development, Birkbeck, University of London; György Gergely, Central European University, Budapest, Hungary; Gergely Csibra, Centre for Brain and Cognitive Development, Birkbeck, University of London, and Central European University, Budapest, Hungary.

We thank John S. Watson for helpful comments. This work is supported by the James S. McDonnell Foundation, the EU6 Framework Programme (NEUROCOM Grant 12738), and an ERC Advanced Investigator Grant (OSTREFCOM).

Correspondence concerning this article should be addressed to Atsushi Senju, Centre for Brain and Cognitive Development, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom. E-mail: a.senju@bbk.ac.uk

responded contingently to the infants' behaviors (Johnson, Slaughter, & Carey, 1998; Movellan & Watson, 1987). First-person (response–stimulus) contingent reactivity of an object can be conceived of as a cue of perceptual and attentional abilities (Johnson, 2003; Movellan & Watson, 2002), or as a cue of communicative intention (Csibra, 2010), or both. Either way, that such an abstract, amodal signal can act as a triggering condition of orientation following suggests that this response may play a functional role in the development of skills for social interactions that characteristically exhibit such temporal adjustments.

The present study aimed to establish whether infants even as young as 8 months of age are sensitive to the purely amodal cue of contingent reactivity that their responses induce in a stimulus object with no human features. Our study differs from earlier studies in a number of significant respects. We employed nonhuman objects (animated on a computer screen) with distinct motion and sound (see Figure 1a). These objects acted as either contingent or noncontingent agents, but they did not possess any humanlike features, and their behavior was fully automated and could not be contaminated by experimenter-induced human biases. The automated eye-tracker methodology we employed allowed us to exert better control over the pattern and degree of contingent reactivity than what was possible in earlier studies with 10- and 12-month-olds (Johnson et al., 1998; Movellan & Watson, 1987). In these previous studies the pattern and degree of response–stimulus contingencies induced, while clearly high but imperfect overall, were nevertheless variable in degree over time, involved variable combinations of different responses by the infant (such as vocalizations and sudden movements detected by the experimenter) and variations in the types and combinations of stimulus effects that these responses induced in the stimulus object (beeping, flashing lights, or turning away). In contrast, by employing a high-precision interactive eye-tracking technique with a gaze-contingent display, we could use young infants' spontaneous gaze-fixations of the stimulus object as the single type of response that induced the pattern of contingent reactivity in the object. The induced stimulus effect in the target object on the computer screen always involved a distinctive animated motion pattern and a sound response, both with fixed durations. In the contingent condition, each time the infant fixated the agent in the middle of the screen, the latter produced a specific contingent response. Following a number of such contingent events, the agent produced a novel response: It changed its orientation to simulate turning toward another noncontingent target object positioned either at the lower left or right side of the screen (see Figure 1b). For the control condition, the playback of this animation was shown to the next infant in a yoked design.

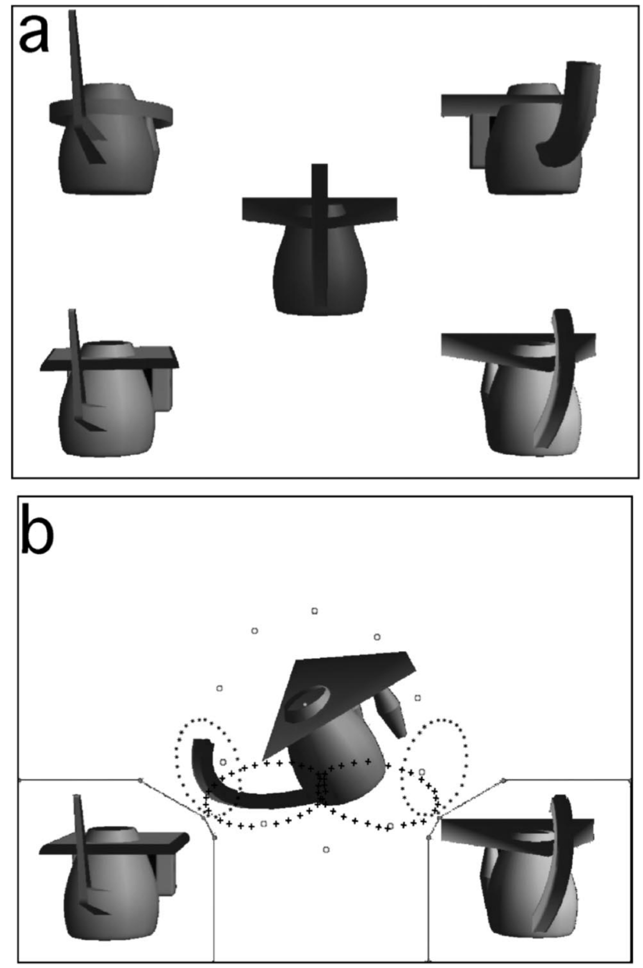

Figure 1. Stimulus displays in the experiment. (a) During the familiarization phase all five objects were present. The objects in the four corners were animated unpredictably. The central object responded to the infant's gaze contingently in the interactive/contingent (IC) condition but was independent of the infant's gazing behavior in the noninteractive/playback (NP) condition. (b) Stimulus configuration during the last 5 s of a trial during the test phase. The region of interest (ROI) of the central object is shown as an ellipse of empty circles. In the IC condition, each time the infant fixated inside this ROI the central object animated. The additional ellipses made of asterisks and crosses cover the extra space occupied by the central object when it turned to one side. The solid line around the corner objects defines their ROIs.

Method

Participants

Thirty-six 8-month-old infants (16 girls, 20 boys) completed the study. Their mean age was 244.2 days (range: 228–261 days). Eighteen infants (7 girls, 11 boys, mean age: 242.1 days) were assigned for the interactive/contingent (IC) condition, and another 18 infants (9 girls, 9 boys, mean age: 246.3 days) were assigned for the noninteractive/playback (NP) condition. A further 8 infants were excluded from the analyses because of inattentiveness and technical problems. All infants were born full term and were recruited by advertisements in local magazines for parents.

Apparatus

A Tobii 1750 Eye Tracker (Stockholm, Sweden) was used to collect gaze data. Stimulus was displayed by MATLAB, using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). A specially developed interface between MATLAB and the Tobii 1750's TET server (Talk2Tobii) allowed online collection and recording of the gaze direction during stimulus presentation.

Stimuli and Procedure

The stimulus presentation started with five 3-D objects, one at the center and four at each corner of a screen of black background (Figure 1a). The 3-D objects were adapted from the Fribble stimulus set (Tarr, 2003). Each of these objects was associated with a distinctive animation scheme and a sound, with duration of 500 ms and 400 to 500 ms, respectively. The animation of each object was implemented on the basis of linear morphing between two predefined shape instances. Each of the corner objects was animated independently from each other at a rate of once in every 10 s with uniform probability distribution. A distinctive color was also associated with each of the four corner objects. In order to get more attention to the central object, its color changed continuously, smoothly and cyclically from green to red to blue throughout the experiment.

The experiment consisted of a familiarization phase and a test phase. In the IC condition, the central object behaved contingently on the infant's looking pattern. When the infant's gaze landed on the central region of interest (ROI) for at least 150 ms (an ellipse around the central object, see Figure 1b), it waited for another interval randomly chosen between 0 and 150 ms and became animated if the infant was still looking at it at the end of this interval. If the infant looked at the central object without interruption for 2 s, it generated an additional response. After four responses, the two corner objects at the upper part of the screen disappeared, the familiarization phase ended, and the test phase started immediately.

The test phase was divided into trials. Each trial included a training phase and a response phase. The training phase was identical to the familiarization phase with the exception that only three objects were present. The training phase ended when the infant's behavior elicited three contingent animations of the central object. Thus, the actual duration of a trial varied according to infants' looking behavior. The response phase started 200 ms later, when the central object "turned" either toward the left or to the right bottom object (Figure 1b). The duration of the turning was 1 s, and the object remained in its final position for an additional 5 s. The other two objects remained still during the response phase. In each trial, the shapes of the two corner objects were selected randomly from the four initial object shapes. The direction of the turning of the central object varied quasirandomly between left and right, ensuring balanced occurrence of both sides. The exact parameters of each animation frame were recorded and saved along with an accurate presentation time stamp. At the end of the trial, all objects disappeared from the screen, and colorful swirling geometric shapes were presented to attract infants' attention to the screen. The experimenter manually controlled the start of each trial.

In the IC condition, the experiment lasted as long as the infants were attentive. In the NP condition, each infant was yoked with an infant in the IC condition and watched the exact same animation sequence that the yoked IC partner produced. In the NP condition, the experiment was terminated either when the playback was over or when the baby became inattentive.

Data Analysis

To analyze the eye-tracking data, saccades and fixations were calculated on the basis of a velocity-based saccade extraction algorithm (Duchowski, 2007). Regions of interest (ROIs) were defined as shown in Figure 1b. Three measures of preferential looking were extracted automatically from the eye-tracking data on the basis of the predefined ROIs during the last 5 s of each trial. The first measurement was based on the direction of the first saccade from the central ROI, which included the additional cover of the "nose" of the turning object (see Figure 1b), to one of the side objects' ROI, showing either congruent or incongruent response with the central object's turning direction. The normalized difference between congruent and incongruent saccades was calculated for each infant by subtracting the number of trials with incongruent saccades from the number of trials with congruent saccades and dividing this value by the total number of trials that generated a central object-to-corner object saccade (Corkum & Moore, 1998; Moore & Corkum, 1998). As a second measure, the frequency of central-to-corner object saccades was estimated on the basis of the difference between the number of saccades toward the congruent object and those toward the incongruent object. Finally, the total duration of the fixation on each corner object was calculated. The frequency and duration measures were normalized the same way as the first measure. All reported statistics are two-tailed, with $\alpha = .05$.

Results

In the IC condition, an average of 17.2 trials (min: 12, max: 31) were completed. In the NP condition, 15 infants completed all trials, while 3 infants watched 10 out of 21, 14 out of 22 and 15 out of 31 trials, respectively. On average, the number of valid trials was 16.2 (min 9, max 31) in the IC condition and 14.1 (min 7, max 22) in the NP condition. A trial was defined as valid if the infant looked at the central object at least once during its turning.

We measured the proportion of looking times toward the central object at various phases of the events to assess the amount of attention infants paid to this object. During the whole experiment, infants fixated at the central object for 52.34% and 50.82% of the time in the IC and NP condition, respectively. This difference is not statistically significant, $t(17) = 1.24$, $p = .231$. (When comparing the two groups, we used paired t tests, treating the infant in the IC condition and his or her NP-yoked partner as a single statistical unit.) During the animation of the central object, infants fixated this object more in the IC (75.57% of time) than in the NP (58.53% of time) condition, $t(17) = 3.59$, $p = .002$. This is not surprising, as animation started only when infants had already fixated this object in the IC condition, while in the NP condition infants most of the times had to make a saccade to the animated object. However, we found no statistically significant difference, $t(17) = 0.17$, $p = .863$, between the two groups in their fixation time at the central object during the turning phase (IC: 54.76%, NP: 55.49%). Thus, all infants had equal chance to observe and follow the orientation of the central object.

In the IC condition, preferential looking toward the contingent object was significantly above chance level for all three measures. Infants were significantly more likely to direct their first saccades toward the target object, $t(17) = 4.07$, $p < .001$, $d = 1.11$, look more often toward, $t(17) = 3.86$, $p = .001$, $d = 0.91$, and longer, $t(17) = 4.06$, $p < .001$, $d = 0.96$, at it (see Figure 2). However, none of these measures differed from baseline in the NP condition: direction of first saccade, $t(17) = 0.68$, $p = .506$, $d = 0.16$;

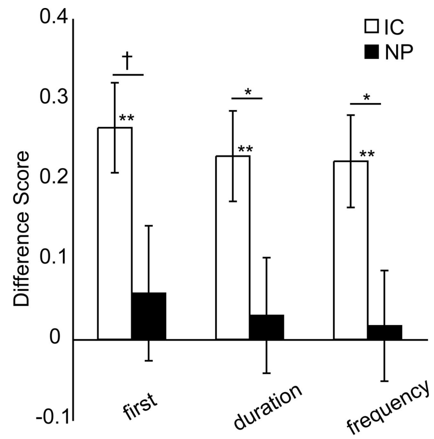

Figure 2. Measures of preferential looking. Average normalized difference scores are depicted for direction of the first central-to-corner-object saccade (First), duration of object fixation (Duration), and frequency of object directed saccades (Frequency) in each condition. IC = interactive contingent condition; NP = noninteractive/playback condition. † $p < .075$. * $p < .05$. ** $p < .01$. Error bars represent standard error.

frequency of fixations, $t(17) = 0.24$, $p = .813$, $d = 0.06$; duration of fixations, $t(17) = 0.41$, $p = .687$, $d = 0.10$.

We also compared the measures of orientation following between the conditions. These tests demonstrated a significant difference in the frequency of saccades toward, $t(17) = 2.69$, $p = .016$, $d = 0.63$, and in the duration of fixations on, $t(17) = 2.51$, $p = .023$, $d = 0.59$, the target object (see Figure 2). For the direction of the first saccade, the test indicated a marginally significant difference, $t(17) = 1.91$, $p = .073$, $d = 0.45$. We also calculated nonparametric Fischer's exact tests to compare the number of infants who produced positive versus negative difference scores on the three measures in the two conditions (here we ignored infants whose difference score was 0). In the IC condition, 15 infants produced more first saccades toward the target than to the other object, while only 1 infant showed the opposite pattern. In the NP condition the corresponding numbers are 10 and 7, respectively, and the difference between groups was statistically significant ($p = .039$). Similar calculations resulted in significant group difference for the duration of fixations (IC: 15 vs. 3, NP: 8 vs. 10; $p = .035$), but not for saccade frequencies (IC: 13 vs. 4, NP: 7 vs. 8; $p = .127$). These results generally confirm those of the parametric tests while suggesting substantial individual variability in various parameters of eye-movement patterns.

Discussion

We found that 8-month-old human infants followed the turning of an object toward a target stimulus, but they did so only if the object had repeatedly responded to them in a contingent fashion. Furthermore, the response-induced contingent reactivity of the object was itself a sufficient cue to induce the infants' orientation-following response as the mechanically animated agent on the screen had no human features and was operated automatically without human intervention. In contrast, observing an accurate playback of what another infant had seen earlier in the contingent condition did not elicit such orientation-following reactions in the noncontingent condition. Thus, the current results extend previous findings (Johnson et al., 1998; Movellan & Watson, 1987) by establishing that infants as young as 8 months of age are sensitive to the purely amodal cue of contingent reactivity that is sufficient in and of itself to elicit the orientation-following gaze response from infants.

The current results suggest that infants' ability to detect contingency and consequent tendency to follow the direction of the responsive agent's subsequent target-oriented behavior are not dependent on the presence of any other social cues, such as faces or human voice, or on the communicative nature of the behavior (such as vocalization) that themselves could trigger the infant's orientation following response (see Senju & Csibra, 2008, for a similar orientation-following effect as a function of preceding direct eye contact or infant-directed speech in 6-month-olds). In the present study the contingent reactivity of a 2-D object without human features on a computer screen programmed to repeatedly react to being visually fixated with the production of an identical motion and acoustic response proved in itself sufficient to elicit subsequent "gaze-following" behavior in the infants. This finding suggests that the 8-month-olds may have interpreted such an object as an agent with communicative and referential intention (Csibra & Gergely, 2006, 2009) toward them and/or attributed attentional and perceptual abilities (Johnson, 2003) and a "line of regard" (Movellan & Watson, 2002) to it. In order to succeed, infants did not only have to pay attention to what the central object did but also had to be able to detect its direction of turning and had to interpret the protuberance on the object as a potential device of perception (like an eye) or reference (like a pointing hand). For this purpose, "Fribbles" offered an optimal stimulus class, as they have an unambiguously marked direction without exhibiting featural similarity to humans.

These results are consistent with the idea that the abstract temporal structure of turn-taking contingency (i.e., alternating and temporally self-adjusted pattern of activity of a stimulus source) functions as an amodal cue of ostensive communication that is not tied to any particular modality-specific instantiation and could generate referential expectation in infants in the same way as do other ostensive signals (such as eye contact or infant-directed speech) that function to communicatively address the infant (e.g., Csibra, 2010). Because young infants follow human gaze shifts only when they are preceded by an ostensive signal (Senju & Csibra, 2008), the current results support the idea that turn-taking contingent reactivity is perceived as functionally equivalent with other infant-directed communicative signals and may generate an illusion of perceived communicative intention attributed to the nonhuman gaze-contingent object.

What is the benefit of relying on such an abstract and amodal structural cue as an ostensive signal? Different communicative signals usually occur together, especially when they are directed to infants. Such multimodal stimuli support each other, and turn-taking contingency can disambiguate whether a signal in a sensory modality (e.g., speech) is directed to the infant. Indeed, contingency can serve as a training signal that helps the infant figure out, and eventually learn, whether an unusual and unfamiliar stimulus configuration is in fact an ostensive signal. For example, while newborns can detect direct gaze in a full frontal face (Farroni, Csibra, Simion, & Johnson, 2002), they are unable to do so when a face is turned 3/4 to the side (Farroni, Menon, & Johnson, 2006).

Evidence shows that by 4 months of age, they can detect whether the eyes in a face that is oriented away are looking at them (Farroni, Johnson, & Csibra, 2004), and the neural responses to this configuration suggest that they treat it the same way as other ostensive stimuli, like smiles (Grossmann et al., 2008). Turn-taking contingency, which normally accompanies eye contact and infant-directed communication, may contribute to such fine-tuning of the recognition of modality-specific ostensive signals.

References

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436. doi:10.1163/156856897X00357

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*. doi:10.2307/1166214

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology, 34,* 28–38. doi:10.1037/0012-1649.34.1.28

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25,* 141–168. doi:10.1111/j.1468-0017.2009.01384.x

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of change in brain and cognitive development: Attention and performance XXI* (pp. 249–274). Oxford, England: Oxford University Press.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148–153. doi:10.1016/j.tics.2009.01.005

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. doi:10.1348/026151007X185987

D'Entremont, B., Hains, S. M. J., & Muir, D. W. (1997). A demonstration of gaze following in 3-to 6- months-olds. *Infant Behavior & Development, 20,* 569–572. doi:10.1016/S0163-6383(97)90048-5

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice.* London, England: Springer Science.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *PANAS: Proceedings of the National Academy of Sciences of the United States, 99,* 9602–9605. doi:10.1073/pnas.152159999

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience, 16,* 1320–1326. doi:10.1162/0898929042304787

Farroni, T., Massaccesi, S., Pividori, D., Simion, F., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5,* 39–60. doi:10.1207/s15327078in0501_2

Farroni, T., Menon, E., & Johnson, M. H. (2006). Factors influencing newborns' preference for faces with eye contact. *Journal of Experimental Child Psychology, 95,* 298–308. doi:10.1016/j.jecp.2006.08.001

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 & Development, 27,* 181–194. doi:10.1016/j.infbeh.2003.09.007

Flom, R., & Pick, A. D. (2005). Experimenter affective expression and gaze following in 7-month-olds. *Infancy, 7,* 207–218. doi:10.1207/s15327078in0702_5

Gredebäck, G., Theuring, C., Hauf, P., & Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy, 13,* 533–543. doi:10.1080/15250000802329529

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., & Elwell, C. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings of the Royal Society B: Biological Sciences, 275,* 2803–2811. doi:10.1098/rspb.2008.0986

Johnson, S. (2003). Detecting agents. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358,* 549–559. doi:10.1098/rstb.2002.1237

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1,* 233–238. doi:10.1111/1467-7687.00036

Meltzoff, A. N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology, 44,* 1257–1265. doi:10.1037/a0012888

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

Moore, C., & Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology, 16,* 495–503. doi:10.1111/j.2044-835X.1998.tb00767.x

Movellan, J. R., & Watson, J. S. (1987). *Perception of directional attention.* Paper presented at the Sixth International Conference of Infant Studies.

Movellan, J. R., & Watson, J. S. (2002). *The development of gaze following as a Bayesian systems identification problem* (Report No. 2002.01). San Diego, CA: University of California at San Diego Machine Perception Laboratory.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. doi:10.1163/156856897X00366

Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology, 18,* 668–671. doi:10.1016/j.cub.2008.03.059

Tarr, M. J. (2003). Perception of faces, objects, and scenes: Analytic and holistic processes. In M. A. Peterson & G. Rhodes (Eds.), *Visual object recognition: Can a single mechanism suffice?* (pp. 177–211). Oxford, England: Oxford University Press.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science, 6,* 297–311. doi:10.1111/1467-7687.00286

Received June 21, 2010
Revision received July 18, 2011
Accepted July 28, 2011 ■