Cognitive Development 60 (2021) 101121

Contents lists available at ScienceDirect

Cognitive Development

journal homepage: www.elsevier.com/locate/cogdev

Gaze following emergence relies on both perceptual cues and social awareness

Kim Astor[a,*], Maleen Thiele[b], Gustaf Gredebäck[a]

[a] Uppsala University, Department of psychology, Uppsala Child & Baby Lab, Sweden
[b] Max Planck Institute for Evolutionary Anthropology, Department of Comparative Cultural Psychology, Germany

ABSTRACT

Decades of research have emphasized the significance of gaze following in early development. Yet, the developmental origin of this ability has remained poorly understood. We tested the claims made by two prominent theoretical perspectives to answer whether infants gaze following response is based on perceptual (motion of the head) or social cues (gaze direction). We found that 12-month-olds ($N = 30$) are able to inhibit motion cues and exclusively follow the direction of others' gaze. Six- ($N = 29$) and 4-month-olds ($N = 30$) can follow gaze, with a sensitivity to both perceptual and social cues. These results align with the perceptual narrowing hypothesis of gaze following emergence, suggesting that social and perceptual cueing are non-exclusive paths to early developing gaze following.

1. Introduction

Gaze following, the ability to align one's own gaze with others to focus on external objects, is observed in many social species; ranging from primates to corvids (Catala, Mang, Wallis, & Huber, 2017; Téglás, Gergely, Kupán, Miklósi, & Topál, 2012; Tomasello, Hare, Lehmann, & Call, 2007). As the eyes carry information about both direction of attention and emotional states, the ability to interpret others' gaze allows for quick social assessment and efficient transfer of information (Dezecache, Crockford, & Zuberbühler, 2019; Emery, 2000). In humans, gaze following starts to emerge between 4 and 6 months of age (Astor & Gredebäck, 2019; D'Entremont, 2000; Gredebäck, Theuring, Hauf, & Kenward, 2008). It represents a crucial condition for learning language and other socially transmitted skills (Barry-Anwar, Burris, Graf Estes, & Rivera, 2017; Brooks, & Meltzoff, 2015; Morales et al., 2000, Tenenbaum, Sobel, Sheinkopf, Malle, & Morgan, 2015). Lower levels of gaze following have been related to difficulties in emotion regulation (Morales, Mundy, Crowson, Neal, & Delgado, 2005) and represent a defining feature of autism (APA, 2013; Mundy, 2016). However, despite decades of extensive research, the developmental origin of gaze following remains poorly understood (for a recent review see: Del Bianco, Falck-Ytter, Thorup, & Gredebäck, 2018).

In the mid-1970s, pioneering gaze following researchers Scaife and Bruner (1975) suggested: "In so far as mutual orientation implies a degree of knowledge in some form about another person's perspective, then the child in its first year may be considered as less then completely egocentric" (p.266). The idea that early gaze following indicates a certain degree of social awareness remains a matter of debate. We will refer to theories in line with Scaife and Bruner's notion as the *social-first* perspective. According to this perspective, the early onset of infants' gaze following is thought to be guided by the motivation to engage with others (Gredebäck, Astor, & Fawcett, 2018; Gredebäck, Fikke, & Melinder, 2010); by their ability to see others as similar to themselves (Brooks & Meltzoff, 2014; Meltzoff, 2007); by a specific sensitivity to communicative signals (Csibra & Gergely, 2009); or by innate and specialized processing modules detecting others' gaze direction and shared attention (Baron-Cohen, 1995). While there are critical differences between the social-first

* Correspondence to: Box 1225, 751 42 Uppsala, Sweden.
E-mail address: kim.astor@psyk.uu.se (K. Astor).

https://doi.org/10.1016/j.cogdev.2021.101121
Received 16 March 2021; Received in revised form 10 September 2021; Accepted 15 September 2021
Available online 1 October 2021
0885-2014/

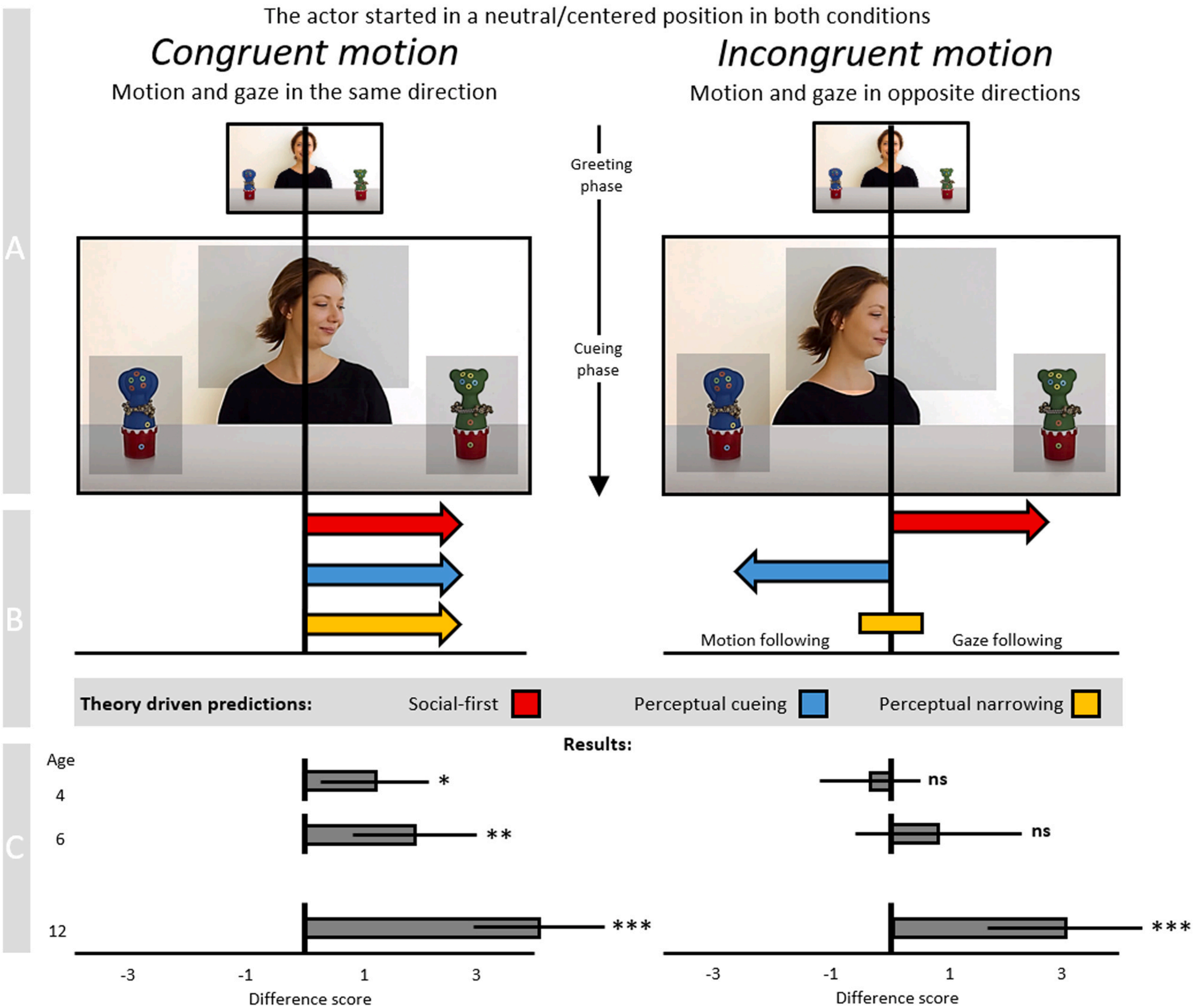

Fig. 1. Illustration of the trial sequence, hypotheses, and results. (A) In both the congruent and incongruent motion conditions, the trial started with a "greeting phase" with the actor positioned at the center of the screen (topmost screenshots), at the vertical black line on the horizontal plane. The next pair of screenshots illustrate the actor's position in the "cueing phase". In the congruent motion condition, the actor rotated her head and gazed towards one of the toys, which also created a lateral motion cue towards this toy. In the incongruent motion condition, the head rotation was countered by a lateral motion (head and body) in the opposite direction. This resulted in a rotation (gaze orientation) towards one toy and a lateral motion towards the other. In this way, the eyes and head drift laterally differentially between conditions though the gaze points in the same direction. The gray transparent rectangles represent the AOI's, and the black vertical lines indicate the center of the video. Centerline and AOI's serve illustrative purposes and were not visible to the infant. (B) Predictions (red, blue, and yellow arrows) derived from different theoretical perspectives (not the results). In the exemplary screenshot of the incongruent motion condition, the left toy was cued by motion and the right toy was cued by gaze. (C) Results of one-sample *t*-tests for the three age groups, represented by bar graphs with error bars that correspond to 95% CI. ***$p < .001$, **$p < .01$, * $p < .05$.

theories, this study focuses on their shared assumption; that early gaze following is a social response.

In contrast to the social-first perspective, many aspects of early developing gaze following suggest that it is better understood as a domain-general perceptual-motor response that leads to behaviors reminiscent of gaze following, without initially being socially motivated (Deák, 2015). This assumption is referred to as the *perceptual cueing* perspective. For example, newborn infants (Farroni, Massaccesi, Pividori, & Johnson, 2004) and 9-month-olds (Moore, Angelopoulos, & Bennett, 1997) attend to objects cued by the motion of a face rather than the direction of gaze. Moreover, 'gaze following like' behaviors have been observed in 8-month-olds in response to the motion of non-animate 3D objects without any human features (Deligianni, Senju, Gergely, & Csibra, 2011). These findings strengthen the notion that gaze following in infants may not necessarily involve a social sensitivity, but instead represent a form of motion following (Gredebäck & Daum, 2015). In other words, the perceptual cuing perspective suggests that infants' gaze following is the result of following the direction of motion (Deák, 2015), and that this response is not specific to human facial features, but applicable to other stimuli as well, such as hands (Yu & Smith, 2013) or tea-pots (Deligianni et al., 2011). In addition, perceptual cueing can be seen as a foundation for a reinforcement learning perspective on the emergence of gaze following: From this view, infants' existing tendency to orient in the same direction as others is continuously reinforced, as other's motion/gaze cues tend to point towards interesting sights (Corkum & Moore, 1998; Deák, 2015; Triesch, Teuscher, Deák, & Carlson, 2006).

With respect to the cues, there is a lack in experimental studies explicitly testing the social-first and perceptual cueing perspectives against each other (Del Bianco et al., 2018). The majority of gaze following studies has relied on the same standard paradigm in which an actor faces the infant, before moving her gaze (eyes and head) toward one of two objects positioned at her left and right side. Crucially, in this standard scenario, gaze direction (orientation of the eyes and head) is confounded with a motion cue from the actor's head, making it difficult to disentangle whether infants follow motion, social information, or both. Some prior work has stated to unpack the low-level cues present in this paradigm, demonstrating that proximity of the face and eyes to the attended object is not sufficient to elicit 'gaze following like' responses (Astor & Gredebäck, 2019). However, from a theoretical perspective, separating the role of motion cuing from gaze following has the potential to unlock the processes by which gaze following emerges and continues to develop over the first year of life.

1.1. This study

Our aim was to set up a critical test, forcing the social-first and perceptual cueing perspectives to predict different outcomes when tested simultaneously. As noted above, the two umbrella terms social-first and perceptual cuing perspective incorporate several different theories and models of gaze following. Importantly, the current study did not intend to test and differentiate the detailed assumptions of all included accounts, but rather test the common assumption behind the two theories, namely that gaze following originates in a social awareness or emerges from motion cuing. We assessed 4-, 6-, and 12-month-old infants' performance in two conditions: the *congruent motion* condition (i.e., the standard gaze following test paradigm), and a new modified version of the test, the *incongruent motion* condition that aimed to test the two perspectives directly against each other. In the latter, the directional orientation of the actor was separated from the lateral displacement (effectively making the actors' overall body, including the head, move in one direction while turning face and gaze in the opposite direction, see the incongruent motion condition in Fig. 1). From a *perceptual cueing perspective*, the lateral displacement of the actor should guide infants' attention, making them follow the actor's motion rather than gaze direction. In contrast, the *social-first perspective* predict that infants follow gaze, regardless of competing motion cues. A third possibility exist: Early in life, infants may rely on information from a wide range of cues (including both motion and gaze direction), but over the course of development, they narrow down to a smaller set of more specific social cues. This developmental view is referred to as the *perceptual narrowing perspective* (Del Bianco et al., 2018). The perceptual narrowing perspective predict that, when both cues align (congruent motion condition), infants at all ages (here 4, 6, and 12 months) should follow the actor's gaze direction. However, when the two cues compete with one another, such as with a head and body movement to one side and a gaze shift in the other direction (incongruent motion condition), a lack of both gaze and motion following would be expected at younger ages (when gaze following begins to emerge, here at 4 and 6 months). Though inclusive, demonstrating that both motion and social awareness can motivate gaze following early in life, such a result would strengthen the suggestion that early gaze following is less than completely egocentric (Scaife & Bruner, 1975) and involves an interest in others that goes beyond motion cueing. These larger assumptions about the driving forces behind the emergence of gaze following have never been empirically tested.

2. Method

The hypotheses, methods, procedures, and the data analysis plan were pre-registered on AsPredicted (blinded link: http://aspredicted.org/blind.php?x=w3m2vc).

2.1. Participants

In total, 98 human infants were invited to the lab for testing at 4 months $N = 31$, $M = 127$ days, $SD = 6$, 6 months $N = 37$, $M = 184$ days, $SD = 9$, and 12 months $N = 30$, $M = 366$ days, $SD = 10$. Eight of those infants were either canceled, never showed up for testing, or could not provide any data due to technical problems or fussiness. This resulted in a final sample of 90 infants ($N = 30$ per cohort with 15 girls and 15 boys in each). We based the sample size on previous studies with a similar paradigm and tested age range (e.g. Astor & Gredebäck, 2019) to ensure a baseline effect of gaze following (congruent motion condition). To be included in the sample, infants needed to be within the maximum age range of 2 weeks younger or older than the target age. Though no additional demographic information was obtained, participants in this study generally come from urban, White, middle-class families in which parents have a university degree. Participation was rewarded with a gift voucher worth €10. Before the experiment, parents were informed about the purpose and procedure of the study. Written informed consent was obtained from both parents. The study was approved by the local ethics committee (EPN) in accordance with the Declaration of Helsinki (1964).

2.2. Stimuli and design

Each trial consisted of a video showing an actor sitting at a table with one toy on each side (see Fig. 1). The video started with a greeting phase (2 seconds) in which the actor looked into the camera and said “Hi” to the infant in its mother tongue. Then, in the cueing phase, the actor turned her head towards one of the toys, fixating her gaze on it (5 seconds). Each trial was preceded by an attention-grabbing colorful animation with sound, placed in the center of the screen to reset attention between trials.

We used a mixed design, testing the effect of perceptual cueing within subjects, and the effect of age cross-sectionally. Two versions of the stimuli were created (congruent and incongruent motion, see Fig. 1). In the congruent motion condition, we applied the standard version of the gaze following paradigm (Astor & Gredebäck, 2019; Brooks & Meltzoff, 2015; D’Entremont, 2000; Tenenbaum et al., 2015), as described above. In the incongruent motion condition, designed to separate the competing hypotheses (social-first and perceptual cueing), the actor’s gaze shift (in this study operationalized as head and eye orientation) was countered by a simultaneous lateral motion of her upper body (including the head) in the opposite direction. Special care was taken to make sure that all motion cues (both body and facial features) moved in the opposite direction to the gaze. Note that this is different from the way other studies have separated head and eye-gaze (e.g., Corkum & Moore, 1995; Farroni, Johnson, Brockbank, & Simion, 2000), and from studies investigating the effect of direct gaze on information processing (Farroni, Johnson, & Csibra et al. 2004; Yamashita, Kanazawa, & Yamaguchi, 2012) and joint attention (Gredebäck et al., 2018). This resulted in a gaze cue (head and eye orientation) in one direction and a large lateral motion cue (head and body) in the other. In both conditions, the positioning of the actor in the initial greeting phase was exactly centered between the two toys. When comparing the two versions of the stimuli, the eyes and head drift laterally differentially between conditions though the gaze points in the same direction.

Three female actors (each with a unique, but similar, set of toys) performed each of the two conditions (congruent and incongruent motion) in both possible directions (towards the right and left object), resulting in 4 unique trials per actor. The resulting 12 trials (3 actors x 4 trials) constituted a block of trials. All infants saw two identical blocks, resulting in 24 trials in total. Within each block, the same condition, actor or direction never occurred more than two times in a row. We created two counterbalanced versions of the experiment where one version mirrored the presentation order of the other versions’ blocks. One version started with the congruent motion condition and the other with the incongruent motion condition. Quasi-randomized assignment was applied.

We excluded participants from the analysis who provided less than two valid trials in either of the two conditions. This only applied to one participant, who scored in 5 trials in the congruent motion condition but only 1 trial in the incongruent motion condition (see Table 1 for descriptive statistics of the scoring).

2.3. Apparatus

The stimuli were presented on a 24-inch monitor with integrated Tobii TX300 eye-tracker (Tobii Studio version 3.4.8) recording gaze at a sampling rate of 120 Hz. Speakers were placed behind the monitor. The room was illuminated and white curtains shielded infants from visual distractions.

2.4. Procedure

Infants were seated on their parents lap approximately 60 centimeters from the screen during the whole session. A five-point calibration was performed before the experiment (Gredebäck, Johnson, & von Hofsten, 2010). The total duration of the Experiment lasted approximately 5 minutes, the whole visit at the lab around 20 minutes.

2.5. Data reduction

The raw eye-tracking data was analyzed using the open source time series analysis tool TimeStudio (version 3.18), based on MATLAB (version R2017b), available at www.timestudioproject.org (Nyström, Falck-Ytter, & Gredebäck, 2016). Three rectangular

Table 1
Descriptive statistics for the number of valid trials per condition for each cohort.

Age	Condition	N	Min	Max	Mean	SD
4	Congruent	30	2	12	7.27	3.27
	Incongruent	30	2	12	7.43	2.93
6	Congruent	29	3	12	8.83	2.75
	Incongruent	29	2	12	8.93	2.71
12	Congruent	30	3	12	9.83	2.02
	Incongruent	30	4	12	9.73	2.13

areas of interest (AOI) were defined: one area covering the head (24° visual field), and two equally-sized areas covering the two toys (11° visual field), see Fig. 1. The AOIs were separated with 2° visual field. The head AOI was defined so that infants would perceived both critical cues, both the lateral motion and the rotation (gaze shift) of the head. For each trial, we measured infants' first gaze shift (defined as a fixation using Tobii fixation filter) from the actor towards either of the toys, starting from the time when the actor started turning her head (cuing phase). A child received a score of (1) if they looked at the same toy as the actor, (−1) if they looked at the non-cued toy, or (0) if they did not look at any of the toys. A standard difference score was obtained for each condition (see Brooks & Meltzoff, 2002; Corkum & More, 1998; Deák, Flom, & Pick, 2000; Johnson, Slaughter, & Carey, 1998). The score could range from −12 to 12, with positive values indicating gaze following.

3. Results

3.1. Main analyses

To assess infants' gaze following, we tested the mean difference score against chance level (i.e., zero) by using single-sample *t*-tests. We performed six *t*-tests in total, one for each condition and age group. To control for false discovery rate, all *p*-values were adjusted using Benjamini and Hochberg (1995) procedure.

3.1.1. Congruent motion condition

In the congruent motion condition (standard paradigm), infants across all age groups followed the actors' gaze: 4-month-olds, $M = 1.2$, $SD = 2.43$, 95% CI [.29, 2.10], $t(29) = 2.71$, $p = .017$, $d = 0.494$, 6-month-olds, $M = 1.86$, $SD = 2.75$, 95% CI [.82, 2.91], $t(28) = 3.65$, $p = .002$, $d = 0.678$, and 12-month-olds, $M = 3.97$, $SD = 2.94$, 95% CI [2.87, 5.07], $t(29) = 7.39$, $p < .001$, $d = 1.348$.

3.1.2. Incongruent motion condition

In the incongruent motion condition (competing hypotheses test), infants did not follow any cue at 4 months, $t(29) = -0.88$, $p = .386$, $d = -0.161$, $M = -0.37$, $SD = 2.28$, 95% CI [−1.22,.49], and at 6 months, $t(28) = 1.16$, $p = .309$, $d = 0.215$), $M = 0.79$, $SD = 2.69$, 95% CI [−0.61, 2.20]. However, at 12 months, infants followed the actors' gaze, $t(29) = 4.59$, $p < .001$, $d = 0.837$, $M = 2.93$, $SD = 3.50$, 95% CI [1.63, 4.24]. Frequency tables are available in Supplementary 1. The results are illustrated in Fig. 1.

3.2. Complementary analyses

To further investigate the impact of motion cueing and age on gaze following, we ran a two-way mixed ANOVA (not pre-registered) for infants' gaze following difference score (dependent variable), including age and condition as interaction. The analysis did not reveal a significant effect of the interaction, $F(2, 86) = 0.19$, $p = .831$, partial $\eta^2 = .004$, but significant main effects of both factors. Infants performed more gaze following in the congruent motion compared to the incongruent motion condition, $F(1, 86) = 9.27$, $p = .003$, partial $\eta^2 = .079$, and their gaze following increased with age, $F(2, 86) = 13.82$, $p < .001$, partial $\eta^2 = .243$.

4. Discussion

In this study, we could demonstrate that infants have developed sufficient cognitive and oculomotor skills to align their gaze with others well before the age of 6 months. Infants at 12, 6, and even 4 months of age followed the gaze of an actor in a standard gaze following task, thus adding to our understanding of gaze following emergence where previous research has been inconsistent (Astor & Gredebäck, 2019; Gredebäck et al., 2008). In addition, our study provides direct evidence concerning the critical question: *how* infants start to follow gaze.

A modified version of the standard gaze following task (i.e., the incongruent motion condition) allowed us to test two theoretical perspectives on gaze following emergence directly against each other: the social-first perspective and the perceptual cueing perspective. During this modified task, infants' saw an actor turning her gaze (head and eye) in one direction while simultaneously moving her body (including the head) laterally in the opposite direction. This lateral motion negated the motion of the actor's head turn (typically an integral part of a gaze shift). At the same token, this large counter movement of the head and body resulted in a new motion cue in the direction opposite to the gaze. In this way, we not only neutralized the motion cue that typically confound a gaze shift, but we also separated gaze orientation and motion direction into two competing directional cues. This way, we forced the two perspectives to make different predictions: According to the social-first perspective, infants should follow the gaze cue. According to the perceptual cueing perspective, infants should follow the motion cue (Fig. 1). We found that, at 12 months of age, infants' visual attention was cued by others' gaze direction. Even when the motion cue pointed in the opposite direction, 12-month-olds followed the direction of the gaze shift. At the critical age of gaze following emergence, however, 4- to 6-month-old infants followed neither gaze nor motion when the two cues played against each other.

The result is not consistent with either perspective in isolation. Infants' typical response (to follow gaze orientation and head movement when both align) disappeared when the two cues pointed into opposite directions, resulting in an unsystematic response. In fact, our finding rather suggests that social (gaze) and perceptual (motion) cueing are non-exclusive paths to gaze following and that infants may use both cues to guide their attention.

Against this background, we suggest that the perceptual narrowing hypothesis of gaze following emergence most accurately captures the current results (Del Bianco et al., 2018): When infants start to follow gaze, at around the age of 4–6 months of age, they are sensitive to a wide range of cues, including gaze and motion, and there seems to be a parity between the two cues when conflicting. Hence, infants might use either social or perceptual cues to produce the 'same' overtly indistinguishable gaze following response. As endogenous control of visual attention develops across infancy (Colombo, 2001), infants eventually acquire the ability to inhibit low-level responses more effectively. Indeed, toward the end of the first year, infants in the current study become more likely to act on the gaze cue: the group of 12-month-old infants followed the actors gaze direction even in the presence of the competing motion cue. At this age, our data seems to suggest that gaze following become almost exclusively guided by goal-driven endogenous control (even though a sensitivity to low-level perceptual cues likely remains, e.g., in situations where information is sparse). This is the first study to provide empirical support to the notion that both motion and gaze cues are important at the onset of gaze following. Our results argue against the suggestion that perceptual cueing is the one exclusive pathway to gaze following emergence. Instead, there seems to be a social component to gaze following from the start. Consequently, gaze following emergence cannot be explained by a reinforcement process grounded exclusively in motion cuing between 4 and 12 months of age, that is, at the age of onset (see also Astor et al., 2020).

Our findings offer a number of alternative interpretations. For example, it is possible that young infants weight information differently, or get confused by the competing cues. It is also possible that infants get confused if they perceive the action as novel (though there are situations where people turn to look at things while simultaneously moving back to focus on a proximal object or flinch for a threat). From a perceptual cueing perspective, it is assumed that motion per se triggers a gaze shift (in this context a gaze following response), the specifics of the motion are thought to be irrelevant. In other words, it is not expected that only some motion cues (e.g., head or eyes) elicit gaze following, but that others are less viable or confusing (e.g., body including head). Hence, the null results in the incongruent condition (at 4 and 6 months) cannot be explained by cue-specific induced confusion from a perceptual cueing perspective. Assuming that confusion occurs (which is not reflected in the scoring between conditions, see Table 1), this would either have a social-first explanation; suggesting that the incongruent motion cue is socially ambiguous or unfamiliar, or an explanation that accounts for effects of both social and perceptual cues. As we know from previous work that motion is sufficient for cueing, the latter alternative seems more likely. Based on the current study, it is not possible to conclusively disentangle these alternatives from the perceptual narrowing perspective. Nevertheless, the alternative explanation arrives at the same conclusion as the proposed framework: The perceptual cueing perspective alone cannot explain our results and emerging gaze following is likely to have a social element. Instead, in line with the perceptual narrowing hypothesis, it seems that several different cues (social and non-social) can elicit a gaze following response early in life, and that infants with increasing age tune in to social signals as the one reliable cue worth following.

As noted in the Introduction, this study does not focus on the differences between social-first accounts, but rather on how the overarching theoretical framework agrees to interpret infants' gaze following as a social response. All social-first theories, except the theory by Baron-Cohen's, is conceptually compatible with the idea that infants may view others as 'body orienters' (Brooks & Meltzoff, 2005), a less precise but nevertheless social account of gaze following. It is therefore important to recognize the distinction between orientation direction (where the body is facing) and motion direction (lateral displacement). The current study kept eye-gaze and head orientation direction consistent while separating it from motion direction (incongruent motion condition), making it overall consistent with the predictions from different social-first accounts.

In this study we used a screen-based video presentation of pre-recorded stimuli, a common approach to theoretical assessment in the gaze following literature because of the high level of stimuli-control this method offers. In this study stimuli control was critical in the incongruent motion condition as the counter movement (a head turn in one direction was countered by a lateral head and body motion in the opposite direction) had to be exact in both timing and amount. Though a live actor assessment would offer higher levels of ecological validity, it is a well-known fact that infants process many kinds of social information/cues when presented on a screen (Csibra & Gergely, 2009, Farroni et al., 2004, Wahl, Michel, Pauen, & Hoehl, 2013). Moreover, since gaze following is demonstrated in studies using both live (Gredebäck et al., 2010) and screen-based (Astor et al., 2020), we would not predict a different outcome in a live setting. Perhaps virtual reality (head mounted displays) can offer a way forward where higher levels of ecological validity are achieved while still benefitting from the same level of stimuli control as a traditional screen-based medium. Regardless of stimuli presentation method, researchers always have to ensure that infants can process the spatial aspects of the scene in a way that is appropriate for the specific task. In this study, infants completed a five-point calibration, covering the whole screen area, prior to testing. Expected levels of base line gaze following, across all ages, confirmed infants' ability to process the scene as intended (see Astor & Gredebäck, 2019 for a further discussion on infants' visual field in the context of screen-based gaze following assessment).

The findings from this study are reported at group level and need further investigation using an individual differences approach. Some infants may be more sensitive to motion while others may be more sensitive to gaze (in general or dependent on context). Knowing how these differences relate to associated abilities or conditions; such as language development and autism, may help us better understand the role of gaze following in early development. Future studies may benefit from adding a third condition where the actor moves laterally without a head turn. This would capture the effect of motion cueing without competing cues. In such condition the perceptual cueing perspective would predict a gaze shift to the toy in the direction of the lateral motion. There would be no competing prediction from a social-first perspective.

In summary, we could demonstrate that 6- and even 4-month-old infants can follow gaze, and that this early behavior depends on a sensitivity to both motion and gaze cues. Based on our findings, we argue that the perceptual narrowing hypothesis provides the most accurate account of early gaze following emergence. From this view, both perceptual and social cues may facilitate early emerging gaze following. By comparing young infants' performance with infants at 12 months we could furthermore show that, toward the end of the first year of life, infants become better at picking up the more informative social cues and become better at ignoring low-level responses. This study distinguishes the origin of gaze following from domain general perceptual development, supporting Scaife and Bruner's (1975) original notion, that "the child in its first year may be considered as less then completely egocentric" (p.266).

Data availability

The data that support the findings of this study are openly available in open science framework (OSF) at https://osf.io/hs8jq/?view_only=3c16dd9ed89645e089a66bcedd1bcf47Astor, K., Thiele, M., & Gredebäck, G. (2020). GAZMO [Raw eye-tracking data, workflow and the data matrix]. Retrieved from https://osf.io/hs8jq/?view_only=3c16dd9ed89645e089a66bcedd1bcf47

Acknowledgements

This study was funded by the following grants: Gustaf Gredebäck (Knut och Alice Wallenbergs Stiftelse: KAW 2012.0120). We thank Kristina Robertson for her help with data collection, all families that participated, and the actors: Freja Ohlin Ogestadh, Felicia Sundström, and Jessica Gunnarsson.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cogdev.2021.101121.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

Astor, K., & Gredebäck, G. (2019). Gaze following in 4.5- and 6-month-old infants: The impact of proximity on standard gaze following performance tests. *Infancy, 24*(1), 79–89. https://doi.org/10.1111/infa.12261

Astor, K., Lindskog, M., Forssman, L., Kenward, B., Fransson, M., Skalkidou, A., … Gredebäck, G. (2020). Social and emotional contexts predict the development of gaze following in early infancy. *Royal Society Open Science, 7*(9), Article 201178. https://doi.org/10.1098/rsos.201178

Baron-Cohen, S. (1995). Learning, development, and conceptual change. *Mindblindness: An essay on autism and theory of mind.* The MIT Press.

Barry-Anwar, R. A., Burris, J. L., Graf Estes, K., & Rivera, S. M. (2017). Caregivers and strangers: The influence of familiarity on gaze following and learning. *Infant Behavior and Development, 46*, 46–58. https://doi.org/10.1016/j.infbeh.2016.11.005

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B ((Methodological)), 57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*(6), 958–966. https://doi.org/10.1037/0012-1649.38.6.958

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*(6), 535–543. https://doi.org/10.1111/j.1467-7687.2005.00445.x

Brooks, R., & Meltzoff, A. N. (2014). *Gaze following: A mechanism for building social connections between infants and adults* (pp. 167–183). US;Washington;DC: American Psychological Association. https://doi.org/10.1037/14250-010

Brooks, R., & Meltzoff, A. N. (2015). Connecting the dots from infancy to childhood: A longitudinal study connecting gaze following, language, and explicit theory of mind. *Journal of Experimental Child Psychology, 130*, 67–78. https://doi.org/10.1016/j.jecp.2014.09.010

Catala, A., Mang, B., Wallis, L., & Huber, L. (2017). Dogs demonstrate perspective taking based on geometrical gaze following in a Guesser–Knower task. *Animal Cognition, 20*(4), 581–589. https://doi.org/10.1007/s10071-017-1082-x

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, 52*(1), 337–367. https://doi.org/10.1146/annurev.psych.52.1.337

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore, & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61–83). Lawrence Erlbaum Associates, Inc.

Corkum, V., & Moore, C. (1998). Origins of joint visual attention in infants. *Developmental Psychology, 34*(1), 28–38. https://doi.org/10.1037//0012-1649.34.1.28

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Deák, G. O. (2015). When and where do infants follow gaze? In: Proceedings of the Paper presented at the 182–187. Retrieved from ⟨https://doi.org/10.1109/DEVLRN.2015.7346138⟩.

Deák, G. O., Flom, R. A., & Pick, A. D. (2000). Effects of gesture and target on 12- and 18-month-olds' joint visual attention to objects in front of or behind them. *Developmental Psychology, 36*(4), 511–523. https://doi.org/10.1037/0012-1649.36.4.511

Del Bianco, T., Falck-Ytter, T., Thorup, E., & Gredebäck, G. (2018). The developmental origins of gaze-following in human infants. *Infancy, 24*, 433–454. https://doi.org/10.1111/infa.12276

Deligianni, F., Senju, A., Gergely, G., & Csibra, G. (2011). Automated gaze-contingent objects elicit orientation following in 8-month-old infants. *Developmental Psychology, 47*(6), 1499–1503. https://doi.org/10.1037/a0025659

Dezecache, G., Crockford, C., & Zuberbühler, K. (2019). The development of communication in alarm contexts in wild chimpanzees. *Behavioral Ecology and Sociobiology, 73*(8), 1–9. https://doi.org/10.1007/s00265-019-2716-6

D'Entremont, B. (2000). A perceptual–attentional explanation of gaze following in 3- and 6- month-olds. *Developmental Science, 3*(3), 302–311. https://doi.org/10.1111/1467-7687.00124

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*(6), 581–604. https://doi.org/10.1016/S0149-7634(00)00025-7

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition, 7*(6), 705–718. https://doi.org/10.1080/13506280050144399

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy: Special issue on developmental cognitive neuroscience. *Journal of Cognitive Neuroscience, 16*(8), 1320–1326.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy, 5*(1), 39–60. https://doi.org/10.1207/s15327078in0501_2

Gredebäck, G., Astor, K., & Fawcett, C. (2018). Gaze following is not dependent on ostensive cues: A critical test of natural pedagogy. *Child Development, 89*(6), 2091–2098. https://doi.org/10.1111/cdev.13026

Gredebäck, G., & Daum, M. M. (2015). The microstructure of action perception in infancy: Decomposing the temporal structure of social information processing. *Child Development Perspectives, 9*(2), 79–83. https://doi.org/10.1111/cdep.12109

Gredebäck, G., Fikke, L., & Melinder, A. (2010). The development of joint visual attention: A longitudinal study of gaze following during interactions with mothers and strangers. *Developmental Science, 13*(6), 839–848. https://doi.org/10.1111/j.1467-7687.2009.00945.x

Gredebäck, G., Johnson, S., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology, 35*(1), 1–19. https://doi.org/10.1080/87565640903325758

Gredebäck, G., Theuring, C., Hauf, P., & Kenward, B. (2008). The microstructure of infants' gaze as they view adult shifts in overt attention. *Infancy, 13*(5), 533–543. https://doi.org/10.1080/15250000802329529

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? the elicitation of gaze-following in 12–month-olds. *Developmental Science, 1*(2), 233–238. https://doi.org/10.1111/1467-7687.00036

Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental Science, 10*(1), 126–134.

Moore, C., Angelopoulos, M., & Bennett, P. (1997). The role of movement in the development of joint visual attention. *Infant Behavior and Development, 20*(1), 83–92. https://doi.org/10.1016/S0163-6383(97)90063-1

Morales, M., Mundy, P., Crowson, M. M., Neal, A. R., & Delgado, C. E. F. (2005). Individual differences in infant attention skills, joint attention, and emotion regulation behaviour. *International Journal of Behavioral Development, 29*(3), 259–263. https://doi.org/10.1080/01650250444000432

Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology, 21*(3), 283–298. https://doi.org/10.1016/S0193-3973(99)00040-4

Mundy, P. (2016). *Autism and joint attention: Development, neuroscience, and clinical fundamentals*. New York: The Guilford Publications.

Nyström, P., Falck-Ytter, T., & Gredebäck, G. (2016). The TimeStudio project: An open source scientific workflow system for the behavioral and brain sciences. *Behavior Research Methods, 48*(2), 542–552. https://doi.org/10.3758/s13428-015-0616-x

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*(5489), 265–266. https://doi.org/10.1038/253265a0

Téglás, E., Gergely, A., Kupán, K., Miklósi, Á., & Topál, J. (2012). Dogs' gaze following is tuned to human communicative signals. *Current Biology, 22*(3), 209–212. https://doi.org/10.1016/j.cub.2011.12.018

Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language, 42*(6), 1173–1190. https://doi.org/10.1017/S0305000914000725

Tomasello, M., Hare, B., Lehmann, H., & Call, J. (2007). Reliance on head versus eyes in the gaze following of great apes and human infants: The cooperative eye hypothesis. *Journal of Human Evolution, 52*(3), 314–320. https://doi.org/10.1016/j.jhevo1.2006.10.001

Triesch, J., Teuscher, C., Deák, G. O., & Carlson, E. (2006). Gaze following: Why (not) learn it? *Developmental Science, 9*(2), 125–147. https://doi.org/10.1111/j.1467-7687.2006.00470.x

Wahl, S., Michel, C., Pauen, S., & Hoehl, S. (2013). Head and eye movements affect object processing in 4-month-old infants more than an artificial orientation cue. *British Journal of Developmental Psychology, 31*(2), 212–230. https://doi.org/10.1111/bjdp.12001

Yamashita, W., Kanazawa, S., & Yamaguchi, M. K. (2012). The effect of gaze direction on three-dimensional face recognition in infants. *Vision Research, 68*, 14–18. https://doi.org/10.1016/j.visres.2012.06.022

Yu, C., & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye-hand coordination. *PLoS One, 8*(11), Article e79659. https://doi.org/10.1371/journal.pone.0079659