Children’s Use of the Temporal Dimension of Gaze for Inferring Preference

Shiri Einav and Bruce M. Hood
University of Bristol

This study examined 4- and 5-year-olds’ ability to spontaneously use the relative duration and frequency of another’s object-directed gaze for inferring that person’s preference. In Experiment 1, analysis revealed a strong age effect for judgment accuracy, which could not be accounted for by cue-monitoring proficiency. Reducing the saliency of the objects in Experiment 2 yielded significant improvement in the younger children’s performance. Thus, at 4 years, children already show signs of attending to the temporal dimension of gaze for making mentalistic inferences of preferential liking, but their competence may be undermined by the object choices themselves. By 5 years, they appear to overcome this competition. The obtained developmental difference is discussed in terms of concurrent transitions in attention regulation.

Keywords: gaze perception, mental state attribution, attention, nonverbal cues, social cognition

Despite the romanticism of the poetic adage that depicts the eyes as “the windows to the soul,” its underlying premise that the eyes reveal a wealth of information about the gazer’s mental states is widely supported by psychological research. Indeed, studies have shown that gaze behavior is, on the whole, a reliable indicator of a person’s focus of attention and therefore his or her object of reference or desire (for a review, see Argyle & Cook, 1976; Kleinke, 1986). Similarly, the available evidence confirms that adult observers are extremely sensitive to this nonverbal channel, utilizing it for determining people’s intentions, feelings, and beliefs (Baron-Cohen, 1995; Baron-Cohen, Wheelwright, & Joliffe, 1997; Kleinke, 1986).

These findings beg the question as to when young children become active interpreters of this subtle but invaluable social “language of the eyes” (Baron-Cohen, 1995, p. 108). Infants’ early sensitivity to eye gaze is well documented (e.g., Farroni, Csibra, Simion, & Johnson, 2002; Hood, Willen, & Driver, 1998; Symons, Hains, & Muir, 1998), as is their growing ability to reliably follow the gaze of their interactive partner (Butterworth & Jarrett, 1991; Corkum & Moore, 1998; Scaife & Bruner, 1975). Moreover, there is mounting evidence to suggest that between 12 and 18 months, infants begin to attend to the relation between a person and the object of their gaze, becoming increasingly aware of the intentional and referential nature of visual attention (Brooks & Meltzoff, 2002; Butler, Caron, & Brooks, 2000; Woodward, 2003). Social referencing paradigms have demonstrated that 1-year-olds are able to use gaze direction to relate an adult’s emotional display to a specific referent and predict goal-directed action (Baldwin & Moses, 1994; Phillips, Wellman, & Spelke, 2002; Repacholi, 1998; Sodian & Theormer, 2004). By 16 months, infants understand that a novel verbal label refers to the target of the speaker’s gaze and on this basis establish new word–object associations (Baldwin, 1991, 1995; Tomasello, 1995).

However, beyond infancy, comparatively little is known about how children’s use of gaze as a “mind-reading” tool continues to develop. Several attempts have recently been made to bridge the gap between theories of gaze processing in infancy and adults, with the emphasis being placed on children’s growing ability to attend to gaze direction cues for making mentalistic inferences. Converging evidence has revealed that 4-year-olds are proficient at using gaze direction alone for inferring the object of a gazer’s desire, intent, and reference (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995; Lee, Eskritt, Symons, & Muir, 1998). In Baron-Cohen et al.’s (1995) original study, preschoolers were presented with a static display depicting a schematic character, “Charlie,” whose gaze was directed at one of four sweets located in each corner of the picture. The children could accurately infer which sweet Charlie desired or was going to take from the direction of his eyes. Using a similar methodology, Lee et al. (1998) found that although younger children had no difficulty in correctly identifying the character’s focus of attention, they failed to use this knowledge to inform their desire inferences. In contrast, when the gaze cue was presented dynamically by an actor, even 3-year-olds were shown to utilize it for mind-reading purposes. Notably, this ability is absent among children with autism (Baron-Cohen et al., 1995)—a finding that has led to Baron-Cohen’s (1995) claim that sensitivity to the mentalistic significance of the
eyes is one crucial component in the development of a theory of mind (also see Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997).

Although this research has been invaluable in revealing when young children begin to read gaze in mentalistic terms, its focus on eye direction has led to the presentation of gaze as an all-or-nothing cue that is either directed at an object or not. As such, one may argue that the above paradigm goes only so far as to show that preschoolers are able to use gaze for making mentalistic inferences when it acts as a type of natural pointer that connects a mentalizing agent exclusively with the source of their attention in a very obvious way.

We propose that naturalistic gazing is more complicated. Individuals do not look only at those objects in their environment that they desire or intend to act on, while completely ignoring others. Rather, they look at the various potential targets available to them because, as is true of most decision-making situations, it is the very process of assessing the options that enables individuals to make their choice. That is not to say that all objects are gazed at equally. Crucially, as they scan the alternatives, individuals vary their gaze with respect to the degree of visual engagement that they allocate to the alternatives. So, for example, they may look at a particular object for just a brief moment and subject another object to lengthy scrutiny; they may glance at a certain person once and at another person repeatedly. Moreover, because such differences in the quantity of gaze are purposefully controlled by the gazer in accordance with their intentions, these differences reveal fundamental information about the individual’s underlying mental states. Nonetheless, these cues present observers with a more difficult interpretive task than that posed by the gaze direction cue. Because visual contact is made with more than one object, observers must be sensitive to the relative amount of visual attention allotted to different objects and interpret this measure appropriately. In a preferential choice situation, more gaze generally signals greater attentional interest in and a desire to learn more about a particular target, thereby identifying the object of choice (Kleinke, 1986).

There is evidence that adults are sensitive to these temporal aspects of gaze and readily use them for inferring others’ level of interest in their physical and social environments (see Argyle & Cook, 1976; Kleinke, 1986). Post and Hetherington (1974) have noted that for adults, “such variables as the presence, absence, frequency and duration of eye contact [all] influence the perception of a social relationship” (p. 882). However, little evidence exists as to when and how young children develop the ability to attend to the temporal dimension of displayed gaze behavior for making mentalistic inferences. Yet this is a key question that needs addressing if we are to progress in our understanding of how young children come to decode the language of the eyes.

To date, Montgomery, Bach, and Moran (1998) have carried out the only study that begins to examine this issue. They manipulated the length of time that an actor spent looking at each of two objects within a shopping context and asked children to judge which one the actor wanted to buy. Their results revealed that 4-year-olds failed to infer that an object that was gazed at extensively was more likely to be the protagonist’s goal than one that was glanced at momentarily. Instead, they were correct only when the protagonist looked solely at one of the objects in a control condition. That is, they responded to the direction of gaze but not to its duration. In contrast, 6-year-olds consistently based their decision on the actor’s prolonged looking, suggesting that at this age children are able to recognize that temporal gaze differences constitute a telling sign of another’s goals.

A serious limitation of Montgomery et al.’s (1998) experiment, however, is that participants did not have to pick up on the differences in looking time themselves, as they were made explicitly aware of these by narrated statements such as “She is looking here [at Object I] for a long time” (p. 695), which highlighted the protagonist’s extended looking toward the object of interest (Object I). According to Montgomery et al., this was done to help ensure that participants’ focus was maintained on the differences in looking behavior directed at each of the objects. We contend, however, that this methodology not only weakened the ecological validity of the presentation but also precluded a pure measure of children’s spontaneous disposition to attend to gaze duration for inferring a goal. In addition, looking behavior was not displayed solely by the actor’s eyes. Rather, the actor’s body and head were also oriented toward the gazed-at objects. Thus, the critical question of when children can spontaneously derive temporal gaze information from another’s eyes and independently use it as the basis for mentalistic inference has been left empirically open.

The present study sought to redress this by comparing 4- and 5-year-olds’ ability to spontaneously use directional and temporal eye gaze cues for inferring an actor’s preferred object. We chose these age groups because piloting found 6-year-olds to perform at ceiling across all tasks. In contrast to the above methodology, the experimenter did not in any way draw the child’s attention toward the actor’s exhibited gaze behavior during the trials, thus ensuring a purely spontaneous response measure. As well as assessing sensitivity to gaze duration, we were interested in examining children’s responsivity to gaze frequency. Although both temporal measures ultimately serve the equivalent function of varying the amount of visual attention directed at objects, they are perceptually very different, with gaze frequency presenting a more dynamic display toward the target. For the duration cue, the gazer looked at each of the objects once but displayed a comparatively prolonged look toward one in particular; for the frequency cue, the gazer displayed glances of equal length to each of the objects but exhibited a greater number of glances toward one. Thus, the gazer also spent cumulatively more time attending to the preferred object in the frequency condition. To control for this factor, we could have equated the length of single glances to the nontarget objects with the total duration of the repeated glances toward the target, but we reasoned that this would literally be sending out a mixed message to the child.

In summary, previous research has presented gaze simply as an all-or-nothing directional cue, whereas naturalistic gaze is in fact relative, varying along not only a spatial but also a temporal continuum, with mentalistic information being communicated by both dimensions. The current study captures rather than ignores this richness of looking behavior, thereby affording an investigation of gaze perception that is more commensurate with the way in which gaze presents itself to children in everyday life. In Experiment 1, we compared 4- and 5-year-olds on their ability to use gaze direction, duration, and frequency cues for inferring an actor’s preferred object. In Experiment 2, the saliency of the object choices was reduced to lower attention competition, which may have been masking the younger children’s competence on the temporal tasks.
Experiment 1

Gaze behavior was presented dynamically to the child through the use of short video vignettes of a real-life actor. This allowed the incorporation of movement as well as time into the presentation, thus offering more naturalistic stimuli than have previous studies that relied on static pictures alone (e.g., Baron-Cohen et al., 1995; Post & Hetherington, 1974). The stimuli were prerecorded as this had the advantage of ensuring the same controlled variation of the gaze variable for all participants (see Mumme & Fernald, 2003).

The direction cue condition was modeled after Baron-Cohen’s (1995) Charlie task, with the protagonist looking in the direction of only one picture from an array in front of her. In the temporal cue conditions, the protagonist looked at all of the pictures in turn but showed increased gaze duration or frequency toward one in particular. We decided to use pictures as the gazed-at objects because they are naturally interacted with through purely visual means in everyday life. Other objects (such as toys) are tactically manipulated and therefore simply gazing at them would not have represented a situation with which children are familiar.

In the preference tasks, participants were required to watch each vignette and decide which picture the protagonist liked the best for a total of four trials per condition. In addition to obtaining a forced-choice behavioral response from participants, we assessed their reflective awareness of the gaze cue and its meaning. We were interested in finding out whether children who provided correct preference judgments were cognizant of their use of the gaze cue or were responding to it implicitly. Therefore, following each preference judgment, participants were required to provide a justification for their decision.

Finally, we wished to verify whether errors on the temporal tasks could be attributed to problems in detecting the gaze cue at the perceptual level. Was it the case that these children could not even discriminate that one picture was being gazed at for a longer duration or a greater number of times? Or was it the case that they could accurately identify the relevant gaze cue but failed to recognize the importance of this information for inferring the gaze’s preference? To assess this, we also administered a posttest cue-detection task in which participants simply had to identify the picture that had been targeted by the gaze cue (i.e., the picture that had been gazed at the longest or the most times). On the basis of previous findings (Doherty & Anderson, 2000; Lee et al., 1998), we anticipated that few children would have problems in identifying the direction cue; however, it was conceivable that 4- and 5-year-olds would have more difficulty with the duration and frequency cues for three reasons. First, discrimination of the cues might have been harder because it involved the actor making perceptual contact with all of the pictures, not just one. Second, sustained attention was required to a greater degree as these cues entailed monitoring quantities over time. Third, children had to remember the cues as they were not available at the time of decision making, in contrast to the direction cue that did remain on display.

Method

Participants. Forty 4-year-olds (19 boys and 21 girls; mean age = 4 years 8 months; range = 4 years 4 months–4 years 11 months) and forty 5-year-olds (19 boys and 21 girls; mean age = 5 years 4 months; range = 5 years 0 months–5 years 10 months) took part in this experiment after their parents provided informed consent. An additional 4 participants were omitted from the original sample and replaced because of noncompliance. The children were all native English speakers, recruited from two schools in a predominantly middle-class area in Bristol, England.

Design. To conduct a strict comparison, we opted for a within-subject design, with each child being tested on all three cue tasks: gaze direction, gaze duration, and gaze frequency. Participants were randomly assigned to one of four task orders: Order 1 = direction–frequency–duration; Order 2 = direction–frequency–duration; Order 3 = duration–frequency–direction; and Order 4 = frequency–duration–direction. Thus, the direction task was presented either before (Orders 1 and 2) or after (Orders 3 and 4) the two temporal tasks, and the relative positions of the duration and frequency tasks were fully counterbalanced (Orders 1 and 3 presented duration before frequency, and Orders 2 and 4 presented frequency before duration). A follow-up cue-detection task was administered to all participants who failed to score full marks on the corresponding preference task. This constituted a check of their ability to accurately discriminate and monitor the relevant gaze cue.

Materials. The stimuli consisted of short video vignettes, presented as QuickTime movies on a Mac G4 laptop, and featured a head and shoulders view of an adult female protagonist sitting at a table, surrounded by pictures. The pictures depicted everyday objects, animals, and plants as well as abstract shapes and patterns. Each condition was enacted by a different protagonist to avoid response bias. (See Figures 1a and 1b for examples of the two stimuli setups.) The direction condition consisted of four pictures, located at each corner of the screen, with the protagonist situated in the center. The duration and frequency conditions included only three pictures because of the increased concentration and memory demands imposed by these trials. The pictures were raised up in front of the protagonist, one to her left, one to her right, and one in the middle. Following Baron-Cohen et al.’s (1995) methodology, the participant was first asked to identify his or her own favorite picture in the direction condition. The target was then randomly selected out of the remaining three options to ensure that it did not coincide with the child’s preference. This method was not applied to the temporal conditions, however, for two reasons. First, Baron-Cohen et al.’s (1995) Charlie paradigm enabled the experimenter to choose online a target that was not preferred by the child because the apparatus was a simple paper and acetate presentation that could be easily manipulated. We produced multiple video versions of the direction cue that allowed us to replicate this procedure in our direction task, but practical constraints prevented this for the more complex duration and frequency vignettes. Second, we were concerned that during the experiment, participants would learn to rule out the picture that they had chosen as their favorite as a potential target of the protagonist. Therefore, in the temporal tasks, the position of the target picture was varied according to a predetermined random order, subject to the constraint that over the course of the four trials, the target would be featured in each of the three positions on at least one occasion. The position of the target on the fourth trial was fully counterbalanced across participants. Within conditions, novel pictures were used from one trial to the next. To keep the stimuli constant across temporal conditions, we always equated the picture trios used in the duration and frequency tasks. On certain trials, the preferred picture was the same for the two conditions, and on other trials it was different, according to the randomization procedure.

Gaze cues followed a standardized pattern. In the direction condition, each trial consisted of the protagonist looking straight ahead and then shifting her gaze toward only one of the four pictures. This end position was captured in a still image and retained on screen as the participant was questioned, in accordance with Baron-Cohen et al.’s (1995) method. In the duration condition, the protagonist gazed at every picture in turn (proceeding from left to right), indicating her preference by staring at one of the pictures for 10 s and glancing only momentarily (~1 s) at the other two.
Richie and Bickhard (1988) have demonstrated that 3- to 6-year-olds can perceptually distinguish between similar lengths of time (e.g., 1 s vs. 7 s). In the frequency condition, the protagonist also viewed every picture in turn (proceeding from left to right), indicating her preference with a 1:2:6 glance ratio. Each glance lasted approximately 1 s. In both temporal conditions, the protagonists’ eyes returned to face the camera before the image was stilled and questioning began. Vignettes were approximately 17 s. For all of these trials, looking was expressed using eye gaze only, without accompanying head movements. (The head remained straight, facing the participants throughout.) All other nonverbal cues, such as facial expressions, were kept as neutral as possible to minimize confounding. Piloting with 10 adults found that they readily perceived the intended preference. In the detection task, participants were presented with novel vignettes for each condition, which followed exactly the same format as those used in the preference task.

**Procedure.** All participants were tested individually in a quiet room. The experimenter explained to the participants that they would be playing a game in which they had to watch a woman looking at pictures and decide which one she liked the best each time. Participants were instructed that they could respond by either naming or pointing to their answer. Every condition began with a brief videotaped narration, introducing the protagonist with her pictures to the participant. To ensure that participants realized that although the protagonist was seated behind the pictures she could still see the same pictures that they could see, the actor turned over each of the cards to demonstrate that identical pictures were painted on both sides. A practice trial was then presented to familiarize the child with the task and the explanatory context of preferential liking. Specifically, children viewed the protagonist enacting the relevant gaze cue and were asked to choose her preferred picture and provide a reason for their answer. Following the participant’s response, the experimenter said, “Let’s see if you are right. Now she is going to tell us which one was her favorite,” and a clip was shown in which the protagonist pointed to the correct answer while smiling and saying, “I like this one the best!” Thus, the feedback revealed to participants only whether they had made the correct picture choice but made no reference whatsoever to the gaze cue. Four experimental trials were then administered. At the end of each vignette, the video was paused, leaving a still frame of the protagonist and the pictures in view on the screen. The participant’s gaze-reading ability was then assessed first by a preference judgment question, “Which picture do you think she liked the best?” and second by a justification question, “How do you know that was her favorite one?” No feedback concerning accuracy was given on any of the experimental items. Throughout the experiment, participants were allowed to take as much time as needed to answer the questions, and a single repetition of the vignette was permitted if the child’s attention had not adequately focused on the screen during the trial. If the child was unsure of responding, the experimenter explained that guessing was allowed.

Following analysis, children who were found to have scored less than full marks on any particular preference task were administered a detection task on the corresponding gaze cue. This assessment took place in a separate testing session, approximately a week later. In this task, participants were instructed to indicate the picture that the protagonist was looking at, had looked at the longest, or had looked at the most times for the direction, duration, and frequency conditions, respectively. There were four detection trials for each gaze condition.

**Scoring.** Participants received 1 point for each correct preference judgment to obtain an accuracy score between 0 and 4 per gaze cue condition. Justifications for responses were coded into four categories: (a) gaze (e.g., “She looked at it for the longest”; “She looked at that one lots of times”), (b) description of chosen picture (e.g., “Because it’s magic”; “Because girls like pink”), (c) egocentric (e.g., “Because I love flowers”; “Because I like that one the best”), and (d) residual (any other response, including “Don’t know,” “It’s just a guess,” paraphrasing, and nonsense justifications). This coding scheme was used to individually code all trials for each
participant, awarding 1 point for every gaze-based justification given to explain their response (out of 4 per gaze cue). In addition, participants received an overall categorization that reflected their predominant explanation type for each condition. An independent rater coded half of the responses from each age group (Cohen’s $\kappa = .91$). Disagreements were resolved by discussion. Those participants who underwent the cue-detection task received a score between 0 and 4 per gaze condition.

Results

Preference judgments. Preliminary analysis indicated no effect of gender on performance; therefore the data were collapsed on this dimension in all further analyses. Table 1 presents the mean score for each of the gaze cue conditions, as a function of age. The overall chance judgment score was 1.00 for the direction task (probability of success = $\frac{1}{4}$ for each trial $\times 4$) and 1.33 for the temporal tasks (probability of success = $\frac{1}{3}$ for each trial $\times 4$). Initial planned comparisons (one-sample $t$ tests) between the chance level and the means of each age group revealed that at both ages children performed significantly above chance on all three tasks ($p < .01$). However, as Table 1 clearly shows, there was a considerable improvement in judgment accuracy with age, particularly on the temporal tasks, and children performed better on the direction cue than on the duration and frequency cues.

These observations were confirmed by a 3 (gaze cue: direction, duration, or frequency) $\times$ 2 (age: 4- or 5-year-olds) $\times$ 4 (task order: direction–duration–frequency, direction–frequency–duration, duration–frequency–direction, or frequency–duration–direction) mixed analysis of variance, with repeated measures on order: direction–duration–frequency, direction–frequency–duration, or frequency–duration–direction) mixed analysis of variance, with repeated measures on order. However, a significant interaction between gaze cue and order emerged, $F(6, 144) = 3.45, p < .005, \eta^2 = .13$. Further analyses revealed that this stemmed from the relative positions of the direction and frequency tasks affecting performance, rather than from the position of the direction task with respect to the two temporal tasks. Moreover, post hoc independent $t$ tests showed that the interaction with the gaze cue was a result of only the duration task being affected by this factor. Performance on the duration task was significantly higher if the task was administered after the frequency task rather than before it, $t(78) = 3.02, p < .005, \eta^2 = .10$, whereas performance levels on the direction and frequency tasks were consistent, regardless of task position.

Individual data. In addition to the group analysis, we wished to consider performance at an individual level to gauge, in particular, the relative frequencies of 4- and 5-year-olds who judged the protagonist’s preference correctly over the majority of the trials. Table 2 shows the number of children per age group answering zero, one, two, three, or all four trials correctly for each of the gaze cue conditions. Inspection of Table 2 shows that a high proportion of participants of both ages scored at least 3 correct trials on the direction task, but only the older children maintained this level on the duration (88%) and frequency (95%) tasks. The 4-year-olds were a more variable group, with just 55% managing a score of 3 or 4 on each of the temporal tasks. To determine whether response errors reflected positional biases (e.g., a propensity to typically pick the last picture that was gazed at as the answer), we analyzed the error pattern for individual participants. This analysis revealed that no single picture position produced a consistent bias and that the most common pattern for children’s errors was mixed across positions.

Justifications. As shown in Table 1, performance levels on the justification component of the tasks fell short of the demonstrated judgment accuracy, particularly among the 4-year-olds. This suggests that some trials were solved by participants who were not explicitly aware of their use of the gaze cue or, at least, were not able to articulate it. Additionally, it is possible that in certain instances, the correct answer was selected by chance or because the participant’s own preference happened to coincide with the protagonist’s target.

However, looking specifically at the data for the subset of participants who consistently picked the correct target on the majority of the trials (i.e., those achieving a judgment score of at least 3 out of 4), we found that these children were also very good at providing gaze-based explanations for each of their correct responses. On the direction task, all of the 5-year-olds and 86% of the 4-year-olds who judged the protagonist’s preference correctly on at least three occasions justified all their correct responses with reference to the gaze cue. These proportions were comparable on the temporal tasks (duration: 5-year-olds = 86%, 4-year-olds = 90%; frequency: 5-year-olds = 97%, 4-year-olds = 88%), indicating that most of these participants were fully aware of their use of the gaze cues and were able to articulate it to the experimenter.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean Judgment and Justification Scores as a Function of Gaze Cue for Each Age Group in Experiment 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gaze cue</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaze cue</td>
<td>Score</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>Judgment score</td>
</tr>
<tr>
<td></td>
<td>Justification score</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>Judgment score</td>
</tr>
<tr>
<td></td>
<td>Justification score</td>
</tr>
</tbody>
</table>

Note. Maximum score = 4; $n = 40$ in each age group.
We also wished to inspect the frequencies of the nongaze explanations that were offered by participants whose preference scores were incorrect on at least half of the trials (i.e., those not achieving a score of 3 or 4). The relative proportions of the various predominant explanation types are displayed in Figure 2 for the direction, duration, and frequency tasks, respectively. A notable trend that is evident for all three cue tasks is the large percentage (50%–65%) of participants who primarily justified their response by alluding to the picture content. Such descriptions usually referred to the appearance or function of the depicted object (e.g., “Because it’s shiny” or “Because it is comfortable to sit in”), even in the case of abstract pictures (e.g., “Because there is a green square and a red square”).

Cue-detection task. A follow-up cue-detection task was administered to any participant who did not achieve full marks on the corresponding preference task (direction, n = 15; duration, n = 33; frequency, n = 32). We found that all of the children who underwent the direction-detection task obtained the maximum score of 4, thereby demonstrating that they could detect the direction cue perfectly. On the temporal detection tasks, every participant managed to correctly identify the picture that was gazed at the most by the protagonist on at least three of the four trials (duration: 60% of the participants obtained a maximum score of 4, with the remainder all scoring 3; frequency: 44% of the children scored 4, and the remainder scored 3). The data suggest that these children had no trouble in monitoring and discriminating the relevant temporal gaze cue. Independent-samples t tests confirmed that there was no systematic difference in preference task performance between those participants who were at ceiling on the detection task and those who erred on one trial. Specifically, the findings revealed that participants who answered all four detection trials accurately had not performed any better on the corresponding preference task than those children who scored only 3 correct trials, for both the duration and frequency cues, t(28) = 0.94, and t(29) = 0.06, respectively.

**Discussion**

Experiment 1 revealed that when asked to judge an actor’s preferred picture, 5-year-olds spontaneously and competently made use of the actor’s displayed direction, duration, and frequency of gaze toward the pictures. Moreover, when asked to justify the reasoning for their judgment, they could explicitly articulate their use of the gaze cues. In contrast, 4-year-olds performed similarly well on the direction task, replicating previous findings (Baron-Cohen et al., 1995; Lee et al., 1998; Montgomery et al., 1998), but had difficulty with the duration and frequency

---

**Table 2**

*Observed Frequencies of Correct Responses as a Function of Gaze Cue and Age in Experiment 1*

<table>
<thead>
<tr>
<th>Condition and age</th>
<th>Observed score frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Observed score frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>0</td>
</tr>
<tr>
<td>4 years</td>
<td>1</td>
</tr>
<tr>
<td>5 years</td>
<td>3</td>
</tr>
<tr>
<td>Duration</td>
<td>2</td>
</tr>
<tr>
<td>4 years</td>
<td>1</td>
</tr>
<tr>
<td>5 years</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Figure 2.** Frequency distributions of predominant explanation types given by participants whose judgment score was less than 3 out of 4 on the direction (n = 8), duration (n = 23), and frequency (n = 20) tasks in Experiment 1.
tasks. Although as a group they did perform above chance level, only around half of these younger participants showed consistent evidence of using the protagonist’s relative gaze behavior as the basis for their preference judgments.

Overall, participants’ responses were comparable across the two temporal cues, implying that they did not find one easier to use than the other, despite the perceptual differences in cue composition. However, an Order × Gaze Cue interaction emerged from our data, which suggests that in fact a subtle yet significant difference may exist between children’s sensitivity to these measures. The interaction revealed that those participants who had been administered the frequency task prior to the duration task performed significantly better on this latter cue than did those who were presented with the duration task first. In contrast, performance on the frequency task was consistent irrespective of whether it had been preceded by the duration task. It seems that the level of performance on the duration task was dependent on the priming afforded by the preceding frequency gaze cue (for a related scaffolding effect, see Lee et al., 1998). This result makes intuitive sense given the combined salience of the more dynamic activity and cumulative duration information displayed by the frequency cue toward the target, which may have served to sensitize participants to the relative nature of the protagonist’s gazing in the subsequent duration task.

So, what accounts for the comparatively poor performance of the younger children on the temporal gaze tasks? Results from the detection tasks indicate that they had no problems in registering the temporal cues perceptually. A caveat is necessary here, however, to acknowledge that our detection trials go only so far as to demonstrate children’s ability when their attention is specifically directed at the gaze cue. This procedure cannot guarantee that these children were equally able to extract the gaze information spontaneously during the preference task. Nevertheless, running the detection task separately was preferable as incorporating the detection question into the preference task itself would have risked alerting children to the role of the gaze cue (see Lee et al., 1998, for similar reasoning).

A second, more likely, explanation is that 4-year-olds’ difficulties with the duration and frequency preference tasks were at the level of cue interpretation rather than cue perception. According to this hypothesis, participants registered that the protagonist gazed at one picture more than the others but did not recognize the underlying mentalistic significance of such an act and so did not use this information when making their judgment.

A third possibility must also be considered—namely, that the children were capable of both perceiving and interpreting the cues appropriately but their competence was somehow masked in our experiment. An examination of the justification data revealed the picture content as a potential source of such masking. We found that a large proportion of the low-scoring participants in each of the cue conditions predominately justified their choice by alluding to the picture content. We cannot be certain that the children’s retrospective justifications for their preference judgments were accurate reflections of their actual decision-making process. It is conceivable that the picture content had no bearing on the judgment itself but its description simply provided participants with a convenient way of rationalizing a randomly chosen picture.

Alternatively, it may have been the case that the picture content had in some significant way influenced the preference judgments. The highly salient nature of the colorful picture content may have proven too compelling a distraction for some of the younger participants to ignore when watching the vignette and making their decision. Such a surmise is compatible with recent demonstrations of preschoolers’ limitations in executive control (e.g., Hala & Russell, 2001; Zelazo, Carter, Reznick, & Frye, 1997). Thus, before we can come to any firm conclusions about 4-year-olds’ abilities and speculate on the underlying mechanisms for the obtained developmental change between 4- and 5-year-olds, we propose a modification to our paradigm that will allow us to assess the above possibility.

Experiment 2

Experiment 2 examined whether 4-year-olds’ performance on the temporal tasks could be significantly improved by decreasing the saliency of the pictures. This modification was achieved by creating a new blank condition in which the picture content was removed from the stimuli such that only the blank backs of the pictures were observable. It was expected that younger children would perform better in the blank than in the original picture condition if the distraction posed by the saliency of the picture content was a key factor hindering their performance in Experiment 1. If, conversely, 4-year-olds’ difficulties with the task were fundamentally due to a lack of insight into the meaning of the temporal gaze dimension, then such a manipulation should not be sufficient to improve performance levels.

Because of the obtained order effect in Experiment 1, we decided not to test both temporal gaze cues within participants but, rather, to adopt a between-subjects design such that participants were tested on either the duration or the frequency cue. In addition, the picture–blank manipulation was varied between participants to avoid any carryover effects and ensure purely spontaneous measures for each condition.

Method

Participants. One hundred predominantly White middle-class children (50 boys and 50 girls) took part in this experiment after their parents gave informed consent. None had participated in Experiment 1. The mean age was 4 years 7 months, ranging from 4 years 1 month to 5 years 0 months. An additional 4 participants were omitted from the original sample and replaced because of attentional limitations (2) or failure to comprehend task demands (2).

Materials and procedure. Participants were tested individually. They were randomly assigned to one of four groups: picture duration, blank duration, picture frequency, or blank frequency, on the provision that these would comprise approximately equal numbers of girls and boys. The picture duration and picture frequency conditions used the same stimuli that were used in the duration and frequency tasks in Experiment 1. The blank duration and blank frequency conditions were identical in format to their respective picture conditions, except that only the blank backs of the pictures were visible to the participant rather than the pictures themselves (see Figures 1b and 1c for a comparison of the stimuli setup). During the initial introduction to the blank condition, the protagonist turned over each of the cards and showed the participant that pictures did in fact exist on her side of the cards. However, it was made clear that these pictures were just an example and were not the ones viewed by the protagonist during the game, which were to remain concealed. Participants received one practice trial (with accuracy feedback) and four experimental trials. The location of the preferred picture on each trial was varied according to the pseudoran-
domination technique described in Experiment 1 and equated across all conditions. Each video vignette lasted approximately 17 s. The testing procedure was identical to Experiment 1, with participants first asked to make the preference judgment (those in the blank condition responded by pointing to their answer, and children in the picture condition could either name or point to their choice) and second to provide a justification for their decision. The same scoring and coding methods were used as in Experiment 1. An independent rater coded half of the responses from each condition (Cohen’s $\kappa = .97$). Disagreements were resolved by discussion.

## Results

**Preference judgments.** Preliminary analysis revealed no gender effects; therefore the data were collapsed on this dimension in all further analyses. Table 3 presents the mean score for each condition group. Initial planned comparisons (one-sample $t$ tests) against chance level revealed that all groups performed significantly above chance on the judgment question ($p < .02$). However, it is clear from Table 3 that the two blank groups performed considerably better than the two picture groups.

A 2 (gaze cue: duration or frequency) $\times$ 2 (condition: picture or blank) between-subjects analysis of variance yielded a main effect of condition, $F(1, 96) = 9.10$, $p < .005$, $\eta^2 = .09$, confirming that participants’ performance was significantly higher in the blank than in the picture condition. Although scores obtained by the two frequency cue groups were overall higher than those obtained by the two duration cue groups, the effect of cue and the interaction between cue and condition were not significant, $F(1, 96) = 0.92$, $ns$, $\eta^2 = .01$, and $F(1, 96) = 0.01$, $ns$, $\eta^2 = .00$, respectively. Errors were not accounted for by a positional bias in responding.

**Justifications.** Table 3 shows that performance on the justification question was once again lower than on the judgment question in all treatment groups. However, as in Experiment 1, we found that a large majority of participants who obtained a score of 3 or 4 were capable of justifying all of their correct responses with reference to the appropriate temporal aspect of the actor’s gaze behavior, thereby demonstrating full awareness of their use of the cue (89%, 75%, 92%, and 90% of the picture duration, blank duration, picture frequency, and blank frequency groups, respectively). Among the remaining children in the two picture groups, (i.e., those who did not score 3 or 4), a description of the picture content was once more the most common form of justification.

## Discussion

The results of Experiment 2 demonstrate that 4-year-olds were much more competent at spontaneously using gaze duration or frequency for inferring another’s preference when the object choices were not visible to them. The obtained difference in performance between the picture and blank conditions indicates that the salient picture content did indeed play an active and hampering role in children’s responses. Moreover, it suggests that those failing participants, who predominantly justified their answers by alluding to the picture content, were not simply rationalizing random choices but, rather, were reliably reporting factors that had influenced their decision. Further, 4-year-olds’ high attainment on the blank trials helps to rule out the possibility that factors inherent in the procedure (e.g., comprehension of task instructions, memory requirements, or the actors’ neutral facial expressions) were responsible for their problems in identifying the preferred object in the picture condition.

It appears that the majority of 4-year-olds are sensitive to the relative amount of visual attention allotted to different objects by others and can interpret the underlying implications of this measure appropriately. Nevertheless, when faced with a salient, attention-grabbing object array from which to choose, their performance levels are significantly compromised. We propose two alternative explanations to account for these findings. The first is that while watching the video sequence, children involuntarily or deliberately fixated their limited attentional resources on the pictures, at the expense of the displayed gaze behavior, and this may have interfered with their encoding of the gaze cue. Our detection trials in Experiment 1 demonstrated that participants could accurately monitor the temporal cues, which would suggest that they were not being unduly distracted by the pictures while watching the vignettes. However, these results do not necessarily guarantee that this was equally the case within the context of the preference trials, in which the child’s attention was not explicitly directed at the actor’s gaze behavior.

A second explanation is that children did attend to the gaze cue while watching the video sequence but when it came to making a choice about the protagonist’s preference, their prepotent response was to base their judgment on the picture content. This may have been due to participants egocentrically attributing their own preferences to the protagonist or selecting pictures that they believed the protagonist would like the best. In this way, the picture content may have weighed heavily in the decision-making process, overriding any information derived from the temporal gaze cue.

## General Discussion

To our knowledge, the present study is the first to address the fundamental question of whether young children can spontaneously utilize temporal eye gaze information as a cue to an agent’s preference. Experiment 1 found that 5-year-olds could readily attend to the relative duration and frequency of an actor’s displayed gaze toward a selection of pictures to infer her preference. In contrast, 4-year-olds were comparatively poor at this task, even though they were able to identify which picture was being looked

<table>
<thead>
<tr>
<th>Condition</th>
<th>Pictures</th>
<th></th>
<th>Blanks</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>SD</td>
<td>$M$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Duration</td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td>Judgment score</td>
<td>2.08</td>
<td>1.47</td>
<td>2.88</td>
<td>1.30</td>
</tr>
<tr>
<td>Justification score</td>
<td>1.48</td>
<td>1.78</td>
<td>2.00</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Judgment score</td>
<td>2.32</td>
<td>1.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Justification score</td>
<td>1.88</td>
<td>1.86</td>
</tr>
</tbody>
</table>

Note. Maximum score = 4; $n = 25$ per cell.
at the longest or the most frequently in a follow-up assessment. However, Experiment 2 revealed that when 4-year-olds were presented with a modified version of the paradigm, in which the picture contents were no longer visible to participants, their ability to make consistent use of these gaze cues significantly improved.

Thus, the main theoretical implication of the present findings is that 4- and 5-year-olds’ gaze-reading abilities are considerably more advanced than is currently assumed in the developmental literature. To date, children of this age have been credited only with the capacity to interpret directional information afforded by the eyes (Baron-Cohen et al., 1995; Lee et al., 1998; Montgomery et al., 1998). Yet, it is evident from our results that their understanding extends to both the spatial and temporal dimensions of gaze behavior. Not only can they infer the desired object of an agent who looks exclusively at one target, but they can also infer the preference of an agent who looks at all possible options but displays comparatively more gaze toward one in particular. We believe that the latter situation constitutes a much more naturalistic representation of the way in which people visually inspect potential targets. Consequently, the importance of children’s developing sensitivity to the temporal nuances of gaze behavior cannot be overstated.

This is especially true in the social arena in which variations in looking time and frequency are often reliable indicators of the nature of the relation (e.g., affectionate vs. distant) that exists between interacting individuals (Burgoon, Buller, Hale, & de-Turck, 1984; Kleinke, 1986). An ability to correctly interpret such cues is clearly an essential tool for generating accurate social representations, both when the target relation applies to observable others and when it involves the self. Indeed, as Montgomery, Moran, and Bach (1996) have pointed out, the value of this skill for adaptive social interaction is underscored by the social impairments that children with autism and Asperger syndrome present, which are thought to be related to gaze-reading deficits (Baron-Cohen, 1995; Baron-Cohen et al., 1995; Leekam et al., 1997). There is at present no empirical evidence regarding children’s use of temporal variations in eye contact for making social relational judgments. Whereas our work has concentrated on gaze behavior that is object focused, we propose that future research efforts be directed toward building on the current basic paradigm by placing it in a socially embedded context.

An important question to consider is whether our findings can be said to reflect a deep conceptual understanding of how the temporal aspect of gaze and liking are related or a less profound behavioral association that does not represent the connection itself. The behavioral association that does not represent the connection itself.

Temporal aspect of gaze and liking are related or a less profound vectorial component of gaze as they understand (like adults) that individuals are intentional beings who deliberately control the amount of visual attention that they allocate to objects, in accordance with their internal desires. Alternatively, a more conservative interpretation would maintain that children of this age are merely following a behavioristic length-of-association strategy that connects the individual with the target that is looked at the most, without imputing the mediating process of purposeful selective attention.

This issue of levels of interpretation has been raised in numerous infancy studies (e.g., Phillips et al., 2002; Repacholi & Gopnik, 1997) and is clearly worth highlighting to avoid unjustifiably overestimating the extent of children’s mentalistic understanding (Montgomery et al., 1998). We acknowledge that the methodology used in the current study does not enable us to assert with any certainty that our participants’ performance was in line with the richer interpretation. However, given that 3- and 4-year-olds’ use of gaze direction has been taken as evidence of their ability to read behavior in terms of attentional and volitional mental states (Baron-Cohen & Cross, 1992; Baron-Cohen et al., 1995), we might expect 4- and 5-year-olds’ use of gaze duration and frequency to be an extension of this ability. Moreover, this view is supported by what we know about children’s capabilities at this age. It has been established that by 3 years, children routinely construe action via an intentional stance, viewing it as “purposeful and thus shaped by and manifesting the actor’s intentional states such as his or her goals, preferences, attention and knowledge” (Phillips et al., 2002, p. 72) rather than as mere behavioral regularities (Bartsch & Wellman, 1995; Flavell & Miller, 1998). In light of these findings, it is reasonable to assume that the protagonist’s extended looking constituted much more than an opaque agent–object association for our 4- and 5-year-olds.

Nonetheless, the possibility exists that preschoolers might mistakenly believe that increased gaze invariably indicates preference, whereas in reality it can signal a range of emotional and cognitive reactions (such as fear, vigilance, etc.), according to situation. In building on previous gaze-reading studies (e.g., Baron-Cohen et al., 1995; Lee et al., 1998; Montgomery et al., 1998), we chose to present the gaze cues within the specific context of preferential liking. Further research is needed to clarify whether, given different setting conditions, children’s use of temporal gaze parameters will extend to the inference of other mental states.

Although the present study has provided evidence of 4-year-olds using temporal differences in gaze behavior for inferring preference, it has simultaneously demonstrated their limitations on this task when the cues occur in the context of a competitively salient object array. There are two related explanations to consider. First, children may have failed to use gaze in the presence of the pictures because their responses were primarily driven by expectations about what should be the appropriate choice for the actor, irrespective of her gaze behavior. This possibility is consistent with a recent finding by Rieff, Tervoq, Koops, Stegge, and Oomen (2001) that 4-year-olds have difficulty judging another’s preference independently of their beliefs about desirability (e.g., gender-stereotyped expectations about preferences).

A second major source of competition may be attributed to the need to divide limited attentional resources between the actor’s gaze and potential targets that are actively recruiting attention. It is possible that during the trial presentation, the preschoolers’ focus of attention was primarily directed at the pictures, at the expense of the displayed gaze behavior, thereby decreasing their chances of encoding the gaze cue properly. In contrast, the finding that 5-year-olds passed the picture condition with ease suggests an increasing ability to disengage one’s attention from perceptually salient but irrelevant information. This account is supported by evidence that attentional regulation undergoes a transition period.
around this age. In a study tracking the ability of 4-, 5-, 7-, and 9-year-olds to disregard task-irrelevant information, Pasto and Burack (1997) found that the most significant changes occurred between the ages of 4 and 5 years. Four-year-olds showed a comparative deficit in narrowing the focus of their attentional lens and ignoring distractors. Moreover, they were the only group whose performance was facilitated by the presence of a window cue that served to artificially narrow the visual area to which attentional resources should be directed (a manipulation that resem-bles our blank condition, which essentially reduced the scene to the protagonist’s face as the pictures were no longer visible). In addition, there is evidence that 4-year-olds have difficulty exercising inhibitory control over prepotent responses. Using a Stroop-like day–night task, Gerstadt, Joo Hong, and Diamond (1994) found that children less than 5 years had great difficulty suppressing a response of “day” when presented with a card showing a bright sun and a response of “night” when shown a black card with moon and stars (also see Livesey & Morgan, 1991). Both sets of findings have obvious implications for the present experiments in which these developmental differences may have been further compounded by the fact that our participants were not explicitly instructed to disregard the pictures while watching the vignettes or to inhibit any tendency to base their choice on the picture content. They had to essentially enforce such attentional self-regulation spontaneously, which may have proven too much for the younger children.

By substituting the pictures with nondistracting and uninforma-tive blank cards, the blank condition afforded 4-year-olds the opportunity to fully attend to and interpret the protagonist’s gaze behavior without potential competition from the pictures arising because of either, or indeed both, of these relevant factors. In this way they were better able to demonstrate their underlying competence, and performance levels drew nearer to that of the older children. However, an important point to make is that in real life, observers of a gazing individual are unlikely to be granted this kind of assistance, because everyday objects are not blank and do not tend to be hidden from view. On the contrary, the targets of people’s gazing are likely to be salient and pose attentional competition for nearby observers who may have their own conflicting desires and intentions with regards to these targets, be they objects or other individuals. The present findings indicate that although the foundational sensitivity to the temporal dimension of gaze is in place at 4 years, the ability to make functional use of this information within a naturally salient context is still developing and might be fully attained only at 5 years.

References


Received July 28, 2004
Revision received May 11, 2005
Accepted May 16, 2005

---

**Call for Nominations**

The Publications and Communications (P&C) Board has opened nominations for the editorships of *Behavioral Neuroscience*, *JEP: Applied*, *JEP: General*, *Neuropsychology*, *Psychological Methods*, and *Psychology and Aging* for the years 2008–2013. John F. Disterhoft, PhD; Phillip L. Ackerman, PhD; D. Stephen Lindsay, PhD; James T. Becker, PhD; Stephen G. West, PhD; and Rose T. Zacks, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2007 to prepare for issues published in 2008. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations also are encouraged.

Search chairs have been appointed as follows:

- **Behavioral Neuroscience**: Linda P. Spear, PhD, and J. Gilbert Benedict, PhD
- **JEP: Applied**: William C. Howell, PhD
- **JEP: General**: Peter A. Ornstein, PhD
- **Neuropsychology**: Susan H. McDaniel, PhD, and Robert G. Frank, PhD
- **Psychological Methods**: Mark Appelbaum, PhD
- **Psychology and Aging**: David C. Funder, PhD, and Leah L. Light, PhD

Candidates should be nominated by accessing APA’s EditorQuest site on the Web. Using your Web browser, go to http://editorquest.apa.org. On the Home menu on the left, find Guests. Next, click on the link “Submit a Nomination,” enter your nominee’s information, and click “Submit.”

Prepared statements of one page or less in support of a nominee can also be submitted by e-mail to Karen Sellman, P&C Board Search Liaison, at ksellman@apa.org.

Deadline for accepting nominations is **January 20, 2006**, when reviews will begin.