

**Gaze communicates both cue direction and agent mental states**

Florence Mayrand

Department of Psychology

McGill University, Montreal

April 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree  
of Master of Science in Psychology - Thesis

©Florence Mayrand, 2022

## Table of Contents

|                                                                                |    |
|--------------------------------------------------------------------------------|----|
| Abstract.....                                                                  | 3  |
| Résumé.....                                                                    | 4  |
| Acknowledgements.....                                                          | 5  |
| Contribution to original knowledge.....                                        | 6  |
| Contribution of Authors.....                                                   | 7  |
| Manuscript “Gaze communicates both cue direction and agent mental states”..... | 8  |
| Title Page.....                                                                | 8  |
| Introduction.....                                                              | 9  |
| Methods.....                                                                   | 12 |
| Results.....                                                                   | 16 |
| Discussion.....                                                                | 20 |
| References.....                                                                | 23 |
| Appendices.....                                                                | 26 |
| Supplementary Materials.....                                                   | 26 |

### **Abstract**

Although it is well established that humans spontaneously attend to where others are looking, it remains debated whether this gaze following behavior occurs because gaze communicates directional information (i.e., where an agent is attending) or the agent's inferred mental content (i.e., what an agent perceives). To address this question, we used a novel task to measure how spatially dissociated and spatially combined effects of gaze cue direction and the inferred agent's mental content influence target performance. Our data revealed that performance was compromised when cue direction and mental content dissociated relative to when they combined. Performance for dissociated gaze direction and mental content was especially prominent when a social avatar served as a cue relative to a nonsocial arrow. Thus, these data show that human gaze communicates information about both where an agent is attending and what they are attending to.

## Résumé

Même s'il est reconnu que les humains orientent spontanément leur attention là où les autres regardent, un débat sur la nature de ce comportement demeure présent. Certains proposent que le regard des autres communique uniquement des informations directionnelles (où quelqu'un porte attention), lorsque d'autres proposent qu'il communique plutôt l'état mental d'autrui (qu'est-ce que quelqu'un regarde). Pour adresser ce débat, nous utilisons une nouvelle tâche pour mesurer les effets combinés et dissociés de la direction du regard et de l'état mental d'un agent lors d'une recherche de cible. Nos résultats démontrent que la performance était compromise lorsque la direction du regard et le contenu mental étaient dissociés relativement à lorsqu'ils étaient combinés. Ceci était particulièrement évident lorsqu'un avatar était utilisé comme signal comparé à une flèche. Par conséquent, ces résultats démontrent que le regard humain communique des informations sur où une personne regarde ainsi que sur quoi cette personne perçoit à cet endroit.

### **Acknowledgements**

I would like to thank the Natural Sciences and Engineering Research Council of Canada – CREATE funds for providing financial support towards the completion of this work. In addition, I would like to acknowledge the support from the Department of Psychology at McGill throughout the completion of my degree.

I would like to thank the Attention and Social Cognition Lab at McGill for allowing me to complete this work within its facilities, as well as all the members of the lab who contributed to the design of this study, data collection, and data analyses.

Finally, I would like to thank Jelena Ristic, Sarah McCrackin and Francesca Capozzi for their guidance and supervision during every step of this project.

### **Contribution to original knowledge**

Eye gaze is one of the most basic and powerful means of human social communication. Understanding the nature of gaze communicated signals would significantly elevate our understanding of the kinds of messages that are relayed by these social cues. While it is well known that humans communicate via gaze through behaviors such as gaze following, it remains debated if this behavior results from gaze communicating directionality, the gazer's mental content, or a combination of these factors. We provide an answer to this question and a resolution to the ongoing debate in the field by demonstrating that social gaze, but not similar directional non-social signals, communicates both directionality and agent's mental state. Thus, this finding brings a fundamentally novel perspective to our understanding of the nature of human gaze signals, and opens multiple new avenues for future research (e.g., the role of learning, development, social factors, or individual differences).

### **Contribution of Authors**

Florence Mayrand led all aspects of this project, including literature review, stimuli creation, experimental design, task programming, data collection, analyses, and manuscript write-up. Guidance and support were provided at each stage of this research project from Dr. Sarah McCrackin, Dr. Jelena Ristic and Dr. Francesca Capozzi. Final edits to the manuscript were done in collaboration between all authors.

## **Manuscript**

### **Gaze communicates both cue direction and agent mental states**

Florence Mayrand, Sarah D. McCrackin, Francesca Capozzi, Jelena Ristic

*Department of Psychology, McGill University, Montreal, QC, Canada.*

*Manuscript written in preparation for submission to a peer-reviewed journal.*



## Introduction

Visual information conveyed by gaze (eyes, head, or body deviation) enables quick communication of social messages (Capozzi & Ristic, 2018). As such, the ability to follow gaze has been implicated in both basic social functions like gaze following and joint attention (Frischen et al., 2007) as well as more complex social behaviors such as attitude formation (Toscano et al., 2018) and social status inference (Capozzi et al., 2019). Strikingly however, the large body of research on gaze following remains agnostic on the nature of messages conveyed by gaze (Capozzi & Ristic 2020; Frischen et al. 2007). Specifically, there remains a key outstanding question of whether humans follow gaze because it conveys directional information about *where* the items of a gazer's interest are located or because it conveys information about *what* the gazer perceives (Capozzi & Ristic, 2020). Using a novel procedure, here we dissociate the contributions of directional and mentalistic components of the gaze signal and show that gaze communicates *both* where the gazer is looking and what the gazer is perceiving.

The nature of signals communicated by gaze has been the subject of a longstanding debate (e.g., Capozzi & Ristic, 2020). The proponents of the directionality account argue that spontaneous following of gaze, often experimentally demonstrated by the gaze cuing (Friesen & Kingstone, 1998) or dot perspective tasks (Samson et al., 2010), reflects the cue's direction indicating where in space the gazer's attention is directed (Cole & Millett, 2019; Santiesteban et al., 2017). In contrast, the proponents of the mentalizing account maintain that this same behavior is instead driven by the spontaneous adoption of the gazer's visual perspective, which aids with inferring and sharing the representation of the gazed-at object with the gazer (Apperly & Butterfill, 2009). Supporting the directionality account, attention orienting has been found to occur similarly regardless of whether it is elicited by the direction of social gaze or a nonsocial cue like an arrow

(e.g.; Kingstone et al., 2019). Supporting the mentalizing account, gaze following is reduced when participants believe gaze cues are delivered by a randomized computer sequence (Wiese et al., 2012) or a person whose line of sight is obstructed (Baker et al., 2016).

While this debate has traditionally been framed within the context of the gaze signal relaying either simple directional information or more complex mental perspective, it is important to note that computing gaze directionality, and the resulting effects this computation has on behavior, can also be considered a form of visual perspective taking. When contextualized in this manner, the two sides of the directionality vs. the mentalizing debate appear to align with available definitions of the two different perspective taking complexity levels. Specifically, the directional account appears to align with the processes associated with simpler Level-1 visual perspective taking, in which one understands what lies within someone else's line of sight (Kessler & Rutherford, 2010). Comparatively, the mentalizing account appears to align well with the processes associated with more complex Level-2 visual perspective taking, in which one understands the visual aspects of a scene relative to an imagined viewpoint of another agent (Flavell et al., 1981). To maintain connection with the existing research and the ongoing debate (e.g. Capozzi & Ristic, 2020), here we refer to the two processes implicated in gaze signal as the directional component and the mentalistic component, whereby the directional component reflects the influence of gaze cue direction while the mentalizing component refers to understanding of the gazer's mental perspective.

To investigate the contribution of directional and mentalizing processes in gaze communication, we designed a novel task in which we measured combined and dissociated effects of these two processes on target performance. Participants were presented with a central cue (a human avatar or an arrow) and were asked to localize a peripheral target which was flanked by a

response target or a non-target distractor (Figure 1). Providing directional information to the observer, the avatar randomly looks to the left or to right, that is, either towards the response target or towards a non-target distractor. Providing mentalistic information, from its perspective, the avatar perceives either the observer's response target or the non-target distractor. Critically, in the combined conditions, the directional and mentalistic processes are congruent, like they would be in a typical gaze signal, such that the avatar both gazes at and mentally perceives the target from its perspective. Conversely, in the dissociated conditions the directional and mentalizing processes are incongruent such that the avatar may gaze at the target but perceive a distractor and vice versa.

If gaze communicates both cue direction and inferred mental content of the gazer, target performance should suffer the most when these two components spatially dissociate relative to when they combine. That is, a condition in which the avatar is looking at a peripheral location containing the participant's target but perceives a distractor from its own mental perspective should result in slower target responses relative to a condition in which the avatar is looking at a peripheral location containing the participant's target and perceives the same target from its own mental perspective. This is because in the former the directional and mentalistic information communicated by the avatar are dissociated while in the latter the two pieces of information are combined. Hence, when directional and mentalistic processes occur together in a congruent manner (i.e., the avatar is both looking at and perceiving the target from its visual perspective) performance is facilitated. To understand whether any these effects may be unique to directional cues delivered by social agents, we also subjected a nonsocial central arrow cues to the same test.

## Methods

The study was pre-registered ([osf.io/fwptb](https://osf.io/fwptb)). Deidentified data are available at [osf.io/3xcqk](https://osf.io/3xcqk).

### Participants

An a priori power analysis using an estimated moderate size estimate of  $r = .25$  for the variability in the magnitude of the gaze cuing with mental state attribution (approximated from Moriguchi et al., 2006; Sulpizio et al., 2015; Tomei et al., 2017) indicated that data from about 95 participants would yield power of .8 and data from 130 participants would yield power of .9 (Alpha=.05). Data from 136 participants (114 women, 22 men; mean age = 20.51 years, SD = 1.34) were analyzed<sup>1,2</sup>. All procedures were approved by the University's Research Ethics Board.

### Apparatus & Stimuli

Figure 1 illustrates the stimuli, sample trial sequence, and the four test conditions. Cue stimuli were images of a human avatar and an arrow, which were equated for length (Figure 1A). The cues were positioned at fixation. Target and distractor stimuli, which were shown on the left and right of fixation respectively, were a letter **E**, a number **3**, and a number **8** (Figure 1C). They were equated for size. Each target, either **3**, **E**, or **8**, was always flanked by a unique distractor, creating three unique target-distractor combination (**3**/**E** ; **3**/**8** ; **E**/**8**).

The study was administered online via Testable (<https://www.testable.org/>). The experiment launched on participants' personal computers. The entire stimulus response display

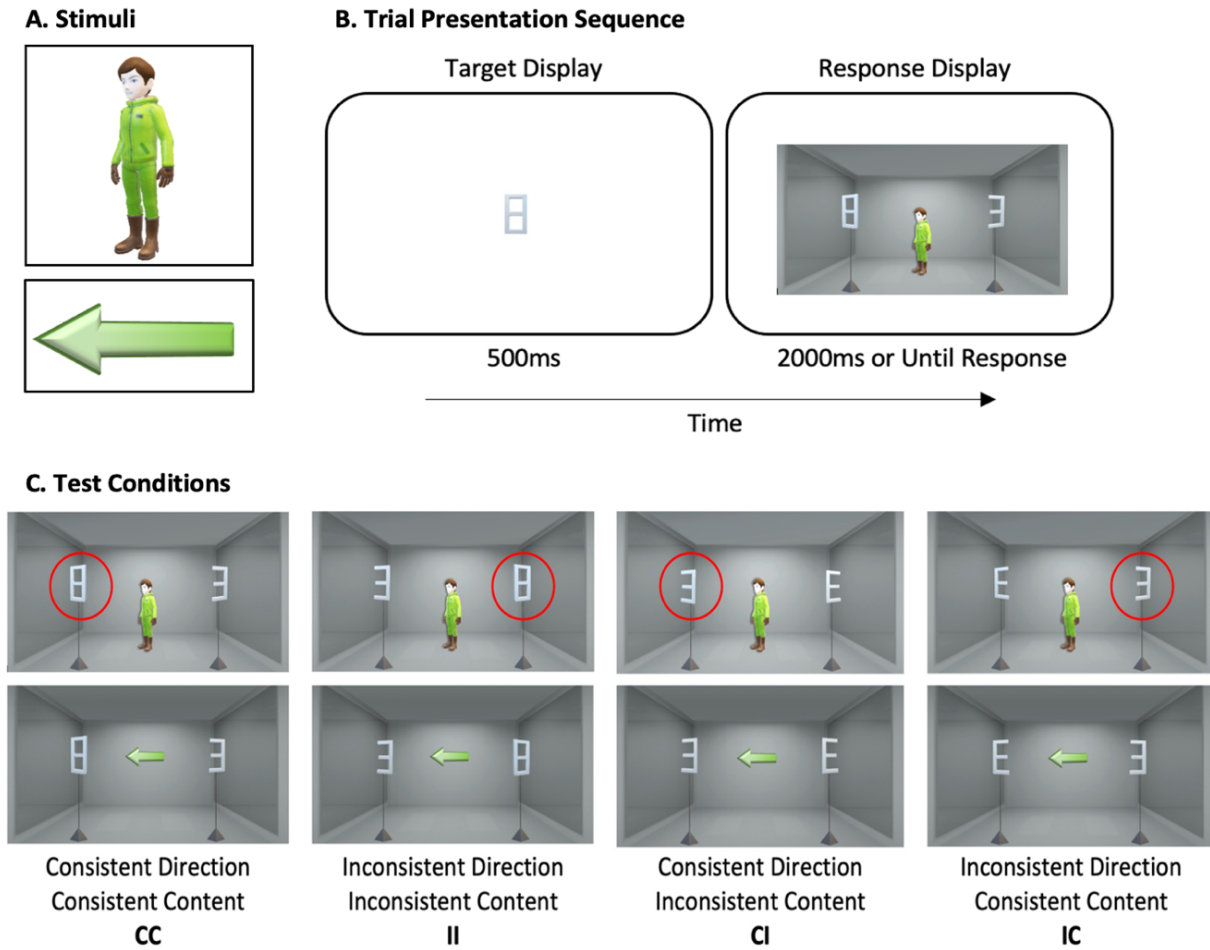
---

<sup>1</sup> A total of 162 undergraduate students completed the study. Data from 26 participants were removed based on the pre-registered exclusion criteria of having less than 80% overall accuracy (i.e., over 20% of trials for each participant lost due to response anticipations, timeouts, or incorrect responses; see e.g., McCrackin and Itier, 2019; Van Selst and Jolicoeur, 1994). All participants reported native English fluency, no history of psychiatric or neurological conditions, and normal or corrected-to-normal vision<sup>1</sup>. Informed consent was obtained from all participants.

<sup>2</sup> As per the registration plan, participants also completed the 28-item Interpersonal Reactivity Index (IRI; Davis, 1983) where, for each question (e.g., "I try to look at everybody's side of a disagreement before I make a decision"), they responded using a 5-point Likert scale ranging from 1 - "Does not describe me well" to 5 - "Describes me very well". Exploratory correlation analyses linking individual RT performance in each of the four test conditions (CC, II, CI, IC) with the overall IRI score for yielded no reliable variability ( $-.004 < r < -.088$ ,  $p > .308$ ).

image including all stimuli in proportions was scaled to approximately 50% of individual participants' screens. The display target images were rendered in grayscale; the central cues were rendered in green.

Figure 1.



**Figure 1.** Stimuli, example stimulus presentation sequence, and test conditions. **1A.** Illustration of cue and target stimuli. **1B.** Example trial sequence. Participants were first presented with an image of the trial response target for 500ms. The response display was then presented for 2000ms or until participants responded. **1C.** Illustration of the four test conditions for each cue type. Target is highlighted for the reader with a red circle. Stimuli are not drawn to scale.

## Design

The study was a repeated measures design, with *Cue Type* (Avatar; Arrow), *Cue Direction Consistency* (Consistent; Inconsistent), and *Target Content Consistency* (Consistent; Inconsistent) included as variables. Cue type was blocked while Cue direction consistency and Target content consistency were intermixed. Each target-distractor combination appeared equally often. Target identity was intermixed and presented equiprobably across trials.

*Cue Type* manipulated the type of the central cue. It varied between an avatar and an arrow. Half of test blocks (i.e., 4) presented an avatar cue and the other half presented an arrow cue. The order of block presentation was randomized across participants.

*Cue Direction Consistency* manipulated whether the response target appeared at the spatial location directionally indicated by the central cue (i.e., Consistent; Figure 1C) or at the opposite location (i.e., Inconsistent). This variable is a composite of cue direction and target location factors, each of which varied independently and equiprobably between left and right spatial positions.

*Target Content Consistency* manipulated the observed and inferred avatar's mental content. The avatar's mental content (i.e., what they would perceive from their vantage point) was either consistent or inconsistent with the participant's perception of the target. Consistent target content between the observer and the avatar was induced using stimuli that invoke the same percept from the observer's and the avatar's perspectives (i.e., number 5; Figure 1C). That is, the response target 5 would appear the same from both the observer's and the avatar's perspective. Inconsistent target content between the observer and the avatar was induced using target stimuli that are mentally perceived as a target from the observer's perspective but a distractor from avatar's perspective and vice versa. That is, for the avatar, the observer's response target 3 would appear

as a distractor letter **E** while the response target **E** would appear as a distractor number **3**. Target identity and target location varied independently and equiprobably with cue direction.

Thus, in the *Consistent direction/Consistent content (CC)* and the *Inconsistent direction/Inconsistent content (II)* conditions, cue direction and target mental content combined, such that both were either consistent (CC) or inconsistent (II). That is, the cue either pointed at or away from the target, which was perceived the same from the observer's and the avatar's perspective. Conversely, in the *Consistent direction/Inconsistent content (CI)* and the *Inconsistent direction/Consistent content (IC)* conditions, cue direction and target mental content dissociated. In the CI condition, the avatar indicated the response target by its gaze but perceived a distractor (i.e., the observer's response target **3** appears as a distractor **E**). In contrast, in the IC condition, the avatar pointed to a distractor but perceived the target at that location, since the non-target distractor may be perceived as the target letter from the avatar's perspective (i.e., distractor **3** appears as the response target **E**).

## Procedure

Sample trial sequence is shown in Figure 1B. Trials began with an image of a response target (**3**, **E**, **B**) for 500ms to inform participants about the response target for that trial. Then, the response display showing the central cue and the target/distractor combination was presented. The display remained visible for 2000ms or until participants responded. Participants were informed that the gaze cue was not predictive of target location. They were asked to localize the target quickly and accurately by pressing the 'b' and 'h' keys on the keyboard. Target location (left, right) – key response ('b' or 'h') pairing was counterbalanced between participants.

The experiment consisted of 512 trials divided evenly across eight testing blocks. Each block contained 16 repetitions of each of the four test conditions. All variables were randomized

and presented in an equiprobable manner. Sixteen practice trials with performance feedback were run first.

## Results

Response anticipations ( $RT < 200\text{ms}$ ) and timeouts ( $RT > 1800\text{s}$ ) accounted for 1.99% of trials and were removed from analyses. When Mauchly's test was significant, the Greenhouse-Geisser degrees of freedom are reported. All follow-up paired, two-tailed t-tests were Bonferroni corrected. Data were analyzed using SPSS 27.

Overall, the task was well done with 93.03% response accuracy. Mean accuracy was examined for each *Cue Type* (Avatar, Arrow) and *Condition* (CC, II, CI, IC) using a repeated measures ANOVA. The analyses indicated no speed-accuracy trade-off, as CC ( $M=.97$ ) and II ( $M=.96$ ) conditions were overall responded to more accurately (and faster, as described in Results) than CI ( $M=.90$ ) and IC ( $M=.90$ ) conditions (*Condition* main effect,  $F(2.39, 323.23)=208.51$ ,  $p<.0001$ ,  $MSE=.002$ ,  $\eta p^2=.61$ ) for both cue types. Main effect of *Cue Type* ( $F<1$ ) and the *Cue Type x Condition* interaction,  $F(2.75, 370.99)=2.63$ ,  $p=0.055$ ,  $MSE=.001$ ,  $\eta p^2=.019$ , were not reliable.

To remind, we reasoned that if gaze communicates both directional and mentalistic content, a detriment in target performance should emerge when these two components of a gaze signal are dissociated in the CI and IC conditions. We also predicted that this performance detriment would be more pronounced when the social avatar vs. a nonsocial arrow served as a cue.

A repeated measures ANOVA with *Cue Type* (Avatar; Arrow), *Cue Direction Consistency* (Consistent; Inconsistent), and *Target Content Consistency* (Consistent; Inconsistent) was used to



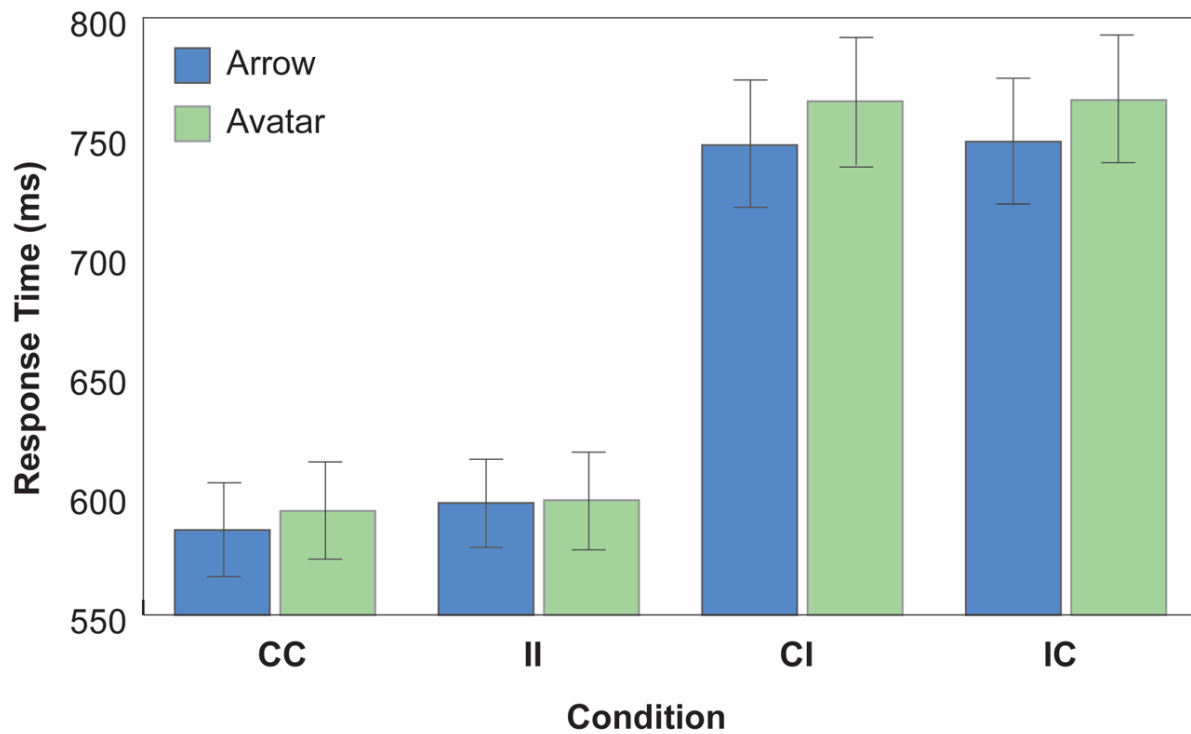
examine mean correct RTs. These means are plotted in Figure 2 as a function of Cue type, Cue Direction consistency, and Target Content consistency conditions.

The results supported our predictions by returning a reliable two-way interaction between *Cue Direction* and *Target Content Consistency* ( $F(1, 135) = 681.68, p < .001, MSE = 10580.12, \eta^2 = 0.835$ ) as well as a reliable three-way interaction between *Cue Type*, *Cue Direction Consistency* and *Target Content Consistency* ( $F(1, 135) = 12.914, p < 0.001, MSE = 12111.288, \eta^2 = 0.087$ ). The first interaction indicated that, overall, targets cued by consistent cue direction were responded to faster when that cue direction was paired with consistent target content relative to when it was paired with inconsistent target content (CC vs. CI; Avatar:  $t(135) = -24.92, p < 0.001$ ; Arrow:  $t(135) = -22.670, p < 0.001$ ). Likewise, uncued targets (i.e., inconsistent direction trials) were responded to faster when paired with inconsistent target content relative to when paired with consistent target content (II vs. IC; Avatar:  $t(135) = -23.84, p < 0.001$ ; Arrow:  $t(135) = -22.25, p < 0.001$ ).

This was further supported by a significant three-way interaction between *Cue Type*, *Cue Direction Consistency*, and *Target Content Consistency* ( $F(1, 135) = 12.914, p < 0.001, MSE = 12111.288, \eta^2 = 0.087$ ). This interaction indicated that there was a larger difference during the dissociated CI and IC conditions, in which the cue directionality and target mental content are dissociated (CI:  $t(135) = -3.126, p = 0.002$ ; IC:  $t(135) = -3.088, p = 0.002$ ), when the social avatar served as a cue relative to when a nonsocial arrow served as a cue. In other words, responses to targets following gaze cues were slower than responses to targets following arrow cues in conditions in which directionality and mental content are dissociated. Thus, the dissociation of the two components induced a larger performance detriment during avatar trials than during arrow trials. In contrast, there was no difference in response to the targets across the two cue types in the

CC and II conditions, where cue directionality and target mental content combined (CC:  $t(135)=-1.756$ ,  $p=0.081$ ; II:  $t(135)=-0.298$ ,  $p=0.766$ ; *Cue Type x Cue Direction Consistency*,  $F(1,135)=1.501$ ,  $p=.223$ ). This replicates a well-known finding in the literature showing that social and nonsocial cues typically induce similar cuing effects (Ristic et al., 2002). The interaction between *Cue Type* and *Target Content Consistency* was not reliable ( $F<1$ ).

**Figure 2. Results**



**Figure 2.** Mean correct Response Time (RT) as a function of Cue type, Cue Direction consistency and Target Content consistency conditions (CC= Consistent direction/Consistent content; II=Inconsistent direction/Inconsistent content; CI=Consistent direction/Inconsistent content; IC=Inconsistent direction/Consistent content). Error bars depict the 95% CI.

The ANOVA additionally returned main effects of *Cue Type* ( $F(1, 135) = 7.59$ ,  $p= .007$ ,  $MSE = 4531.75$ ,  $\eta p^2 = .053$ ) and *Cue Direction Consistency* ( $F(1, 135) = 7.951$ ,  $p = 0.006$ ,  $MSE = 5366.289$ ,  $\eta p^2 = 0.056$ ), with overall faster responses to targets cued by the arrow's direction than

targets cued by the avatar's gaze, and overall faster responses in the CC condition than the II condition across both cue types. A main effect of *Target Content Consistency* approached significance,  $F(1, 135) = 3.815, p = .053, MSE = 858.22, \eta p^2 = .027$ , with trials on which the target content was consistent with the observer's mental representation generating faster responses than trials on which the target content representation diverged across the observer and the cue.

Thus, to summarize, in line with our predictions, we found a deterrent in target performance when gaze information about cue directionality and target mental content dissociated (IC and CI conditions) relative to when they combined (CC and II conditions). Further, this deterrent was found to be larger when a social avatar served as a cue relative to when a nonsocial arrow served as a cue. As such, this finding supports the notion that gaze conveys information both about *where* the gazer is directing their attention and *what* they see from their visual perspective in a joint manner. When these processes are combined, typical effects of facilitated target performance are found<sup>3</sup>. When they are dissociated, a large slowdown in performance is observed.

One might wonder whether the reported deterrent in performance for the dissociated conditions may reflect differences in targets, as the targets in the dissociated conditions afford multiple representations (i.e., **3** and **E**) while the targets in the combined condition (i.e., **B**) afford a single representation. There are two converging pieces of evidence that argue against this interpretation. First, our analysis is indicating a critical interaction across social vs. nonsocial cue type indicated a larger deterrent in performance for dissociated conditions when an avatar cue that engages social mental processes served as a cue relative to when a nonsocial arrow, which presumably engages those mentalizing processes to the lesser extent, was used. Importantly, this

---

<sup>3</sup> The same effects of dissociated cue direction and target mental content were observed in a separate additional experiment in which cue type was manipulated between participants. This experiment is reported in Supplementary Materials section.

dissociation emerged when the same stimulus served as the target. Thus, when social gaze cues are used, performance is significantly more slowed down than by the dissociation between cue direction and gazer's mental content relative to when nonsocial arrows are used irrespectively from target properties. Second, existing research on the processes relating to the line of sight computation and mental representations (Zacks et al., 2000; Voyer et al., 2017; Michelon & Zacks, 2006) shows that target location tasks, such as the one we have used here, invoke egocentric reference of frames computation in observers. This means that target responses reflect the computation of mental perspective from the gazer's perspective rather than alternative processes like mental rotation for instance. In other words, when responding to targets in tasks such as this one, observers adopt the other's agent mental perspective to respond to targets rather than mentally rotating the target from their own point of view.

### **Discussion**

Grasping the nature of messages conveyed by human gaze is important for understanding the means and the underlying mechanisms of our nonverbal social communication. Using a novel paradigm which measures both combined and dissociated contributions of cue directionality and target mental content, here we showed that when these two aspects of the gaze signal are dissociated, significant slowing in target performance is found. This stood in contrast to the conditions in which cue direction and avatar's mental content were congruent or combined, in which performance was reliably facilitated by such combined representation. As such, these results provide one of the first experimental demonstrations of the influence of both cue directionality *and* mental content in gaze communication, and strongly suggest that the gaze signal typically communicates information about both cue direction and the gazer's visual perspective, i.e., both

*where* agents are attending and *what* they perceive. In other words, understanding of the line of sight depends both on the computation of the direction of the cue and inferred mental perspective of the gazer. We next bring up and discuss two potential implications of this finding.

First, the typical gaze following response appears to reflect the computation of both the gaze cue's direction, i.e., the line of sight *and* the representation of gazer's mental content. This conclusion follows from our finding showing impaired performance in conditions in which the cue's directionality and target mental content were dissociated. That is, target performance was slowed in a similar manner by both the absence of the cue's directional content *and* the absence of the mental representation of the target. Thus, in typical gaze following target-related response, both cue's directional information and agents' mental state appear to be relayed in a joint fashion. When these two processes are dissociated (i.e., in CI and IC conditions), the gaze following response is slowed. This result provides empirical resolution to the longstanding debate in the field of whether gaze conveys directional information or mental content, showing the involvement of both the domain general attentional processes linked to processing of cue directionality and domain specific processes linked to computations of social mental content of the gazer (Capozzi & Ristic, 2020).

Dovetailing with this point, our results also showed that dissociating the cue direction from mental content impacted target performance the most when the social avatar relative to a nonsocial arrow served as a cue. Thus, it seems likely that while the components of cue direction and mental content may be tightly combined in social signals, they are less intertwined and thus more easily dissociated in nonsocial ones (Marotta et al., 2012). For example, while a simple "line of sight" computation may be easily attributed to a directional cue such as an arrow, this nonsocial stimulus does not possess a mental or visual perspective. That said, while it may be counterintuitive to

attribute ‘mental state’ representation to a nonsocial cue, a joint representation of cue directionality and its intentional meaning may still occur in this condition but to a lesser extent given that behaviorally relevant cues like arrows may convey highly meaningful messages in daily life (e.g., Santiesteban et al., 2014). The difference between the magnitudes of the dissociated effects between social and nonsocial cues thus may reflect the increased difficulty in disjoining the combined representation of directionality and meaning for social vs. overlearned nonsocial cues (e.g., Ristic & Kingstone, 2012). Future studies are needed to examine how dissociating cue direction from its mental content representation may be affected by the differences in the cues’ social and learned values.

In sum, using a novel paradigm, here we showed experimentally that that gaze signal conveys both information about *where* an agent is attending and *what* they perceive. This finding provides a new perspective on social signaling by gaze indicating its complex signal content. As such, this finding also opens fruitful new avenues for research on the properties of social gaze communication such as relative contributions of directionality vs. mental content, developmental trajectory and expressions in special groups, variations with individual differences, and/or the underlying neural mechanisms.

## References

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116(4), 953.
- Baker, L. J., Levin, D. T., & Saylor, M. M. (2016). The extent of default visual perspective taking in complex layouts. *Journal of Experimental Psychology: Human Perception and Performance*, 42(4), 508.
- Capozzi, F., Beyan, C., Pierro, A., Koul, A., Murino, V., Livi, S., et al. (2019). Tracking the Leader: Gaze Behavior in Group Interactions. *IScience*, 16, 242-249.
- Capozzi, F., & Ristic, J. (2018). How attention gates social interactions. *Ann. NY Acad. Sci*, 1426, 179-198.
- Capozzi, F., & Ristic, J. (2020). Attention AND mentalizing? Reframing a debate on social orienting of attention. *Visual Cognition*, 28(2), 97-105.
- Cole, G. G., & Millett, A. C. (2019). The closing of the theory of mind: A critique of perspective taking. *Psychonomic Bulletin & Review*, 26(6), 1787-1802.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1), 113
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490-495.
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694-724.
- Kingstone, A., Kachkovski, G., Vasilyev, D., Kuk, M., & Welsh, T. N. (2019). Mental attribution is not sufficient or necessary to trigger attentional orienting to gaze. *Cognition*, 189, 35-40.

- Marotta, A., Lupiáñez, J., Martella, D., & Casagrande, M. (2012). Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection. *Journal of Experimental Psychology: Human Perception and Performance*, 38(2), 326.
- McCrackin, S. D., & Itier, R. J. (2019). Individual differences in the emotional modulation of gaze-cuing. *Cognition and Emotion*, 33(4), 768-800.
- Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., ... , et al. (2006). Impaired self-awareness and theory of mind: an fMRI study of mentalizing in alexithymia. *Neuroimage*, 32(3), 1472-1482.
- Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, 9(3), 507–513.
- Ristic, J., & Kingstone, A. (2012). Automated symbolic orienting: A unique form of human spatial attention. *Visual Cognition*, 20(3), 244-264.
- Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1255.
- Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, 40(3), 929.
- Santiesteban, I., Kaur, S., Bird, G., & Catmur, C. (2017). Attentional processes, not implicit mentalizing, mediate performance in a perspective-taking task: Evidence from stimulation of the temporoparietal junction. *NeuroImage*, 155, 305-311.
- Sulpizio, V., Committeri, G., Metta, E., Lambrey, S., Berthoz, A., & Galati, G. (2015). Visuospatial transformations and personality: evidence of a relationship between



- visuospatial perspective taking and self-reported emotional empathy. *Experimental Brain Research*, 233(7), 2091-2102.
- Tomei, A., Besson, J., & Grivel, J. (2017). Linking empathy to visuospatial perspective-taking in gambling addiction. *Psychiatry Research*, 250, 177-184.
- Toscano, R., Price, G., & Scheepers, C. (2018). The impact of CEO arrogance on top management team attitudes. *European Business Review*.
- Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology Section A*, 47(3), 631–650.
- Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PloS One*, 7(9), e45391.

## Appendices

### *Supplementary Materials*

#### *Between-Subjects Study*

#### *Participants, Apparatus, Stimuli, Design & Procedure*

The data from an additional 220 participants were analyzed<sup>4</sup>, with 111 participants randomly assigned to view the avatar cue (97 women, 14 men, Mean age = 20.53 years,  $SD = 2.34$ ) and 109 participants (92 women, 15 men, 2 other, Mean age = 20.67 years,  $SD = 2.82$ ) assigned to view the arrow cue.

Stimuli, apparatus, design, and procedure were identical as in the main study except that participants always viewed the same cue, either an avatar or an arrow. Four blocks of 80 trials were presented, for a total of 320 trials. Each block contained 20 trials for each of the 4 test conditions (CC, II, CI, IC). All conditions were intermixed and presented equally often using a random sequence.

#### *Results*

Response anticipations ( $RT < 200ms$ ) and timeouts ( $RT > 1800s$ ) accounted for 1.49% of all trials and were removed from analyses. Performance was high overall at 92.22%. Accuracy was examined across *Cue Type* (Avatar, Arrow; between-subjects) and the four test *Conditions* (CC, II, CI, IC; within-subjects) using a mixed-effects ANOVA. No speed-accuracy trade-off was evident in the data, with CC ( $M = .96$ ) and II ( $M = .96$ ) conditions returning higher overall accuracy than CI ( $M = .89$ ) and IC ( $M = .88$ ) conditions,  $F(2.30, 502.11) = 303.67, p < .001, MSE = .002$ ,

---

<sup>4</sup> A total of 275 participants completed the study. Participants achieving less than 80% overall accuracy (i.e., over 20% of trials lost due to anticipations, timeouts, or incorrect responses) were excluded from analyses ( $n=55$ ). As in the main study, participants also completed the 28-item Interpersonal Reactivity Index (IRI). The same correlation analyses were performed (Avatar:  $.111 < r < .189, .246 < p < .047$ ; Arrow:  $-.023 < r < -.059, .545 < p < .808$ ) and suggested that there was a positive relationship between the IRI score and RTs in the II ( $r = .189; p = .047$ ) and CC ( $r = .186; p = .051$ ) conditions in the Avatar group only.

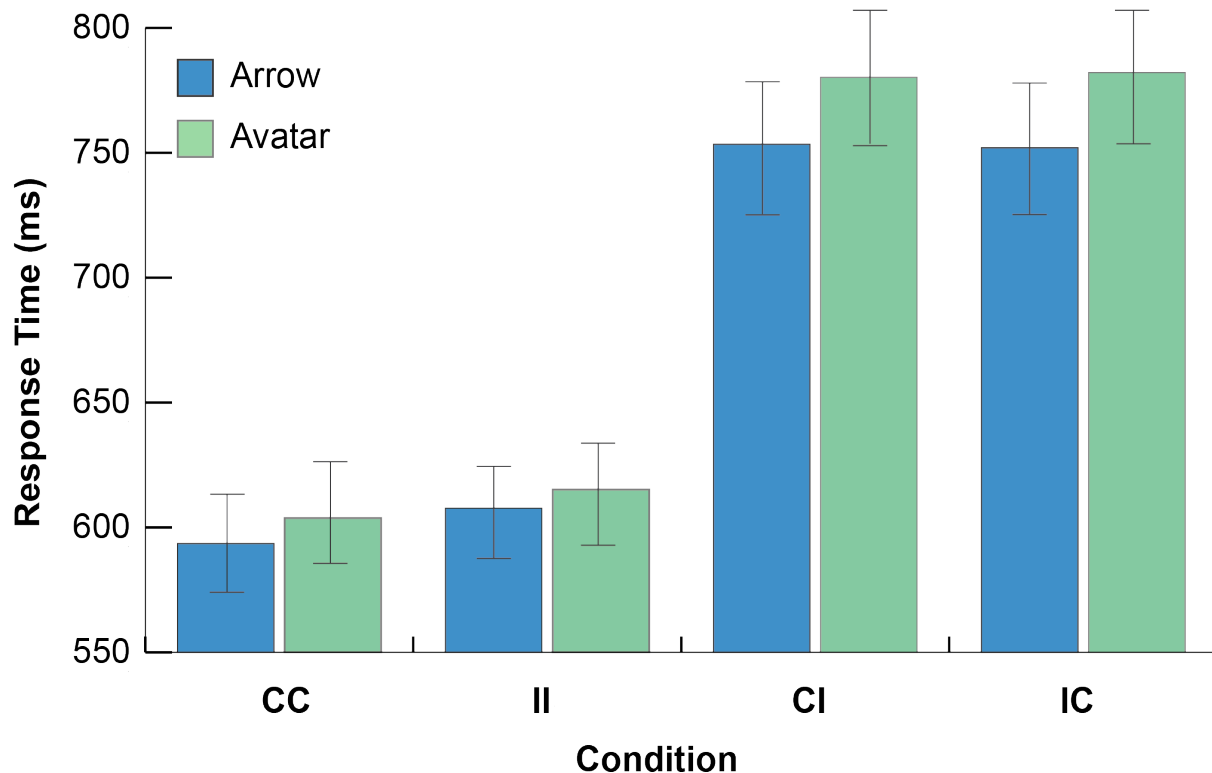
$\eta^2=.58$ ). *Cue Type x Condition* interaction was not reliable,  $F(3, 654) = .19, p=.90, MSE = .002, \eta^2=.001$ .

Response time data were analyzed using a mixed-effects ANOVA with *Cue Type* (Avatar; Arrow) included as a between-subjects variable, and *Cue Direction Consistency* (Consistent; Inconsistent) and *Target Content Consistency* (Consistent; Inconsistent) included as a within-subjects variables. Supplementary Figure 1 shows mean RTs as a function of Cue type, Cue Direction consistency and Target Content consistency conditions.

As in the main experiment, a *Cue Direction Consistency x Target Content Consistency* interaction,  $F(1, 218) = 1214.10, p < 0.001, MSE = 4748.76, \eta^2 = 0.85$ , indicated that consistent direction trials were responded to faster when paired with consistent compared to inconsistent target content (CC vs CI for arrow:  $t(108) = -24.26, p<.001$ ; for avatar:  $t(110) = -23.73, p<.001$ ) and inconsistent direction trials were responded to faster when paired with inconsistent compared to consistent target content (IC vs II for arrow:  $t(108) = -21.80, p<0.001$ ; for avatar:  $t(110) = -22.77, p<0.001$ ).

The critical three-way interaction between *Cue Type*, *Cue Direction Consistency*, and *Target Content Consistency* was significant as well,  $F(1, 218) = 4.48, p = .035, MSE = 4748.76, \eta^2 = 0.02$ , indicating slower performance for the avatar group relative to the arrow group for the CI (780ms avatar vs. 753ms arrow) and IC ( $M = 782$ ms avatar vs. 752ms arrow) conditions, while there was less of a cue type difference for the CC (604ms Avatar vs. 594ms Arrow) and II (615ms Avatar vs. 608ms Arrow) conditions. Follow-up independent group t-tests comparing the mean RTs for the Avatar and Arrow cue CI ( $t(218) = 1.48, p = 0.14$ ) and IC ( $t(218) = 1.64, p = 0.10$ ) conditions did not reach significance.

All main effects were reliable. A main effect of *Cue Type* ( $1, 218$ ) = 7789,  $p < .0001$ ,  $MSE = 53165$ ,  $\eta^2 = .973$ , indicated overall faster responses to targets cued by the arrow's direction than targets cued by the avatar's gaze. A main effect *Cue Direction Consistency* indicated that overall consistent CC trials were faster than II trials,  $F(1, 218) = 11.87$ ,  $p < .001$ ,  $MSE = 788.22$ ,  $\eta^2 = .052$ , replicating the attentional cuing effects. Finally, a main effect of *Target Content Consistency* indicated that overall Consistent Content generated faster response times than Inconsistent Content,  $F(1, 218) = 13.10$ ,  $p < .001$ ,  $MSE = 660.19$ ,  $\eta^2 = .057$ . No other effects or interactions were reliable (*Cue Type*  $\times$  *Cue Direction Consistency*; *Cue Type*  $\times$  *Target Content Consistency*, both  $F_s < 1$ ).



**Supplementary Figure 1.** Mean Response Time as a function of Cue Type, Cue Direction consistency, and Target Content consistency conditions (CC = Consistent Cue Direction/Consistent Target Content; II = Inconsistent Cue Direction/Inconsistent Target Content; CI = Consistent Cue Direction/Inconsistent Target Content; IC = Inconsistent Cue Direction/Consistent Target Content). Error bars depict the 95% CI.