Cognition 180 (2018) 1–9

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Mental states modulate gaze following, but not automatically

Gustav Kuhn[a], Ieva Vacaityte[b], Antonia D.C. D'Souza[c], Abbie C. Millett[c], Geoff G. Cole[c,*]

[a] Department of Psychology, Goldsmiths, University of London, UK
[b] Department of Psychology, Brunel University, UK
[c] Centre for Brain Science, University of Essex, UK

ARTICLE INFO

Keywords:
Gaze following
Mental states
Social attention
Automaticity
Eye movements

ABSTRACT

A number of authors have suggested that the computation of another person's visual perspective occurs automatically. In the current work we examined whether perspective-taking is indeed automatic or more likely to be due to mechanisms associated with conscious control. Participants viewed everyday scenes in which a single human model looked towards a target object. Importantly, the model's view of the object was either visible or occluded by a physical barrier (e.g., briefcase). Results showed that when observers were given five seconds to freely view the scenes, eye movements were faster to fixate the object when the model could see it compared to when it was occluded. By contrast, when observers were required to rapidly discriminate a target superimposed upon the same object no such visibility effect occurred. We also employed the barrier procedure together with the most recent method (i.e., the ambiguous number paradigm) to have been employed in assessing the perspective-taking theory. Results showed that the model's gaze facilitated responses even when this agent could not see the critical stimuli. We argue that although humans do take into account the perspective of other people this does not occur automatically.

1. Introduction

We often gaze towards locations that are looked at by others, and this form of social attention is an essential part of human interaction and cognition in general. At the centre of this orienting mechanism is the need to know what others are looking at, a process that involves the computation of another person's mental state, i.e., Theory of Mind (ToM). Although early work did consider ToM mechanisms in this so-called gaze following, later social attention workers tended to conceive gaze-induced attentional behaviour as a bottom-up process, rather than involving higher mechanisms (Driver et al., 1999; Friesen & Kingstone, 1998; Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002). More recently, a number of authors have explicitly suggested that gaze following is influenced by what the gazer can see and that ToM forms an essential component of gaze cueing (e.g., Teufel, Fletcher, & Davis, 2010). Some authors have even argued that the computation of what others see occurs spontaneously (Samson, Apperly, Braithwaite, Andrews, & Scott, 2010).

Mechanisms associated with gaze following are typically investigated using some variant of a paradigm in which participants are asked to respond to targets that either appear in locations looked at by another agent (i.e., 'valid' trials), or they appear elsewhere (i.e., 'invalid' trials; Driver et al., 1999; Friesen & Kingstone, 1998). Response times (RTs) are generally shorter on valid compared with invalid trials, an effect that has been observed for both manual responses (Frischen, Bayliss, & Tipper, 2007) as well as saccadic eye movements (Kuhn & Benson, 2007; Ricciardelli et al., 2002). To examine whether ToM processes modulate gaze following, Teufel, Alexis, Clayton, and Davis (2010) used a modified version of this task in which the gazing agent wore mirrored goggles. Participants were informed that the goggles were either transparent or opaque, thereby manipulating whether the agent could see the targets or not. Results revealed a larger gaze cueing effect when participants were informed that the agent could see, thus supporting the view that gaze following is modulated by mental state attribution (see also Teufel et al., 2009).

Although the above results have been taken as evidence that gaze cueing *can* be modulated by mental states, some authors have argued that humans spontaneously compute the perspective of others. This view has come from results obtained in the 'dot perspective' paradigm in which participants are presented with an image of a room that contains an avatar who looks either towards the left or the right hand wall (e.g., Samson et al., 2010; Santiesteban, Catmur, Hopkins, Bird, & Heyes, 2014). A number of dots are pinned to either the left, the right, or both walls, and participants are asked to make judgments about the number of dots that are either visible to them or visible to the avatar. The most interesting finding from this procedure is that when making

* Corresponding author at: Centre for Brain Science, University of Essex, Wivenhoe Park, Colchester CO4 3 SQ, UK.
E-mail address: ggcole@essex.ac.uk (G.G. Cole).

https://doi.org/10.1016/j.cognition.2018.05.020
Received 3 November 2016; Received in revised form 25 May 2018; Accepted 26 May 2018

own-perspective judgements, participants make slower responses if the number of dots seen by the avatar does not match that seen by the participant; so-called altercentric intrusion. Samson et al. argued that this occurs because the avatar's perspective is computed, leading to the interference, and that this process is 'spontaneous'. In a later article, Surtees and Apperly (2012) stated that the process is 'automatic'. Samson et al. (2010) also suggested that the processes involved in this effect are similar to those involved in generating gaze following in the gaze cueing effect (described above).

A central challenge to the mental state theory of these gaze-induced effects has come from a series of experiments by Cole and colleagues (Cole, Atkinson, Le, & Smith, 2016; Cole, Smith, & Atkinson, 2015; Cole, Atkinson, D'Souza, & Smith, 2017; see also Langton, 2018) who adopted a procedure often employed in animal and infant ToM research (e.g., Moll & Tomasello, 2004). A physical barrier placed in the line of sight between the gazing agent and the target renders the target non-visible to the gazer. For instance, Hare, Call, and Tomasello (2001) showed that a subordinate chimpanzee knows whether a dominant chimpanzee can see a food item based on whether the latter's view of the food is obscured by a barrier or not. Since a gazer cannot see a target under a barrier condition, Cole et al. reasoned that any gaze cueing-like effect observed when a target is not visible to the gazer cannot be due to the gazer's visual perspective driving the gaze cueing effect. In a series of experiments, including one in which a physically present person acted as the cue, the gazing agent induced strong cueing effects. Importantly, Cole et al. found that these effects were not influenced by whether the target was visible or not, thus challenging the notion that gaze effects are modulated by mental states. In a follow-up study, Cole et al. (2016) found the same pattern of results when the barrier method was employed in the dot perspective task. That is, automatic perspective-taking-like data were observed when the avatar could not see any dots due to the location of a barrier.

The findings of Cole et al. are however in contrast to other recent work that has also employed the barrier method. Using a variant of the dot perspective task, Baker, Levin, and Saylor (2016) reported that dot judgements *were* found to be influenced by whether the avatar could see the targets or not (see also Morgan, Freeth, & Smith, 2018). Furthermore, the challenge to the theory that ToM influences gaze following does not concur with the common observation that we, as social beings, often find ourselves following another person's gaze precisely because of an explicit mental state attribution. Many of us have often said to ourselves "I wonder what she is looking at" before trying to determine what the viewer is observing. This is a clear example of visual perspective modulating gaze following. The real issue may therefore be whether the process in which ToM modulates gaze following can occur automatically.

The primary aim of the present work was to examine the question of whether humans do indeed compute the perspective of other individuals. In four experiments we adopted the visibility manipulation described above in which a gazing agent either sees the target stimuli or does not. Furthermore, our experiments were particularly concerned with the theory that not only does such perspective computation occur but that it does so automatically. As we briefly review in the General Discussion, the notion of automaticity has been somewhat problematic, with different authors suggesting a number of (related) definitions. In the present work, we employed the common, and perhaps uncontroversial, view that a necessary condition of automaticity is that the process should be fast and goal-independent (see Moors & De Houwer, 2006, for a review). Thus, if perspective-taking is automatic, one should expect it to occur when participants are engaged in a secondary task (i.e., detecting a target), and when several seconds of scene viewing are not required for the effect to occur. We also aimed to test the perspective-taking theory using the most recent paradigm that has been employed in support of the theory, that is, the 'ambiguous number' paradigm.

We examined the perspective-taking theory via the use of eye movement measurement. Eye movements provide a relatively non-intrusive online measure of attention which allows attentional mechanisms to be studied under more naturalistic conditions than many other visual cognition paradigms (Findlay & Gilchrist, 2003). Several studies have shown that eye movements are influenced by social cues and illustrate how people generally look at objects that are looked at by others (Fletcher-Watson, Leekam, Benson, Frank, & Findlay, 2008; Kuhn, Tatler, & Cole, 2009; Leekam, Hunnisett, & Moore, 1998; Zwickel & Vo, 2010). Furthermore, the overall time spent inspecting an object, in addition to the time taken to fixate the object, provides a valuable index of attentional allocation. Although overt gaze following has been central to our understanding of ToM in infants (Butler, Caron, & Brooks, 2000; Caron, Kiel, Dayton, & Butler, 2002), eye movement measures have not been typically used to investigate the automatic perspective-taking claim (see Ferguson, Apperly, & Cane, 2017).

In the present Experiment 1, participants freely viewed everyday scenes that contained a model who either looked towards an object/area of interest or looked elsewhere. Orthogonally to the gaze direction, we manipulated whether the model's view of the object was occluded (by a natural barrier) or in full view. We predicted that participants would be faster to fixate the object when the person depicted in the scene looked towards it, i.e., a basic gaze cueing effect. Crucially, we predicted that this social facilitation *will* be modulated by whether the target object is visible or not to the model. Although some studies challenge the notion that mental state attributions influence gaze following (see above), this prediction was based on the fact that freely viewing a scene allows observers to employ higher mechanisms concerned with ToM. It is precisely under these circumstances that mental state attributions should influence gaze following. That is, when participants have time to consider what the model is looking at, i.e., non-automatically. In Experiment 2, participants viewed the same scenes but, rather than freely viewing the images, they performed a standard target discrimination task in which the targets were positioned at the object/area of interest employed in Experiment 1. In this scenario, we reasoned that participants' attentional set would be concerned with rapidly finding a prespecified target, thus vastly reducing the likelihood that they would consider what the model is viewing. We also used the barrier technique (Experiment 3) to examine the automatic perspective-taking claim and employed the centrally located gaze cue method. In the final experiment (Experiment 4), we used the 'ambiguous number' paradigm, together with the visibility manipulation, as a relatively new test of the perspective-taking theory.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixty (47 female) Brunel University students took part. Age ranged from 18 to 44 years (M = 20; SD = 5). All participants reported normal or corrected to normal vision.

2.1.2. Stimuli and apparatus

Twelve different scenes were photographed. Each contained one model who either looked towards an object/area of interest (e.g., a cup of coffee) or away from it. Each scene also contained an object that was located such that it could act as a barrier between the model and object (see Fig. 1). Thus, there was a total of 48 images generated, i.e., 12 scenes, each with a valid and invalid gaze cue, and each with the object of interest being visible or occluded. Eye movements were recorded with a head-mounted, video-based eye tracker (EyeLink 1000; SR Research Ltd., Osgoode, Ontario, Canada), using a sampling rate of 1000 Hz. They were recorded monocularly and analysed using Eyelink Data Viewer (SR-Research). The images were presented on a 21-in CRT monitor (1024 × 768; 85HZ) using Experiment Builder presentation software (SR-Research), with a viewing distance of approximately