

Looking ahead: Anticipatory cueing of attention to objects others will look at

Katherine Joyce¹, Kimberley Schenke¹, Andrew Bayliss², Patric Bach^{1*}

¹ School of Psychology, Plymouth University, Devon, UK

² Department of Psychology, East Anglia University, Norwich, UK

* corresponding author patric.bach@plymouth.ac.uk

Tel: +0044 1752 584848

School of Psychology

Plymouth University

Drake Circus, Devon

PL4 8AA

United Kingdom

Total word count: 3996 (including abstract and references)

Key words: gaze cueing, joint attention, predictive coding, prediction, attention

Abstract

Seeing a face gaze at an object elicits rapid attention shifts towards the same object. We tested whether gaze cuing is predictive: do people shift their attention towards objects others are merely *expected* to look at? Participants categorized objects while a face either looked at this object, at another object, or straight ahead. Unbeknownst to participants, one face would only look at drinks and the other at foods. We tested whether attention was drawn towards objects “favoured” by a face even when currently looking straight ahead. Indeed, while gaze expectations initially had a disruptive effect, participants did shift attention to the faces’ favoured objects once learning had been established, as long as emotional expressions had indicated personal relevance of the object to the individual. These data support predictive models of social perception, which assume that predictions can drive perception and action, as if these stimuli were directly perceived.

Acknowledgements: Patric Bach's work was supported by the Economic and Social Research Council grant number ES/J019178/1. We thank Steven P. Tipper, University of York, for helpful comments during data collection and interpretation.

Looking ahead:

Anticipatory cueing of attention to objects others will look at

Humans routinely track what other people know, want and feel, for a large number of individuals, without being explicitly aware of it (e.g., Schneider, Bayliss, Becker & Dux, 2012). A key process in this ability is the human tendency to follow others' gaze. Seeing a face look at an object causes a rapid, spontaneous shift of spatial attention in the observer towards the same target (Driver et al., 1999; see Frischen, Bayliss & Tipper, 2007, for review). By synchronizing attention across interaction partners, this 'gaze cueing' effect may underpin, at least in part, mankind's remarkable ability for social learning, cooperation, and communication. Conversely, disruptions of this process may underlie some of the social difficulties in autism spectrum and related conditions (Baron-Cohen et al., 1995).

The present study tests whether gaze cueing is predictive. Do we just follow the overt gaze of other people, or do we also routinely shift our attention towards what we *expect* other people will look at? Everyday behaviour is ripe with such anticipatory attention shifts, where people's attention is guided by the relevance an object has, not to themselves, but to their interaction partners. Parents report that their attention is often drawn to objects (diggers, tractors, princesses and sweets) that their children would like. People helpfully point out objects others are looking for, and, when shopping, they have a knack of finding items that their friends would like.

Such predictive gaze cueing effects emerge naturally from recent hierarchical feedback models of perception (Friston & Kiebel, 2009; Clarke, 2013). In such models, perception is not only driven by bottom-up sensory information, but is directly informed by prior knowledge about forthcoming events. In many cases, these predictions can "stand in" for what is actually observed, and induce perceptual biases. For example, people misperceive the disappearance point of a moving object further along the trajectory than it really was

(Freyd & Finke, 1984). These predictions of future motion are guided by one's assumptions about the external forces acting on the object (cf. Hubbard, 2005) and are integrated, in a Bayesian manner, with actual stimulation (e.g., Roach, McGraw & Johnson, 2011). In the social domain, it has been shown that heads appear to be rotated further if their gaze suggests a looking goal in this direction (Hudson, Hung-Liu, & Jellema, 2009). Hands are perceived closer to an object when the observer anticipates a reach and further away if they anticipate a withdrawal (Hudson, Nicholson, Ellis & Bach, under review), and observers more speedily identify actions an individual typically carries out with an object, compared to actions that are equally likely but typical for someone else (Schenke, Wyer, & Bach, under review).

These findings provide initial evidence that people constantly predict others' future behaviour and that these predictions can, to some extent, stand in for sensory information, as if these stimuli were directly perceived. Here, we test, for the first time, whether observers make such predictions about others' gaze, and whether these gaze predictions lead to similar attentional shifts as directly perceived gaze. Prior research has provided suggestive evidence for this idea. For example, gaze cueing can be elicited not only by directly perceived gaze but merely by the *belief* that a person looks at an object (Teufel et al., 2010). Moreover, Frischen and Tipper (2006) have reported that, when re-encountering a face one has seen before, attention is guided not only by the face's current gaze, but by its gaze in the previous encounter. Frischen and Tipper interpreted this as a recall of the observer's own prior attentional state, but it might also reflect an anticipatory coding of gaze (cf. Bayliss & Tipper, 2006). The person in the photograph had been looking left so we predict the photograph to show the same gaze when seen again.

To test whether gaze expectations can induce such anticipatory gaze cueing, we adapted the standard gaze cueing paradigm. In each trial, participants saw one of two faces in the middle of two objects. A target appeared on one of the objects and participants indicated,

in a speeded response, whether the object was a food or a drink. In half the trials, the face would gaze – and smile – at one of the two objects just prior to the cue. Participants were instructed to ignore these gaze shifts, and we manipulated, unbeknownst to participants, which objects the two faces would look at. One face would only look and smile at drinks, and the other at foods. According to prior work, these gaze shifts should involuntarily direct attention to the gazed at object, rendering it easier to identify.

The key question is whether people form gaze expectations about the two individuals, such that attention would shift towards their favoured object, even when the individuals looked straight ahead. Participants should then more quickly identify objects the face would typically gaze at, compared with objects it typically looks away from, specifically in the second half of the experiment, when participants have learned the relationships between faces and objects. Such an effect could not be explained by a mere reactivation of the observer's prior attentional state. Overall, both faces equally often looked left and right, and their favoured objects were seen equally often in both locations. Finding gaze cueing effects towards the faces' favoured objects would therefore reveal an anticipatory cueing of attention guided by predictions about the other individuals' behaviour.

Method

Participants

Sixty-two students from Plymouth University (18 to 52 years, 14 male) took part. Sample size was determined by a power analysis conducted on pilot data (see supplementary material). For half of the participants, the presented faces were cartoon faces and for the other half real face photographs. Cartoon faces are typically more effective in evoking gaze shifts, most likely because their visual simplicity facilitates cue extraction (e.g., Hietanen & Leppänen, 2003) and are therefore the most powerful test of the experimental hypothesis. Face photographs, in contrast, allow us to test whether any effects generalize to more realistic

face representations that more plausibly represent different individuals. Two participants, one from each group, were excluded because they made more than 10% errors.

--- Figure 1 about here ---

Material and apparatus

All participants completed the Autism Quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). The experiment proper was administered via Presentation (Neurobehavioral systems, version 14.9), on a Windows XP SP3 1280x1024 32 bit colour 17" display. Stimuli were assembled from 25 images: a fixation cross (the '+' symbol in Microsoft's Trebuchet font), four colour photographs each of foods (orange, cupcake, apple and hotdog) and drinks (cola can, orange juice, milkshake and coffee) and cartoon and photograph versions of a male and a female face. The faces could either look straight ahead with a neutral expression, blink (eyes closed) with a neutral expression, or smile while looking left or right (Figure 1).

Design and Procedure

The participants were seated roughly 60 cm away from the monitor. They filled out the Autism Quotient, were given verbal and on screen instructions, and then completed 16 practise trials. The experiment proper consisted of 256 trials. Participants were allowed to press 'p' at any point to take a break.

Each trial started with a fixation cross (400 ms). After a brief blank (600 ms), one of two faces appeared in the middle of the screen, with a food item and a drink on either side (locations counterbalanced across trials). The face looked straight ahead with a neutral expression for a random time interval between 500 and 1100 ms. and then blinked for 100 ms. The next image showed the face with eyes open, in 50% of the trials looking straight

ahead with a neutral expression, or smiling and looking at one of the objects in the other 50%. One face would always look at the drinks and never at foods, while the other face showed the opposite behaviour (counterbalanced between participants). After 400 ms, a blue square appeared on one object and participants categorised this object as either a drink or a food by pressing either 'h' or the space bar. This image was presented until participants made a response or a maximum trial time of 4 seconds had passed. Feedback was given for error and timeouts reminding participants of the key assignment.

After the experiment, participants were asked whether they noticed a pattern in the stimuli; none did and several expressed surprise when the manipulation was revealed.

Results

As during piloting (supplementary material), trials (1.7% in total) were excluded if they contained: responses before or after the response interval, responses with uncertain presentation timing (> 10 ms.), or responses with RTs beyond 3 standard deviations from the participant's mean in the straight or sideways gaze trials. As participants initiated pauses themselves, trials following or preceding pauses were also excluded. Participants with AQ scores in the clinical range (>32) would have been excluded though none did.

Data were analysed separately for the gaze towards object and gaze straight trials. The gaze trials tested whether our paradigm evoked the expected gaze cuing effects, but do not say anything about gaze expectancies, as the faces looked at their favoured object in all of these trials. In contrast, the trials with straight gaze measured gaze expectancies, and whether attention is directed towards the favoured object even when the face looks straight ahead.

Regular gaze cueing

We first confirmed that our paradigm elicited the typical gaze cueing effects. We analysed the gaze towards object trials with a 2X2 mixed-factors ANOVA with Object (looked at, not looked at) and Block (1, 2) as repeated measures factors and Group (cartoon

faces, photographs) as the between subjects factor. The analysis of RTs (Table 1) only revealed a main effect of Object $F(1,58)=90.8$, $MSE=1563$; $p<.001$, $\eta_p^2=0.610$, but no other effects ($F_s<1$). Participants more quickly categorized a looked-at object than a not looked-at object, in both block one, $F(1,58)=54.7$; $MSE=1248$; $p<0.001$, $\eta_p^2=0.486$) and block two, $F(1,58)=58.6$; $MSE=1259$; $p<.001$, $\eta_p^2=0.502$). The analogous analysis of error rates (Table 1) revealed no effects ($F_s<1$), but an interaction of Object by Group, $F(1,58)=13.3$; $p<.001$, $MSE=.019$; $\eta_p^2=0.19$. For cartoon faces, participants made more errors for not looked-at-objects, and for photographs for looked-at objects (Table 1).

Table 1. Regular gaze cueing effects for both response times and error rates. Values in brackets show the standard deviation in the condition.

Group	Response Times (ms.)		Error rates (%)	
	looked at	not looked at	looked at	not looked at
Cartoons	509 (83)	562 (89)	3.3 (2.9)	5.2 (3.6)
Photographs	544 (116)	587 (114)	4.4 (3.8)	2.8 (2.5)

Anticipatory gaze cueing

The crucial test was whether observers would also shift their attention to the expected object when the face looks straight ahead. RTs and Error rates in the straight gaze trials were analysed with the same ANOVA as the regular gaze trials, with the Object factor now coding whether the target appeared on the object the face typically looked at, or the other object.

The analysis of RTs revealed neither an effect of Object, $F<1$ nor of Block, $F(1,58)=3.5$, $MSE=2052$, $p=.066$, $\eta_p^2=0.057$, but a significant interaction between both factors, $F(1,58)=12.116$, $MSE=761$, $p=.001$, $\eta_p^2=0.173$. Step down ANOVAs revealed that, in Block 1, participants were faster to categorize typically looked away from objects compared

to typically looked at objects, $F(1,58)=6.2$, $MSE=822$, $p=.015$, $\eta_p^2=0.097$, while Block 2 revealed the expected pattern of faster responses for objects the person typically looks at, $F(1,58)=4.9$, $MSE=834$, $p=.031$, $\eta_p^2=0.078$. Group did not interact with any factor ($F_s<1$). The Object by Block interaction was replicated in the photograph group, $F(1,29)=4.6$, $MSE=874$, $p=.040$, $\eta_p^2=0.138$, and the cartoon group $F(1,29)=8.0$, $MSE=647$, $p=.008$, $\eta_p^2=0.217$, when analysed separately.

--- Figure 2 about here ---

The analysis of error rates revealed a significant effect of Block, $F(1,58)=6.7$, $MSE=.001$, $p=.012$, $\eta_p^2=0.104$, with fewer errors in the second half of the experiment. Importantly, there was a significant effect of Object, $F(1,58)=4.7$, $MSE=.001$, $p=.034$, $\eta_p^2=0.075$, but no interaction, $F<1$. Step-down ANOVAs revealed that participants made fewer errors for typically gazed at objects in Block 2, $F(1,58)=4.863$, $MSE=.001$, $p=.031$, $\eta_p^2=0.077$, but not in Block 1, $F(1,58)=1.065$, $MSE=.001$, $p=.306$, $\eta_p^2=0.18$. Group did not interact with any other factor ($F_s<2.7$).

A separate study (see Supplementary Experiment) tested whether the same effects are also obtained when the observed faces did not smile but showed a neutral expression when looking at the objects, but was otherwise identical. While this study replicated the regular gaze cuing effects, with faster responses for looked at compared to not looked at objects, all predictive gaze cuing effects in both RTs and Error Rates were eliminated, all $F<1.1$. Indeed, across experiment comparisons revealed that the predictive gaze cuing effects in the main experiment differed significantly from this control experiment, both for the RT interaction, $F(1,118)=10.5$, $MSE=722$, $p<.002$, $\eta_p^2=0.082$, and the error rate effect in Block 2, $F(1,118)=4.150$, $MSE=0.004$, $p<.044$, $\eta_p^2=0.034$.

General Discussion

We tested whether gaze cueing of attention is driven only by bottom-up gaze cues, or whether it is predictive, and is guided by the gaze a person is only *expected* to show.

Participants categorized objects while a face either looked at this object, at another object, or straight ahead. Unbeknownst to participants, we manipulated the looking behaviour of the faces, such that one face (e.g. the male) would only look at foods and the other (the female) only at drinks. We tested whether participants would derive implicit gaze expectations from these contingencies, and whether these expectations would elicit similar shifts of attention even when the faces looked straight ahead.

As predicted, the second half of the experiment, after participants had learned the individuals' looking behaviours, revealed such anticipatory cueing effects. Participants categorized objects more quickly and reliably if these objects were typically looked and smiled at by the individual, compared to the objects that they typically looked away from. These anticipatory shifts of attention were observed even though (1) neither the face, nor gaze direction was task relevant, (2) gaze direction did not predict the location of the task-relevant object, and (3) across faces, all objects were looked at equally often. The data therefore indicate that, by Block 2, observers had implicitly formed internal models of the individuals' gaze behaviour, which then directed attention towards the expected object, as if the gaze shift were directly observed. Gaze cueing of attention is therefore not only driven by bottom-up sensory information, but also by expectations about others' gaze, supporting the assumption that predicted stimulation can drive action and cognition, as if the event was directly observed (cf. Clarke, 2013; Friston & Kiebel, 2009).

We had expected that these effects would slowly build up, and that anticipatory gaze cueing was small or absent in the first half of the experiment. However, an unexpected pattern emerged in these earlier trials. Participants were *slower* to identify objects that the

other person typically looked at, compared to objects they typically looked away from. The shift from negative to positive gaze cueing effects over the experiment was replicated in both the face photograph and cartoon faces groups. It suggests that participants learn about the predictive relationships between individuals, objects and gaze very quickly. At the start, these predictions have a disruptive effect. Only when fully established do they guide attention similarly as directly observed gaze.

Although not predicted, such negative effects are in line with hierarchical feedback models of perception. These models conceptualize perception as an integration of sensory information and top-down predictions, such that predictions can “stand in” for what was actually perceived (Clarke, 2013; Friston & Kiebel, 2009). These assimilative effects of prior information are, however, typically only observed when the stimuli are not directly attended (Kok et al, 2011). Attended stimuli may cause reverse effects and bias perception away from the predictions. Classic examples are visual after images, where observing – or even imagining – a movement to the left causes subsequent static images to be perceived as moving subtly to the right (Wade, Thompson & Morgan, 2014), and we have observed similar contrastive effects in motor activation when famous athletes were seen *not* performing their typical sports (Bach & Tipper, 2007; Tipper & Bach, 2010). In the same way, here, the negative cueing effects might emerge because the individual’s straight gaze is coded relative to the expected looking behaviour, causing a bias in the opposite direction. When, however, participants become more effective at diverting attention away from the distracting faces over the course of the experiment, these effects revert to positive ones, revealing the expected superposition of sensory information and prior expectations. Of course, at the moment, these considerations are speculative and need to be further verified.

An interesting observation was that the predictive gaze cueing effects depended on the faces showing an emotional response when looking at the objects (see also Bayliss et al.,

2009, 2010). In a separate experiment (Supplementary Material) – initially used for piloting, but increased to full sample size at the request of a reviewer –, we ran exactly the same procedure, but with the faces showing a neutral expression when looking at their objects. While the regular gaze cueing effects were identical across experiments, all predictive gaze cueing effects were now eliminated. This difference is striking given that predictive coding models do not necessarily distinguish between social and other effects on perception. This dependency on an emotional expression reveals a very much social coding of the individuals' gaze, reflecting, perhaps, the mental states that the emotional expressions imply, or the emotional relevance the object has to the individuals (cf. Manera, Elena, Bayliss, & Becchio, 2014). It is in line with the idea that internal models of other people do not only reflect their behaviours towards objects, but also the internal states these behaviours imply (Barresi & Moore, 1996). Our new data therefore provide first evidence for models suggesting that such top-down knowledge can be converted into concrete predictions of others' gaze that drives one's own attention system as if the gaze shifts were directly observed, as predicted by predictive coding of social perception (e.g., Kilner et al., 2007).

An important question is whether the predictive effects on gaze cueing reflect explicit or implicit learning of the contingencies. This question was not central to the current hypotheses, but in a post-experiment screening none of the participants reported awareness of the patterns. This unawareness probably emerges because the faces of the individuals were task irrelevant and there was no relationship between gazed-at-object and object-to-be-categorized. Indeed, many participants stated that they attempted to ignore the faces altogether. Moreover, in a prior study, we showed that participants often remain unaware of similar relationships between individuals, situations and behaviours (Schenke, Wyer & Bach, under review), even when alerted, several times during the experiment, that such a pattern was present. Together with the reports of our participants, these data indicate that the typical

behaviour of others is learned in a largely implicit manner, but nevertheless directly affects perception and guidance of attention.

Conclusions

Gaze cueing of attention is not only driven by bottom-up sensory information, but also by prior expectations about what other individuals will look at in a given situation. Because gaze cueing is central for joint attention, learning from others and fluent social interaction, such a predictive mechanism could have far reaching impacts on social interaction, allowing people to attend to what other's will look at in the future or direct their attention to objects that are relevant to their interaction partners. Future experiments need to explore the role of the implicit/explicit nature of the predictions, and why, in the first half of the experiment, predictions can give rise to paradoxical negative gaze cueing effects.

References

- Bach, P. & Tipper, S.P. (2006). Bend it like Beckham: embodying the motor skills of famous athletes. *Quarterly Journal of Experimental Psychology*, *59*(12), 2033-2039.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31* (1), 5-17.
- Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*(4), 379-398.
- Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*, 107–154.
- Bayliss, A. P. & Tipper, S. P. (2006) Gaze cues evoke both spatial and object-centred shifts of attention. *Perception and Psychophysics*, *68*(2), 310-318.
- Bayliss, A. P., Schuch, S. E., & Tipper, S. P. (2010). Gaze cueing elicited by emotional faces is modulated by emotional context. *Visual Cognition*, *18*(8), 1214-1232.
- Bayliss, A. P., Griffiths, D. & Tipper, S. P. (2009). Predictive gaze cues affect face evaluations: The effect of facial emotion. *European Journal of Cognitive Psychology*, *21*(7), 1072-1084.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 1-73.
doi:10.1017/S0140525X12000477
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, *6*(5), 509-540.
- Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental*

- Psychology: Learning Memory and Cognition*, 10, 126-132.
- Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. *Visual Cognition*, 14(3), 351-364.
- 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.
- Friston, K.J., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, 22, 1093-1104. doi: 10.1016/j.neunet.2009.07.023.
- Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, 12, 822-851.
- Hudson, M, Hong-Liu, C. & Jellema, T. (2009). Anticipating intentional actions: The effect of eye gaze direction on the judgment of head rotation. *Cognition*. 112, 423-434.
- Hudson, M., Nicholson, T., Ellis, R. & Bach, P. (under review). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions
- Kilner, J.M., Friston, K.J., & Frith, C.D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processes*, 8, 159-166. doi: 10.1007/s10339-007-0170-2
- Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2011). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, bhr310.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, 24(8), 1377-1388.
- Manera, V., Elena, M. R., Bayliss, A. P. & Becchio, C. (2014). When seeing is more than looking: Intentional gaze processing modulates object desirability. *Emotion*, 14(4), 824-832.
- Roach, N. W., McGraw, P. V., & Johnston, A. (2011). Visual motion induces a forward prediction of spatial pattern. *Current Biology*, 21(9), 740-745.

- Schenke, K., Wyer, N. & Bach, P. (under review). I know what you will do. Observers implicitly integrate information about actor and object to predict forthcoming actions.
- Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General*, 141(3), 433.
- Teufel, C., Alexis, D. M., Clayton, N. S., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception, & Psychophysics*, 72(3), 695-705.
- Tipper, S.P., Bach, P. (2011). The face inhibition effect: Social contrast or motor competition? *Journal of Cognitive Psychology*, 23(1), 45-51.
- Wade, N. J., Thompson, P., & Morgan, M. (2014). The after-effect of Adolf Wohlgenuth's seen motion. *Perception*, 43, 229-234. doi: 10.1068/p4304ed

Figure captions

Figure 1. Schematic of the trial sequence. After a fixation cross (400 ms.) and a blank screen (600 ms.), participants saw either a male or female face flanked by a drink and a food item. After a random time interval (between 500 to 1100 ms.), the face blinked (for 100 ms.). When the eyes opened again, the face either looked straight ahead (not shown) or at one of the two objects, and a blue square marking the object to be categorized appeared 400 ms. later. We manipulated the gaze behaviour of the two faces such that, if looking sideways, one face would only look at drinks and the other only at food (counterbalanced between participants).

Figure 2. Response times (left panels) and Error Rates (right panels) in the straight gaze trials, for the cartoon faces (top row) and real faces (bottom rows) groups separately. In each figure, the left two bars show the data for the first half of the experiment, and the right two bars show the data for the second half of the experiment. The black bars show categorization response for objects that are typically looked at by the shown individual, and the white bars show objects that this individual typically looks away from. Error bars show the standard error of the mean.