Do the Eyes Have it?
Cues to the Direction of Social Attention

Stephen R. H. Langton, Roger J. Watt and Vicki Bruce

University of Stirling

Correspondence to: Dr. S. R. H. Langton, Department of Psychology, University of Stirling, FK9

4LA, Scotland.

E-mail: srhl1@stirling.ac.uk

Tel: +44 (0) 1786 467659

Fax: +44 (0) 1786 467641

Abstract

The face communicates an impressive amount of visual information. We use it to identify its owner, how they are feeling and to help us understand what they are saying. Models of face processing have considered how we extract these kinds of meaning from the face but have ignored another important facial signal - eye gaze. However, recent neurophysiological and developmental studies have sparked some interest in the perception of gaze on the part of cognitive psychologists. In this article we begin by reviewing evidence suggesting that the eyes may constitute a special stimulus in at least two senses. First the structure of the eyes may have evolved to provide us with a particularly powerful signal to the direction in which someone is looking, and second, we may have evolved neural mechanisms devoted to their processing. As a result, gaze direction is analysed rapidly and automatically, and is able to trigger reflexive shifts of an observer's visual attention. Although the eyes are an undoubtedly important cue, understanding where another individual is directing their attention involves more than simply analysing their gaze direction. We go on to describe research with adult participants, children and non-human primates suggesting that other cues such as head orientation and pointing gestures make significant contributions to the computation of another's direction of attention.

2

Since the early 1980's, considerable progress has been made in understanding the perceptual, cognitive and neurological processes involved in deriving various different kinds of meaning from the human face[1,2]. For example, we now have a much better understanding of the operations involved in recognising a familiar face, categorising the emotional expression carried by the face, and of how we are able to use the configuration of the lips, teeth and tongue to help us interpret what the owner of a face is saying to us. In their influential model of face processing Bruce and Young[3] proposed that these three types of meaning - identity, expression and facial speech - are extracted in parallel by functionally independent processing systems, a suggestion for which there is now converging empirical support[4] (though see Walker et al.[5] and Schweinberger & Soukup[6] for some complications).

However, in common with other cognitive models of face processing, Bruce and Young's account neglected a number of additional facial movements that convey important meaning and make substantial contributions to interpersonal communication. One such signal - gaze - has been widely studied by social psychologists who have long known that it is used in functions such as regulating turn-taking in conversation, expressing intimacy, and exercising social control[7]. Despite this, interest in the perceptual and cognitive processes underlying the analysis of gaze and gaze direction has only emerged in recent years, perhaps stimulated by the work of Perrett[8,9] and Baron- Cohen[10,11]

Perrett and his colleagues have proposed a model which is based on neurophysiological research which we outline later in this article, and describes how we combine information from gaze, head and body to determine where another individual is directing their attention (see Box 4). Baron-Cohen's account has a wider scope, he proposes a "mindreading" system comprising a collection of modules which have evolved to enable humans to attribute mental states to one another (see Box 1). One of these modules functions specifically to detect another's gaze direction and attributes the mental state of "seeing" to the gazer. By implicating the perception of gaze in our understanding of what another person is attending to, or what they are thinking about, these researchers highlight the potentially central role played by the perception and

interpretation of gaze in the processes of social cognition. Central to both models is the importance of eye gaze as a cue to the direction of another's attention. Eye gaze cues override head and posture cues in Perrett's account, and in Baron-Cohen's model, the ability to use gaze direction to establish joint attention underpins the development of a theory of mind. Here we briefly review the evidence for these models which have rendered gaze perception an important issue in contemporary cognitive research. We argue that the emphasis on eye gaze within these models has led to the neglect of other important cues from head, posture and gesture, in the perception and computation of attention direction.

Box 1 - Baron-Cohen

The Perception and Detection of Gaze

Humans and many other species tend to look at things in their environment that are of immediate interest to them. You might be the recipient of another's gaze, for instance, because you are a potential meal, a mate or simply because you are someone with whom they would like to interact. Individuals who are able to detect rapidly when they are the object of another's attention, and who can analyse exactly where another's gaze is directed therefore have considerable adaptive advantage. How might evolution have equipped us to deal with this problem? First, we may have evolved dedicated brain mechanisms for recovering the relevant information from another's eyes early in visual processing. A second possibility is that the physical structure of the eye may have evolved in such a way that eye direction is particularly easy for our visual systems to perceive. Indeed, recent work suggests that the output of simple cells found in the visual cortex can, in principle, signal the direction of gaze (see Box 2). Of course these two viewpoints are not necessarily mutually exclusive; the eye may be a special stimulus and we may have evolved brain mechanisms to perceive it.

Box 2 - Watt

4

As part of his "mindreading" model, Baron-Cohen emphasises the latter position. He has proposed the existence of an Eye Direction Detector (EDD) in humans, a functionally specialised *module* devoted to the task of detecting eyes, and computing where in the environment eye-gaze is directed. Whether or not such a specialised system exists, we might expect the eyes to form a special kind of stimulus that we are able to process rapidly and obligatorily, and to which we are particularly sensitive. If Baron-Cohen's position is correct, we might expect these behavioural properties to be underpinned by some specialised neural circuitry tailored to processing the eyes and the direction in which they are gazing. In the remainder of this section we describe some of the evidence for these claims.

The ability to process eyes and eye-like stimuli appears to develop very early in humans. By the age of 2 months infants show a preference for looking at the eyes over other regions of the face[12], and by 4 months they are able to discriminate between direct and averted gaze[13]. Evidence of this sort doesn't mean that the ability to detect gaze is "hardwired" or present from birth, but it does suggest that by the time infant's visual acuity is sufficiently developed they show a particular preference for the eyes. By adulthood, subjects in experiments are extremely efficient at searching for a direct gaze amongst an array of distracting leftward and rightward gazes, significantly more so than they are at searching for equivalent geometric control stimuli[14]. This suggests that the ability to detect the eyes and the direction of gaze could be based on more than basic low-level perceptual abilities such as visual acuity and contrast sensitivity (but see Box 2).

Fodor[15] argued that putative modules should, in addition to a number of other criteria, operate rapidly and mandatorily (see Coltheart[16] for a recent discussion of modularity). Thus, if we are to take seriously the notion of a gaze module we must show that the processing of gaze also occurs rapidly and obligatorily. Some evidence for this is provided by a recent set of studies that have demonstrated that gaze cues are able to trigger an automatic and rapid shifting of the focus of a viewer's visual attention. In one study[17], three month old infants turned their eyes to a target faster if the location of that target had just been cued by an adult's gaze direction. Other

5

studies[18-20] have used the more traditional cueing paradigm devised by Posner[21] where participants are asked to make a response to a target stimulus whose location may or may not have been cued, for example, by the orientation of a head and/or the direction of eye-gaze (see Box 3). These experiments have shown that gaze cues will trigger rapid, reflexive shifts of adult participants' visual attention, even when the gaze direction does not predict the likely location of a target stimulus, and when participants are explicitly asked to ignore these cues.

Gaze cues therefore seem to be processed obligatorily and cause viewers' attention to be shifted toward the cued region. This has the effect of facilitating the processing of any target which subsequently appears in that location, and somehow primes an infant's eye-movement response in that direction. What is remarkable about these findings is that in the experimental psychology literature only cues presented in the periphery of participants' visual fields have been found to exert these kinds of reflexive orienting effects (see Box 3). Arrows presented in the centre of a computer screen, for instance, do not trigger reflexive shifts[22]. Attention generally seems to be *pulled* toward brief peripheral visual onsets but social cues seem to be unique in causing attention to be automatically *pushed* in the direction they indicate.

Box 3 - Cueing

and Box 4- Perrett

Finally, neurophysiological and neuropsychological work has provided some evidence for the existence of a neural system dedicated to processing the direction of gaze. Using a technique that enables the activity of a single nerve cell to be recorded, Perrett and his co-workers have identified certain cells in the STS region of the macaque temporal lobe which respond maximally to the particular direction in which the eyes in that face are looking (See Box 4). For example, one population of cells fire with maximum frequency when the monkey sees another individual gazing upwards, and another population of cells respond well to gazes directed downwards[9,23] Moreover, when this region of the macaque cortex is removed these monkeys are unable to make

6

gaze direction judgements, but nevertheless perform well on a number of other face processing tasks[24]. Humans suffering damage to the equivalent part of the brain have also been shown to be impaired in gaze recognition tasks[25,24].

The above evidence suggests that gaze direction may indeed be analysed quickly, and that gaze direction apparently cannot be ignored - it seems to trigger reflexive shifts in an observer's visual attention. The extent to which such effects are due to specialisation of the internal perceptual machinery, the nature of the eye itself and the signals it sends, or both requires further consideration (cf Box 2).

The Importance of Other Cues

Whatever the reasons for our sensitivity to shifts in eye-gaze we must not neglect other cues. Where someone is perceived as directing their attention may depend, not just on the direction of eye gaze, but on the orientation of their head, the posture of the body and perhaps where they are pointing their finger. It has been suggested that these cues are *all* processed automatically by observers and *all* make contributions to decisions about another individual's social attention.

Experimental Studies with Adults

As long ago as 1824, William Wollaston[26] noted that judgements of gaze direction are not based solely on the position of the iris and pupil relative to the whites of the eyes. Wollaston's original drawings and our own images (see Figure 1) clearly demonstrate how head orientation can influence the *perception* of gaze. More recently, several authors have systematically investigated this phenomenon. In general, this work has established two types of effect. First, the perceived direction of gaze can be "towed" toward the orientation of the head. In this case, as with the Wollaston images, the direction of gaze is perceived to be somewhere between the angle of the head and the true line of regard of the eyes[27,28]. The second kind of influence of head angle on the perception of gaze is a kind of overshoot effect. Imagine someone standing in front of you with their head 30 degrees or so to your right and with their eyes either staring straight back at

7

you, or back towards your left shoulder. Under these conditions, you might perceive their eyes to be gazing a little further to the left than they actually are[29,30].

Figure 1

Regardless of whether perceived gaze is towed towards the head, or appears to overshoot its target, the point is that the *perception* of gaze must be based on some combination of information extracted from the eyes and information extracted from the orientation of the head. However, at later stages of information processing we also see that computing where another individual is directing their attention depends on a number of further social cues. In fact, some of the neurophysiological work described earlier has hinted that this might indeed be the case[23]. As described earlier, these studies indicated that certain cells in the macaque temporal cortex respond strongly to certain gaze orientations. However, these same cells were also found to be sensitive to *conjunctions* of eye, head and body position, suggesting that all of these cues might contribute to the processing of attention direction. Moreover, Perrett and his colleagues have suggested how these cues might contribute to the computation of attention direction. They contend that information from gaze, head and body is combined hierarchically in a mechanism they called the Direction of Attention Detector (DAD). In this model, attention direction will be signalled by the eyes if these are visible, but if they are obscured, or if the face is viewed at too great a distance, the head will carry the burden of signalling attention direction. If for some reason, information from the eyes *and* head is unavailable, attention direction is signalled by the orientation of the body (see Box 4). Thus, although the model stresses that other cues are involved in the computation of attention direction, these cues play a relatively minor role compared to the part played by the eyes.

Experimental work with human subjects is also beginning to indicate that decisions about the direction of another's attention are based on a number of different cues[31,32]. However, this research has lead to some rather different conclusions concerning the way in which these signals

contribute to the computation of attention direction. In some of these experiments the directional cues of interest are placed into conflict in a Stroop-type interference paradigm. In one study, participants were shown the stimuli illustrated in Figure 2 one at a time on a computer screen and, in one block of trials, they were asked to push a button on a keyboard contingent on the direction of the eye gaze. Although participants were asked to ignore the orientation of the head, the results indicated that they were unable to do so. Reaction times (RTs) were faster when the eye gaze and head were oriented in the same direction than when they were oriented in opposite directions. An identical pattern of results was obtained when participants were asked to do the opposite task, that is respond on the basis of the orientation of the head and ignore the direction of eye gaze (see Figure 3). In a second experiment, participants were again presented with the stimuli shown in Figure 2, but this time they were asked to ignore these images and to make their responses contingent on a spoken directional word (either "up", "down", "left" or "right") which was presented at the same time as each face. Head and eye gaze cues were found to exert equal and independent effects on the speed of participants' responses to the spoken words. So on hearing the word "up", participants responded faster when they saw the gaze and head oriented upwards compared with trials when the head and eye gaze were directed downwards. However, this effect was completely eliminated when the head and gaze were oriented in opposite directions.

Figures 2 & 3

The results of these two experiments clearly suggest that participants process the directional information provided by the orientation of the head as well as that provided by the angle of eye gaze. In addition, they process both sources of information in parallel even when the experimental task demands that they attend to information in a completely different modality. Finally, the findings of these experiments suggest that head and gaze are more equal partners in the computation of attention direction than predicted by Perrett's DAD model where information from the head is overridden by the direction of eye gaze.

9

Other studies have indicated that, in addition to head and eye gaze cues, pointing gestures also contribute to decisions about the direction of another's attention. These cues also produce interference effects on responses to spoken directional words[33] and seem to be processed automatically and in parallel with directional cues provided by the head and eyes[32]. Taken together with the studies described above, it seems as though observers process directional cues provided by the eyes, the head, and pointing gestures, and do so in parallel so that all this information is available when a decision has to be made about where another individual is directing their attention. Clearly these other cues play a greater role in this computation than they do in Perrett's DAD model, and in Baron-Cohen's mindreading system which makes no reference to any other visual cues to attention direction.

Developmental Studies

Baron-Cohen contends that eyes form a particularly salient feature for the developing infant. This may be so, however a number of studies with young children have shown that secondary cues such as head orientation and pointing gestures might actually be more salient signals to the direction of another's attention than eye-gaze direction alone. In gaze following experiments, children sit in front of their mothers who attempt to engage them in eye contact. Having done so, the mothers shift their eyes and/or turn their heads away from the child and their following behaviour is observed. Using this procedure studies have shown that infants as young as 3 - 6 months are able to follow a combination of head and eye cues[34,35], but it is not until between 14 - 18 months that they show any indication of following eye cues alone[36]. Prior to 14 - 18 months it seems as though children actually ignore the orientation of the eyes and simply use the position of the head as an attention following cue[37].

What is not clear from many gaze following studies is whether or not the child actually *understands* the mental experience of their mother. Can the child who follows their mother's gaze to a target object actually represent the fact that the mother "sees" that particular object, or is the behaviour simply an example of the kind of reflexive attentional orienting mechanism observed in the precueing studies described earlier? In order to explore whether young children

are able to use attentional cues such as eye gaze to infer another's mental state researchers have used a different kind of task. Children were shown a picture of a face gazing at one of four objects and are asked "which one is Sam looking at?". In general, children of around 2 - 3 years failed this task, whereas children of around 4 years and older performed well indicating that the latter but not the former are able to infer the mental state of "seeing" from another's gaze direction[38-40]. However, the performance of the younger children was dramatically improved when the gaze cues were presented in concert with cues from the orientation of the head, or were replaced by pointing gestures[40].

In summary, children are able to follow head cues and use information from the orientation of the head to choose which object is being looked at before they are able to perform these tasks on the basis of eye direction alone. Thus although sensitive to gaze from an early age, it seems as though children are most influenced by information from other individuals' gestures and head orientation in order to engage in joint visual attention and gather information about the world.

The Perception of Gaze by Non-Human Primates

Comparative work with non-human primates also suggests that the orientation of the head might provide a stronger cue to another's attention direction than eye-gaze alone. In some of these studies animals had to learn to use an experimenter's attention cues in order to obtain food hidden under one of two objects. Despite undergoing extensive training, capuchin monkeys were unable to use an experimenter's gaze cues in order to make the correct object choice, but learnt to make use of pointing gestures and a combination of head and gaze cues in order to perform the task successfully[41]. Monkeys also failed to orient spontaneously to experimenter-given eye, head or pointing cues in gaze-following experiments similar to those used with human infants[42], but are able to follow eye plus head cues of another individual of the same species[43]. Thus, monkeys need eye-gaze cues to be accompanied by a turn of the head in order to provoke a gaze following response, or to enable them to obtain a reward in an object choice task. In contrast to non-ape species, there is evidence that chimpanzees are able to make use of eye gaze cues in both types

of study [42,44,45]. However, in the object-choice task they required fewer sessions to reach criterion using pointing cues and head/gaze turns than when learning to use gaze-only cues[44].

In general then, it seems that turns of the head are more important cues, at least for monkeys, than movements of the eyes alone. However, this conclusion is perhaps not all that surprising given what we now know about the external morphology of the primate eye. Unlike humans, primates do not have a widely exposed white sclera surrounding a much darker iris[46]. In most species the colour of the sclera is rather similar to that of the skin around the eyes so that, compared with humans, the direction of gaze will be relatively camouflaged, perhaps in order to deceive predators, prey, or even fellow primates who might be in competition for scarce resources. We humans may have evolved eyes with a greater contrast between iris and sclera precisely because the risk of predation is minimal, and the benefits of an enhanced gaze signal in terms of communication and co-operation far outweigh the cost of an inability to deceive.

Conclusion

Recent interest in the study of gaze and social attention on the part of cognitive psychologists has been sparked, in part, by the work of Baron-Cohen and Perrett. These researchers have proposed somewhat different models, linked by the fact that in each, the detection of eye-gaze and gaze direction plays a pivotal role. Our own research has confirmed the importance of cues from direction of gaze. Such cues cannot be ignored even when they are irrelevant to the task in hand, and can create reflexive shifts in visual attention. However, recent experimental work with adults, children and non-human primates has suggested that the orientation of the head makes a larger contribution to the processing of another's direction of attention than these models allow. Studies have shown that the perception of gaze direction depends on the orientation of the surrounding face, whilst others have demonstrated that observers process head and gesture cues automatically and that this information contributes to decisions about social attention direction.

Thus, it is clear that the models proposed by Perrett and Baron-Cohen need some modification. In line with this, Perrett and Emery have already suggested that Baron-Cohen replace his Eye Direction Detector with something akin to their Direction of Attention Detector

12

in order to take account of the neurophysiological findings. However, it is equally clear that the DAD itself needs some modification if it is to accommodate the experimental evidence reviewed in this paper.

Outstanding Questions

• How does the social attention computation fit in with the Bruce and Young[3] framework for face processing? Should the putative Direction of Attention Detector be considered as part of the face processing system? The evidence that body movements (posture and hand gestures) may contribute information to the computation of social attention direction is problematic for this view.

• How do we combine the directional information extracted from the various cues at the perceptual level of analysis and at later stages of processing where decisions are computed? Is the information actually integrated or combined, or does the context provided by one cue somehow modulate the processing of the other cues? How does the system code the orientation of the head?

• How do we disambiguate gaze cues? Shifts of gaze and/or turns of the head serve a number of different functions. They can act as intentionally communicative signals, illustrating the referent of a remark, disambiguating deictic expressions such as "this one" or "that one", expressing intimacy or dominance, or communicating various emotional states[11]. But gaze shifts need not be intentionally communicative at all - for example when we gaze upwards when thinking. Given all the different functions, and the range of meanings the eyes and head might express, it is difficult to imagine how we are able, in the most part, to interpret just what another's gaze actually means. Do different kinds of signals have different spatial and temporal properties which can be used to disambiguate their meaning?

• How does context influence the perception and interpretation of gaze? A sudden shift of gaze is likely to mean one thing in a conversation but something entirely different during a tennis match. More locally, the context provided by a verbal utterance during a conversation might also

13

influence how gaze, head and gesture are interpreted, something children seem able to understand when learning new words[47]. The context provided by the facial expression might also be important. A direct gaze coupled with an angry expression probably means something entirely different to a similar gaze married to a smile.

• What is the relationship between the social function of gaze and its primary cognitive function for the gazer - to foveate and hence analyse in depth a region of the visual world? One recent theory elaborates the role of gaze as a deictic, focussing processing agendas selectively and serially[48]. Can we develop this "active vision" perspective and consider how joint attention, achieved by the perception of another person's gaze leads to the coordination of different individual processing agendas?

• How do we avoid processing overloads when social signals compete with other cognitive processing? Some recent research suggests that concurrent social signals can interfere with ongoing complex cognitive tasks, such as solving difficult mental arithmetic problems or remembering hard-to-retrieve items[49]. Children may suffer interference from visual social signals in some circumstances, where adults would avert their gaze to avoid such conflicts (Doherty-Sneddon et al., unpublished study).

Acknowledgements

15

References

1 Bruce, V. *et al*. (1992) *Processing the facial image*, Oxford University Press

2 Young, A.W. (1998) *Face and mind*, Oxford University Press

3 Bruce, V. and Young, A. (1986) Understanding face recognition. *British Journal of Psychology* 77, 305-327

4 Young, A.W. *et al*. (1993) Face perception after brain injury: Selective impairments affecting identity and expression. *Brain* 116, 941-959

5 Walker, S., Bruce, V. and O'Malley, C. (1995) Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception and Psychophysics* 57, 1124-1133

6 Schweinberger, S.R. and Soukup, G.R. (1998) Asymmetric relationships among perceptions of face identity, emotion and facial speech. *Journal of Experimental Psychology: Human Perception and Performance* 24, 1748-1765

7 Kleinke, C.L. (1986) Gaze and eye contact: A research review. *Psychological Review* 100, 78-100

8 Perrett, D.I. and Emery, N.J. (1994) Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive* 13, 683-694

9 Perrett, D.I. *et al*. (1994) Organisation and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B* 335, 23-30

10 Baron-Cohen, S. *et al*. (1992) How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive* 13, 513-552

11 Baron-Cohen, S. (1995) *Mindblindness: An essay on autism and theory of mind*, MIT Press/Bradford Books.

16

12 Maurer, D. (1985) Infants' perception of facedness, in *Social perception in infants* (Field, T. and Fox, N. eds.), Ablex.

13 Vecera, S. P., & Johnson, M. H. (1995) Gaze detection and the cortical processing of faces: Evidence from infants and adults. *Visual Cognition* 2, 59-87

14 Von Grünau, M., & Anston, C. (1995) The detection of gaze: A stare in the crown effect. *Perception* 24, 1297-1313

15 Fodor, J. (1983) *The modularity of mind: An essay on faculty psychology*, MIT Press

16 Coltheart, M. (1999). Modularity and cognition. *Trends in Cognitive Sciences* 3, 115-120

17 Hood, B.M., Willen, J.D. and Driver, J. (1998) An eye direction detector triggers shifts of visual attention in human infants. *Psychological Science* 9, 131-134

18 Driver, J. *et al*. (1999) Shared attention and the social brain: Gaze perception triggers automatic visuospatial orienting in adults. *Visual Cognition* 6, 509-540

19 Friesen, C. K. & Kingstone, A. (1998) The eyes have it!: reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review* 5, 490 - 495

20 Langton, S.R.H. and Bruce, V. (1999) Reflexive visual orienting in response to the social attention of others. *Visual Cognition* 6, 541-568

21 Posner, M.I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology* 32, 3-25

22 Jonides, J. (1981) Voluntary versus automatic control over the mind's eye's movement, in *Attention and Performance IX* (Field, T. and Fox, N. eds.), pp. 187-203, Lawrence Erlbaum Associates

23 Perrett, D.I., *et al*. (1985) Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London* B223, 293-317

24 Heywood, C.A. and Cowey, A. (1992) The role of the 'face cell' area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society of London, Series B* 335, 31-38

25 Campbell, R., *et al*. (1990) Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia* 28, 1123-1142

26 Wollaston, W.H. (1824). On the apparent direction of eyes in a portrait. *Philosophical Transactions of the Royal Society, London*, 247-256. Cited in Bruce, V. and Young, A. (1998) *In the eye of the Beholder: The science of face perception*, Oxford University Press

27 Cline, M.G. (1967) The perception of where a person is looking. *American Journal of Psychology* 80, 41-50

28 Maruyama, K., & Endo, M. (1983) The effect of face orientation upon apparent direction of gaze. *Tohoku Psychologica Folia* 42, 126-138

29 Anstis, S.M., Mayhew, J.W. and Morley, T. (1969) The perception of where a face or television 'portrait' is looking. *American Journal of Psychology* 82, 474-489

30 Gibson, J.J. and Pick, A. (1963) Perception of another person's looking. *American Journal of Psychology* 76, 86-94

31 Langton, S.R.H. The mutual influence of gaze and head orientation in the analysis of social attention direction. *Quarterly Journal of Experimental Psychology* (in press)

32 Langton, S. R. H., & Bruce, V. You *must* see the point: Automatic processing of cues to the direction of social attention. *Journal of Experimental Psychology: Human Perception and Performance* (in press)

33 Langton, S.R.H., O'Malley, C. and Bruce, V. (1996) Actions speak louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural

information. *Journal of Experimental Psychology: Human Perception and Performance* 22. 1357-1375

34 Scaife, M. and Bruner, J.S. (1975) The capacity for joint visual attention in the infant. *Nature* 253, 265-266

35 Butterworth, G., and Jarrett, N. (1991) What minds have in common is space: Spatial mechansims serving joint visual attention in infancy. *British Journal of Developmental Psychology* 9, 55-72

36 Moore, C. and Corkum, V. (1998) Infant gaze following based on eye direction. *British Journal of Psychology* 16, 495 - 503

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

38 Baron-Cohen, S., *et al*. (1995) Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology* 13, 379-398

39 Anderson, J.R., and Doherty, M.J., (1997). Preschoolers' perception of other people's looking: Photographs and drawings. *Perception* 26, 333-343

40 Doherty, M.J. and Anderson, J.R. A new look at gaze: Preschool children's understanding of eye-direction. *Cognitive Development*. (In press)

41 Itakura, S. and Anderson, J.R. (1996) Learning to use experimenter-given sues during an object-choice task by a capuchin monkey. *Current Psychology of Cognition* 15, 103-112

42 Itakura, S. (1996) An exploratory study of gaze-monitoring in non-human primates. *Japanese Psychological Research* 38, 174-180

43 Emery, N. J. *et al*. (1997) Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*). *Journal of Comparative Psychology* 111, 286-293

44 Itakura, S. and Tanaka, M. (1998) Use of experimenter-given cues during object-choice tasks by chimpanzees (*Pan troglodytes*), an orangutan (*Pongo pygmaeus*), and human infants (*Homo-sapiens*). *Journal of Comparative Psychology* 112, 119 - 126

45 Povinelli, D.J. and Eddy, T.J. (1996) Chimpanzees: Joint visual attention. *Psychological Science* 7, 129- 135

46 Kobayashi, H. and Kohshima, S. (1997) Unique morphology of the human eye. *Nature* 387, 767 - 768

47 Baldwin, D.A. (1995) Understanding the link between joint attention and language, in *Joint attention: Its origins and role in development* (Moore, C. and Dunham, P.J. eds.), pp. 131-158. Lawrence Erlbaum Associates

48 Ballard, D.H. *et al*. (1997) Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20, 723-???

49 Glenberg, A.M., Schroeder, J.L., Robertson, D.A. (1998) Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition* 26, 651-658

Box 1. The Mindreading System

Humans and the vast majority of primate species are social animals, living in groups comprising as many as 200 individuals. Thriving in such an environment requires a particular kind of "social intelligence"; an ability to make sense of another individual's actions and crucially, to predict what they are about to do next. Several authors have argued that we humans are able to do this because we have evolved the ability to read the behaviours of others in terms of mental states such as knowing and believing[a,b]. Each of us acts on what we *know* is true, *believe* to be true or sometimes *pretend* to be true about the world. Baron-Cohen[c,d] has proposed the existence of a "mindreading" system which functions to make these mental state attributions to other agents. When fully developed it comprises four components (Fig.1); the Intentionality Detector (ID), the Eye Direction Detector (EDD), the Shared Attention Mechanism (SAM), and the Theory of Mind Mechanism (ToMM). Each is considered to be a cognitive "module" sharing many, though not all, of the properties of modularity described by Fodor in his influential work[e]

The ID, according to Baron-Cohen, is a primitive perceptual mechanism that interprets self-propelled motion stimuli in terms of its desires and goals. For instance, it is this mechanism which allows us to infer that a cat chasing a mouse "wants" to eat the mouse. The second mechanism is the EDD which has three basic functions. It detects the presence of eyes or eye-like stimuli, it computes the direction of gaze based on the position of the iris in the surrounding sclera, and finally it attributes the mental state of "seeing" to an agent whose eyes are directed toward itself or toward another object or agent. Thus, by the age of about 9 months when the ID and EDD are considered to be fully functioning, an infant is able to read another individual's behaviour in terms of their goals and desires, and understands that these individuals "see" the things to which their eyes are directed. What the infant cannot do at this stage is link the two mechanisms, that is, understand that people often look at the things they want or are about to act on. This feat is achieved by the SAM which is fully developed between 9 and 18 months. Although it serves to link the ID and the EDD, the SAM's main function is to identify when the self and another are attending to the same thing. Essentially it uses information from the EDD -

that another individual is looking at, say, a bus - and compares this with the self's current perceptual state. If the two match, visual attention is shared. The SAM therefore enables its possessor to engage in a "meeting of minds", the recognition that you and another are sharing the same mental state - in this case that of "attending to", "seeing", "wanting" or the state of having a particular goal. Baron-Cohen suggests that this primitive meeting of minds triggers, from between 18 and 48 months, the development of the final module in the mindreading system, the ToMM. The ToMM has two major functions. First it is able to infer the full range of mental states from observable behaviour. These include pretending, thinking, knowing, believing, imagining, and deceiving. Second, the ToMM is able to integrate this mental state knowledge into a useable theory which the child or adult can use to explain and predict other's behaviour.

Baron-Cohen places particular emphasis on the EDD in his model, and in particular its links with the SAM. He maintains that the ability to detect eyes and eye direction, and thence to use gaze to figure out another's mental state is of extreme importance in mindreading. Although mental states can be inferred from other modalities, the eyes, he claims, are the best and most immediate "windows to the mind", and also the best indicators that we have "connected" with another mind when engaging in joint attention.

References

a Premack, D. and Woodruff, D. (1978) Does the chimpanzee have a 'theory of mind'? *Behavioural and Brain Sciences* 1, 515 - 526

b Humphrey, N. (1984) *Consciousness regained*, Oxford University Press

c Baron-Cohen, S. (1994) How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive* 13, 513-552

d Baron-Cohen, S. (1995) *Mindblindness: An essay on autism and theory of mind*, MIT Press/Bradford Books.

e Fodor, J. (1983) *The modularity of mind: An essay on faculty psychology*, MIT Press

Box 2: The Eye as Visual Stimulus

In what sense might there be something special or dedicated about the processing system that leads from the retinal image to a representation of someone's gaze direction? The step that brings about the explicit representation of eye direction must be dedicated. However, in getting to that point, perhaps the system employs specialised neural pathways that otherwise do nothing; or perhaps the system manages to use neurones that already exist for other purposes right up to the input to that very final stage. Does the eye get its special effect because it is an ordinary sort of stimulus treated by a specialised process or because the eye is an especially suitable kind of stimulus for the purpose of communicating eye direction processed in the ordinary way?

It can be easily shown that the eye, as a visual stimulus, has a number of simple and potentially powerful features[a].. Figure 1 shows images of an eye. Superimposed on each is a pattern which shows the spatial variation in the amplitude of the response of vertically oriented simple cells from striate cortex. As can be seen, the cells respond vigorously over the whole of the eye. The response is in three spatially separate parts: one to each of the two visible parts of the sclera and one to the iris/pupil. As the eye turns, the response to the two scleral parts alter in their relative strength (in proportion to the respective areas). Thus, the contrast of the response of the two scleral parts is a monotonic function of eye direction. Eye direction is therefore a particularly simple measurement to perform on an image of the eye. The reason for this lies in the form of the eye and its interaction with the functional properties of cortical simple cells.

Scleral contrast actually computes something that is hybrid between absolute eye direction in space and eye direction relative to head direction. If the gap between the eyelids (the palpebral fissure) were planar and oriented in the fronto-parallel plane, then sclera contrast would measure eye direction entirely relative to head direction. However, this gap actually curves around the eyeball. As a result, scleral contrast measures eye direction entirely absolute in space from any viewing directions where the corner of the eye (the lateral canthus) is out of sight round the eyeball. In fact, the gap between the eyelids extends for about 130° with the corner of the eye being about 75° away from straight ahead. Therefore, if treated as if it were the measure of

absolute eye direction, scleral contrast would lead to small errors in judging eye direction for views of a person when they were not facing you, errors that increased as the head angle turned away from you, exactly as have been found[b].

All of this leads to the suggestion that the eye is a special stimulus in the sense that useful information can be recovered from it with robust simple processing mechanisms. The implication of this is that the processing system involved could be correspondingly un-special.

References

a Watt, R.J. (1999). What your eyes tell my eyes, and how your eyebrows try to stop them. *Paper presented at Tenth International Conference on Perception and Action - ICPA-X, University of Edinburgh, August 1999*

b Anstis, S.M., Mayhew, J.W. & Morley, T. (1969) The perception of where a face or television 'portrait' is looking. *American Journal of Psychology* 82, 474-489

Box 3: Social Directional Cues Trigger Reflexive Shifts of Attention

Most of us have experienced the tendency to look where others are looking. In the middle of your next conversation, for instance, suddenly shift your gaze, or turn your head to look at something, and observe your conversational partner's behaviour. Anecdotally then, there appears to be some suggestion that shifts in another's line of regard might trigger shifts in an observer's visual attention. In fact, recent studies by three independent groups[a-c] have provided some empirical evidence to support this claim. All three groups adapted the cueing paradigm devised by Posner[d] to demonstrate that socially relevant cues such as eye-gaze direction and head orientation trigger reflexive shifts of a viewer's visual attention.

In the study conducted by Langton & Bruce[c], participants were asked to press the space bar on a keyboard as soon as they detected a target letter which could appear at one of four locations on a computer monitor. Either 100 ms or 1000 ms prior to the appearance of the target, a face appeared in the centre of the screen that was oriented toward one of the possible target locations (see Figure 1). Targets could therefore appear in either cued or uncued locations. Participants were told (correctly) that following the appearance of a face, the target letter was equally likely to appear in any of the four possible locations. In other words, the cue was completely uninformative regarding the likely location of the target and therefore should be ignored. However, the results indicated that participants were not able to comply with these instructions. At the shorter, 100 ms cue-target interval, detection times were faster for targets appearing in cued locations than for those appearing in uncued locations. However, this cueing effect had vanished within 1000 ms of the presentation of the face cue (see Figure 2).

On the basis of this pattern of results, Langton and Bruce concluded that the face cues triggered a kind of reflexive or exogenous shift of visual attention that is normally associated with a change in luminance, or the abrupt onset of a stimulus in the periphery of vision. Identical conclusions were reached by Driver et al.[a] and Friesen and Kingstone[b] who obtained broadly similar results using eye-gaze direction as their cueing stimuli.

However, there are reasons to believe that the shifting of attention in response to head and gaze direction represents a rather special form of reflexive orienting. First, directional cues such as arrows do not trigger reflexive shifts of attention[e]. Second, recent evidence suggests that the neural circuitry subserving reflexive shifts of attention in response to socially irrelevant stimuli (e.g., abrupt onsets and luminance changes) is different from that involved in the orienting response triggered by gaze cues[f]. The latter involves lateralised cortical pathways, whilst the former depends on subcortical pathways that are shared between the cerebral hemispheres.

References

a Driver, J. *et al*. (1999) Shared attention and the social brain: Gaze perception triggers automatic visuospatial orienting in adults. *Visual Cognition* 6, 509-540

b Friesen, C.K. and Kingstone, A. (1998) The eyes have it!: reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review* 5, 490 - 495

c Langton, S.R.H., & Bruce, V. (1999) Reflexive visual orienting in response to the social attention of others. *Visual Cognition* 6, 541-568

d Posner, M.I. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology* 32, 3-25

e Jonides, J. (1981) Voluntary versus automatic control over the mind's eye's movement, in *Attention and Performance IX* (Field, T. and Fox, N. eds.), pp. 187-203, Lawrence Erlbaum Associates

f Kingstone, A., Friesen, C.K. and Gazzaniga, M.S. Reflexive joint attention depends on lateralized cortical connections. *Psychological Science* (in press)

Box 4: The Direction of Attention Detector

Perrett and his colleagues[a,b] have suggested that something like an Eye Direction Detector (see Box 1) forms only part of a system designed to compute the direction of social attention. Their single cell studies have indicated that individual cells in the STS region of the macaque temporal lobe are sensitive to *conjunctions* of eye, head and body position. For instance, those cells which are particularly active when presented with a pair of eyes looking downwards also respond strongly when heads are directed downwards or when the body adopts a quadrupedal posture. Accordingly, they postulate the existence of a Direction of Attention Detector (DAD) which combines information from separate detectors analysing the direction of the eyes, head and body. However, how does the system respond when, say, the eyes might be looking downwards whilst the head is angled slightly upwards? Perrett's group have suggested that the DAD is organised such that information from the eyes will override any information provided by the head, and in turn, information provided by the head can override directional signals from the body. This is achieved by a network of inhibitory connections. Information from the eyes can directly inhibit cells coding an inappropriate head direction but not vice-versa, and similarly information specifying a particular head angle can inhibit cells coding an inappropriate body position but not vice versa. To return to our example, if the eyes are visible and are looking downwards whilst the head is directed upwards, the inhibitory connections will ensure that the input to the STS cells is restricted to information provided by the eyes and the direction of attention will therefore be coded as downwards (see Fig. 1). Social attention can also be computed under a variety of viewing conditions. For instance, if the face is viewed at a distance, or if the eyes are obscured by shadow, the system defaults to signalling the direction of attention from the orientation of the head, or if this too is obscured, from the orientation of the body.

References

a Perrett, D.I. and Emery, N.J. (1994) Understanding the intentions of others from visual signals: Neurophysiological evidence. *Cahiers de Psychologie Cognitive* 13, 683-694

b Perrett, D.I. *et al*. (1992) Organisation and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B* 335, 23-30
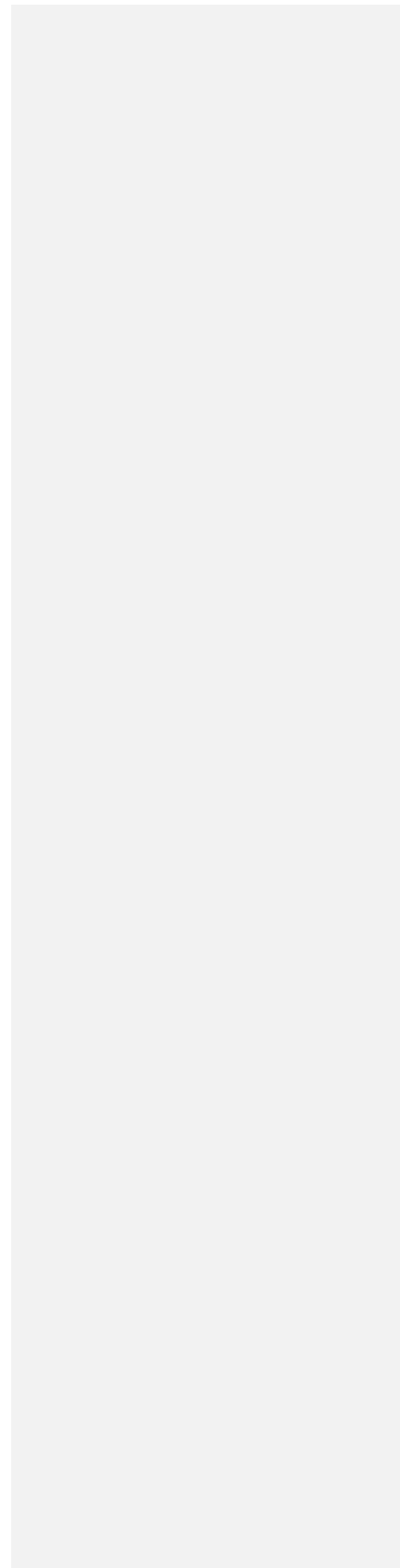
Fig. 1. Head orientation influences the perceived direction of gaze. The top two pictures are taken from Wollaston's original paper. Face B seems to be gazing directly at the viewer whereas Face A appears to be looking slightly to the viewer's right. By covering the lower and upper parts of each face you can see that the eye regions of both are, in fact, identical. The lower two faces illustrate a similar effect with greyscale images. The eye region from D has been pasted onto C where the head is rotated slightly to the viewer's left. Each of fifteen people shown Face D decided the eyes were looking straight ahead rather than to the left. However, thirteen of a further fifteen people shown Face C decided the identical pair of eyes were actually looking towards their left (Langton, unpublished data).
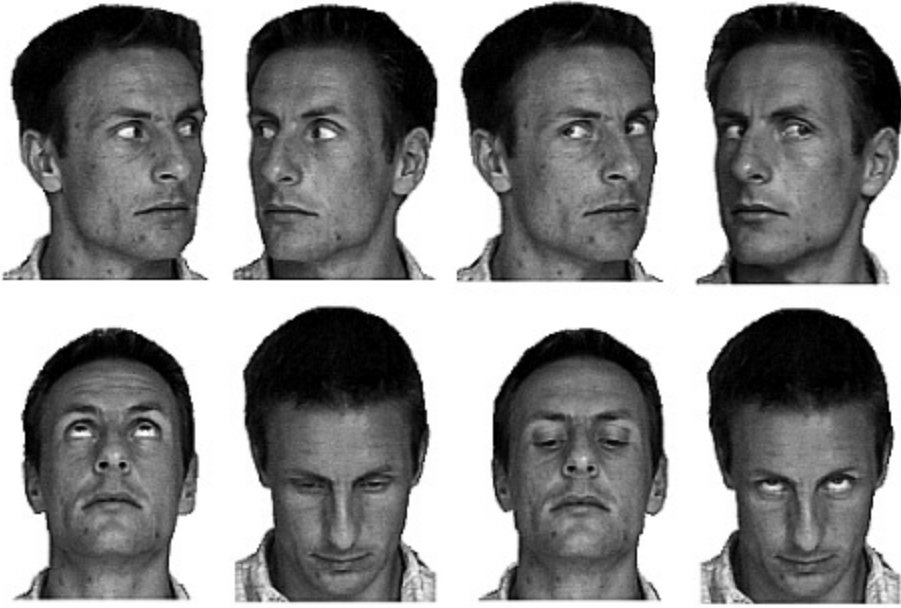
Fig. 2. The stimuli used to examine the mutual influence of gaze and head orientation in the processing of social attention direction.
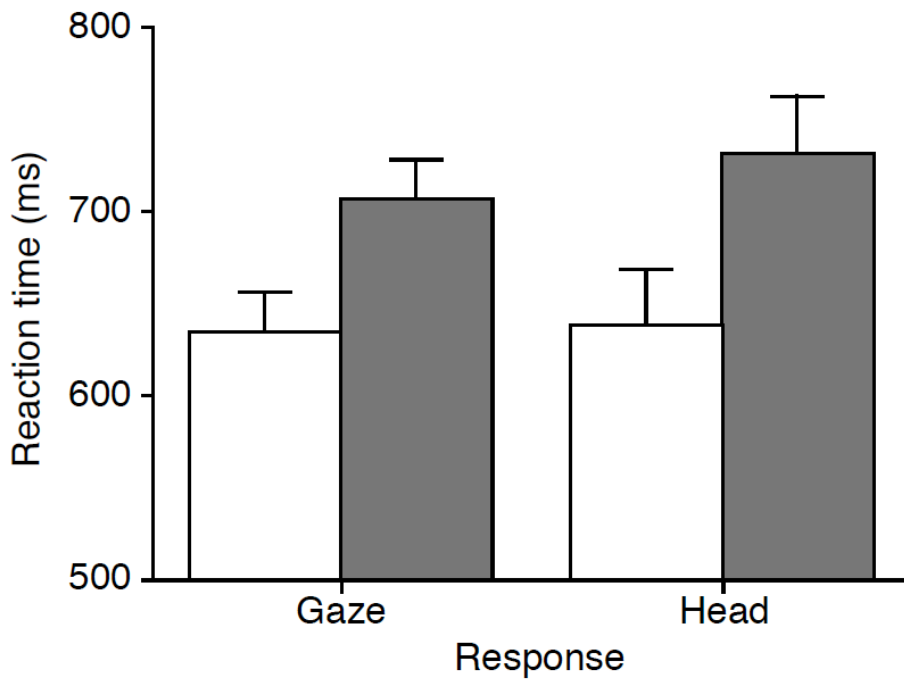
Fig. 3. Time (in milliseconds) to respond to the direction of the gaze, and the orientation of the heads of the images shown in Figure 2. Reaction times to gaze directions were affected by the congruity of the head orientation and reciprocally, the time taken to respond to the head orientation was equally affected by the congruity of the gaze direction. (Data drawn from Ref. 31)
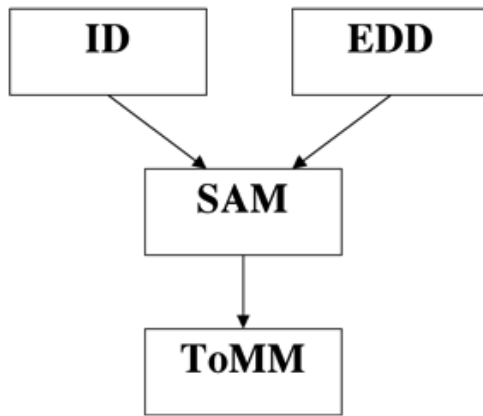
Box 1, Fig.1. The relationship between the four components of Baron-Cohen's Mindreading model. The figure was adapted from one used by Baron-Cohen[c] .
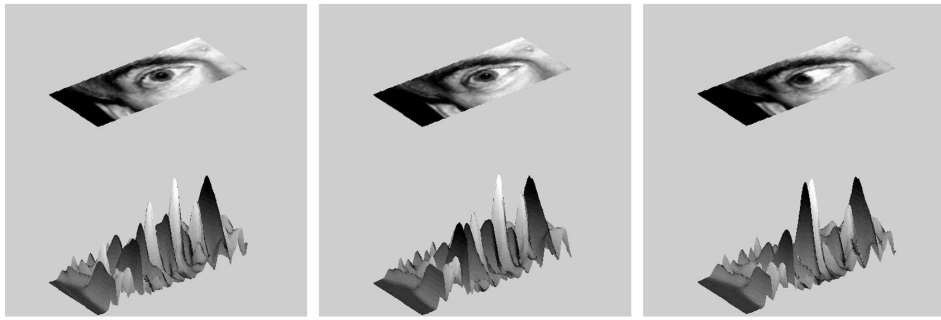
Box 2, Fig. 1. The response of cortical simple cells to eye direction. When the eye is looking straight ahead (A), the output of the cells responding to the area of sclera on either side of the eye is roughly equivalent (the two white peaks shown in the image above the eye). As the eye begins to turn to the left (B and C), the area of sclera to the right of the iris increases relative to the area to the left of the iris. The relative strength of the cells' outputs corresponds to this change. This can be seen as one of the white peaks increases its height from B to C whilst the other peak's height is reduced.
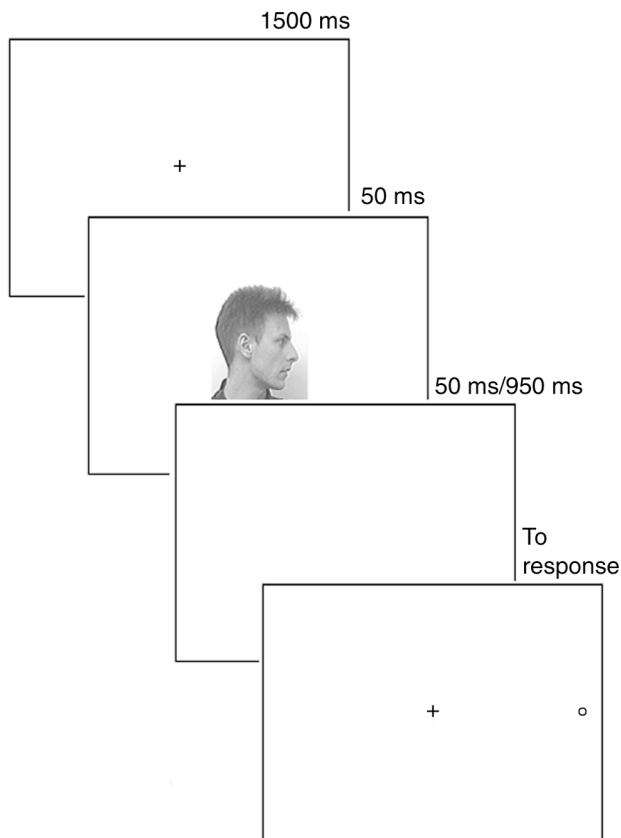
Box 3, Fig.1. The sequence of events in Langton & Bruce's precueing experiment. A fixation cross was presented for 1500 ms, followed by the appearance of a face cue which remained on the screen for 50 ms. The cueing faces were either looking upwards, downwards, to the left, or to the right. The target display was then presented either 50 ms or 950 ms after the disappearance of the cue. The time between the onset of the cues and the onset of the targets - the Stimulus Onset Asynchronies (SOAs) - were therefore 100 ms and 1000 ms respectively. Participants were asked to keep their eyes fixed in the centre of the screen and press the space bar as soon as they detected the target letter "o" regardless of where it appeared on the screen. The target display remained on the screen until participants had made their response. The screen then went blank and remained so for 1000 ms before the beginning of the next trial.
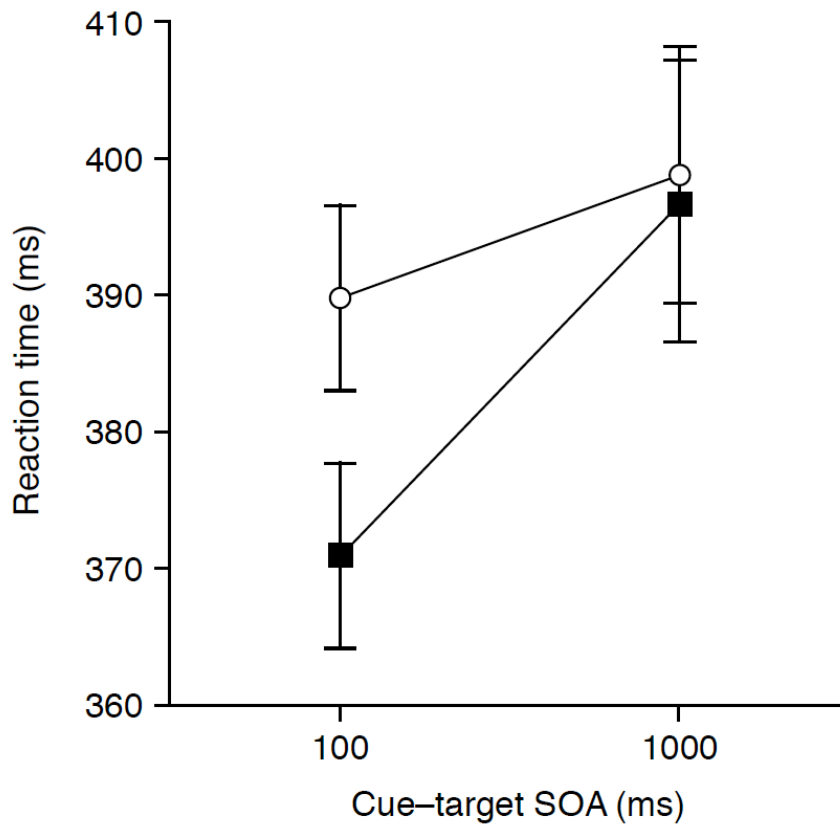
Box 3, Fig. 2. Detection times (in milliseconds) for cued and uncued targets as a function of cue - target stimulus onset asynchrony (SOA). The plot shows that at 100 ms SOAs, participants were faster to detect targets appearing in cued locations compared to uncued locations, but that this effect was absent at 1000 ms SOAs. (Data drawn from Ref. c.)
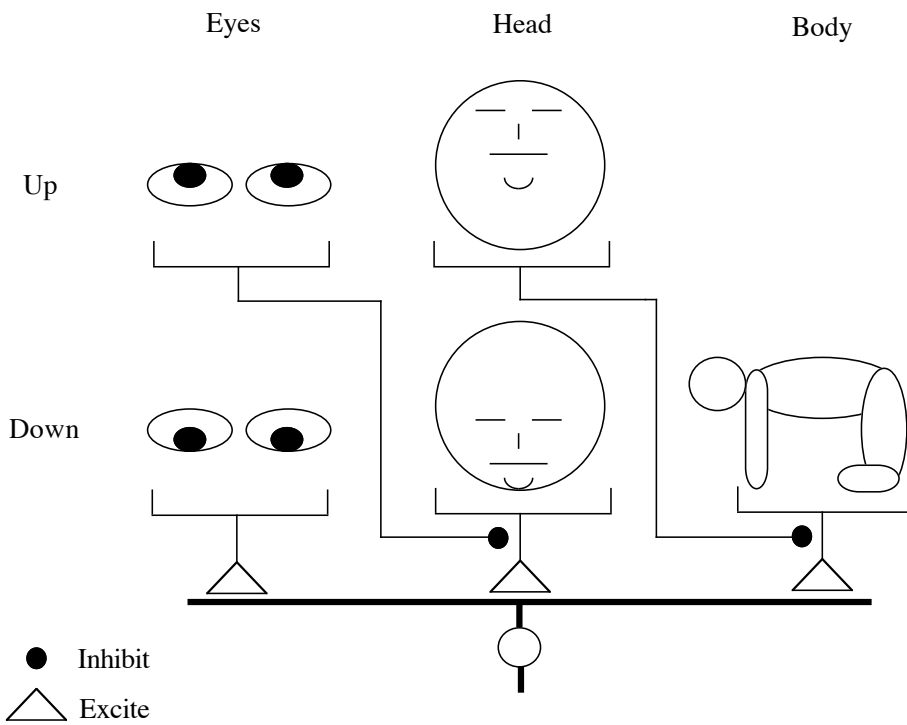
Box 4, Fig. 1. The Direction of Attention Detector. This figure depicts a schematic representation of the connections and visual input to an STS cell (the large circle) signalling that another's attention is directed downwards. The cell receives excitatory connections from cells selective for the appearance of eyes, head and body directed downwards. Should the gaze be directed upwards, inhibitory connections prevent any response to the downward directed head and body cues. The figure was adapted from one used by Perrett et al.[b]