Don't be fooled! Attentional responses to social cues in a face-to-face and video magic trick reveals greater top-down control for overt than covert attention

Gustav Kuhn a,⇑, Robert Teszka a, Natalia Tenaw b, Alan Kingstone c

a Goldsmiths, University of London, UK
b Brunel University, UK
c University of British Columbia, Canada

A R T I C L E   I N F O

Article history:
Received 17 November 2014
Revised 4 May 2015
Accepted 11 August 2015

Keywords:
Attention
Eye movements
Social cognition
Social attention
Misdirection
Top-down control
Inhibition

A B S T R A C T

People's attention is oriented towards faces, but the extent to which these social attention effects are under top down control is more ambiguous. Our first aim was to measure and compare, in real life and in the lab, people's top-down control over overt and covert shifts in reflexive social attention to the face of another. We employed a magic trick in which the magician used social cues (i.e. asking a question whilst establishing eye contact) to misdirect attention towards his face, and thus preventing participants from noticing a visible colour change to a playing card. Our results show that overall people spend more time looking at the magician's face when he is seen on video than in reality. Additionally, although most participants looked at the magician's face when misdirected, this tendency to look at the face was modulated by instruction (i.e., “keep your attention on the cards”), and therefore, by top down control. Moreover, while the card’s colour change was fully visible, the majority of participants failed to notice the change, and critically, change detection (our measure of covert attention) was not affected by where people looked (overt attention). We conclude that there is a tendency to shift overt and covert attention reflexively to faces, but that people exert more top down control over this overt shift in attention. These finding are discussed within a new framework that focuses on the role of eye movements as an attentional process as well as a form of non-verbal communication.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

What we see is strongly influenced by what we attend to, and it is a truism that visual attention is controlled in two basic ways. One is exogenously, through bottom-up stimulation from the external world (Itti & Koch, 2001), and the other is endogenously, through top-down internally generated intentions (e.g., Folk, Remington, & Johnston, 1992). Traditionally, these two forms of attentional control have been investigated using simple stimuli, such as light flashes or computer beeps (Posner, 1980). However, more recently, biologically meaningful stimuli have been used as they seem to be prioritized by the attention system, and are governed by principles that were not necessarily captured by those generated from studies using simpler nonsocial stimuli (Birmingham, Bischof, & Kingstone, 2009; Kingstone, Smilek, & Eastwood, 2008).

To date, however, the bulk of the research studies with social stimuli have themselves been conducted in a fairly socially impoverished manner, with the standard experiment framed around a single participant sitting alone in a testing room looking at social stimuli, e.g., pictures of people. What these studies have revealed is that people prioritize faces, especially the eyes (Birmingham, Bischof, & Kingstone, 2008) and that people reflexively attend to where the eyes of a face are looking (Friesen & Kingstone, 1998). Moreover, it is clear that all faces are not treated equally, with people prioritizing faces on video that are speaking (Foulsham & Sanderson, 2013), suggesting that the effect of social stimuli on visual attention extends to dynamic auditory stimulation. Nevertheless, there do appear to be some important limitations to these lab-based investigations, the most notable one being that the results derived using images of social stimuli often fail to extend to real life situations composed of people. For example, Laidlaw, Foulsham, Kuhn, and Kingstone (2011) have shown that when a stranger is in a room people are far less likely to look at that person than if that person was presented on a computer screen, despite the fact that the image of the person on the computer was far smaller.

⇑ Corresponding author at: Department of Psychology, Goldsmiths, University of London, New Cross, London SE14 6NW, UK.
E-mail address: G.Kuhn@gold.ac.uk (G. Kuhn).

http://dx.doi.org/10.1016/j.cognition.2015.08.005
0010-0277/ © 2015 Elsevier B.V. All rights reserved.
and less salient than that of a real person. Similarly, Gallup et al. (2012) have demonstrated that in real life people tend to avoid looking where someone else is looking (unless they can do it discreetly), which conflicts with lab-studies reporting that people automatically attend to where eyes in images are directed. In other words, the relationship between gaze direction and attentional shifts is far more complex in real life than lab-studies would suggest, a point driven home recently by Wu, Bischof, and Kingstone (2014, 2013) who revealed that when eating with another person, people use their eyes to communicate when they are about to take a bite of their food causing the other person to look away.

One overarching principle to emerge from these comparisons between life and lab is that in every day life people exert top-down control over attention in a manner that is often divergent to what has been observed in the lab. Our working hypothesis for why this disconnection exists is that in real life one's own eyes are used both to observe people and to communicate to them, just as their eyes are used to observe and signal to you and others. This dual function of one's eyes – observation and communication – is absent when one is simply looking at images of individuals. Because images of people neither observe one's gaze nor communicate back, one's own eyes merely serve to observe and do not communicate to the image (Wu et al., 2014). Thus, in the lab it is perfectly acceptable to stare at the eyes of a stranger's image, but in real life, this would be abnormal behaviour (Kingstone, 2009).

Accordingly, in the present study we were very sensitive to the fact that the manifestation of social attention may change dramatically as one shifts between lab and life, and therefore we chose a task – a magic trick to be precise – that we had good reason to believe would engage social attention in a similar manner when the magician was live as when he and the trick were shown on video. Kuhn and Land (2006), Kuhn, Tatler, and Cole (2009) have demonstrated repeatedly that a magician's trick that depends on social cues to misdirect attention is successful whether it is performed live or recorded and played back on video.

To date, measures of real world attention have been restricted exclusively to overt forms of attentional orienting, i.e., shifts in head and eye movements. While at first blush this seems reasonable, as people tend to look at what they attend to, it is well established that people can attend covertly to objects that are positioned at locations away from where their eyes are directed (for review see Smith & Schenk, 2012). What role covert attention plays in natural real world social attention is very much an open question and a crucial one that the present study investigates.

Magicians use misdirection to prevent people from detecting their secret methods (Kuhn, Caffaratti, Teszka, & Rensink, 2014), and many of these misdirection techniques involve manipulating both overt and covert attentional processes (for review see Kuhn & Martinez, 2012). For example, misdirection is not only effective in manipulating where people look, but it is also extremely effective at preventing people from perceiving visually salient events, which in turn provides a valuable index of covert attentional orienting (see Kuhn & Findlay, 2010). Of critical relevance to the present study, it has been demonstrated that misdirection can be used to study attention in the real world (Kuhn & Tatler, 2005), as well as in the lab (Kuhn, Tatler, Findlay, & Cole, 2008), which makes it an ideal tool to compare attentional processes in these different contexts.

Magicians use a wide range of social cues to misdirect attention (Kuhn et al., 2014). For example, directional gaze cues effectively orient covert and covert attention towards a looked at location (Cui, Otero-Millan, Macknik, King, & Martinez-Conde, 2011; Kuhn & Land, 2006; Kuhn et al., 2009). Here we investigate people's reflexive tendency to look at faces by exploiting a powerful social misdirection technique frequently used by magicians. Magicians often draw the spectator's attention towards their face by asking them a question whilst establishing eye contact. Amongst magicians it is commonly accepted that if you ask someone a question, that person will naturally look at your face (Tamariz, 2007). In other words, asking a person a direct question is a social cue that will trigger reflexive social orienting to the face.

In the present study, participants watched a magic trick, either live or on video, in which a magician used social misdirection (a question) to prevent observers from detecting a visually salient colour change. Change blindness is a term for a phenomenon whereby changes to an unattended item go undetected (Rensink, O'Regan, & Clark, 1997). Overt fixation is not sufficient for change detection because covert attention may be allocated elsewhere (e.g. Mack & Rock, 1998; Smith, Lamont, & Henderson, 2012). Whilst some changes can be missed when attended (Rensink, 2000), attention is necessary to notice changes to items, and therefore change detection provides a valuable index of attentional mechanisms that are independent of eye movement (i.e. covert attention).

Previous research indicates that people look at real people less than images (Laidlaw, Risko, & Kingstone, 2012), whilst others have found no difference between the two (Freeth, Foulsham, & Kingstone, 2013). We therefore further explored the conditions under which fixations to the face vary as a function of viewing condition (i.e. live vs. video). In order to assess the reflexivity of attentional orienting, we modulated participants' direct top-down attentional control by instructing half of the participants to avoid being distracted from the card trick. We expected that participants would be able to exert some degree of top-down control over the reflexive tendency to look at the magician's face when they were posed a question (e.g., Laidlaw et al., 2012). However, and of critical importance for the present study, how this instructed top-down control would vary as a function of context (live vs. video) and type of attention (overt vs. covert) was far from clear based on past work. Addressing these two issues were the focus of the present paper.

Unlike in the video context, in the live context the actor can see the participant, and thus there is scope for real social interactions. With regard to context (live vs. video), one prediction is that participants will be able to exert less instructed top-down control when the magician's social distracting question is presented live than in the video because live situations are more social (Freeth et al., 2013). Alternatively, as people are more inclined to look at faces on video than in live situations, instructed top-down inhibition may be less effective for video than live questioning.

As for the effect of social distraction on covert orienting, there are two clear-cut alternatives, derived from the fact that overt and covert orienting are linked but separable (Smith & Schenk, 2012). If overt and covert attention are always linked the overt orienting will be mirrored by covert orienting. On the other hand, if the two forms of orienting are separable, it is possible that the effect of top-down control and context will be very different for overt and covert attention.

In sum, the aims of the present study were twofold. First, we aimed to measure and compare in real life and in the lab people's top-down control over overt shifts in reflexive social attention to the face of another. Second, we chose a task that we anticipated would, qualitatively speaking, behave similarly in real life and in the lab, thereby enabling us to investigate how the presence or absence of a real person may modulate the control of reflexive shifts of covert versus overt attention. Finally, it is worth noting at the outset that in the present study our focus is on the functional relationship between overt and covert orienting, and as such we are agnostic as to whether overt and covert attentional orienting are driven by independent attentional mechanisms (e.g. Hunt & Kingstone, 2003), or a single mechanism in which covert attention
is conceived as a planned, but not yet executed eye movement (e.g. Zirnsak & Moore, 2014).

2. Method

2.1. Participants

One hundred and twelve undergraduate students at Brunel University participated, with equal numbers (28) in each of the four conditions, all of whom were reimbursed with chocolate. Data from 8 participants were excluded as they reported having seen the same trick on a previous occasion (7 live condition, 1 video condition). The study received departmental ethical approval.

2.2. Materials and procedure

The magic trick: The misdirection was embedded within the context of the classic “Princess card trick” (Downs, 1909) in which a thought-of card vanishes. We measured the subjective effectiveness of the misdirection by adding an additional component to the trick whereby the backs of the cards changed colour (from blue to red; see Fig. 1). This colour change took place in full view, and is fully visible when attended to.

All participants were told that they were about to see a magic trick and that their task was to find out how the trick was done. Participants in the misdirection warning condition were instructed: “As you may know magicians often use misdirection to distract your attention. It is therefore very important that you try not to be distracted and always keep your eyes on the cards!”.

We predicted that if instructed top-down control can modulate whether a person’s attention is misdirected by the social cue in the magic trick, then participants’ attention in the warned condition should be misdirected less.

In the face-to-face condition, participants were seated opposite the magician (first author), and watched the magic trick at a distance of approximately 1 m. The performer was blind to which instructions the participants had received. In the video condition, a video clip of the same magic trick was displayed on a computer monitor. Here the magician asked the question directly to the camera. Immediately after the magic trick, participants were asked to fill out a short questionnaire assessing if they had detected the back of the cards change colour.

2.2.1. Eye tracking

In the face-to-face condition participants’ eye movements were recorded using an Arrington Research monocular SceneCamera mobile eye-tracker with sampling rate set to 60 Hz. After calibration, the ViewPoint software (Arrington Research Inc., 2006) superimposed participants’ eye positions on the scene camera’s 30 fps video output, which was used for all subsequent analysis. Participants’ eye-movements were calibrated using a 9-point calibration in ViewPoint software (Arrington Research Inc., 2006).

In the video condition, eye movements were recorded using an Eyelink 1000 eye tracker (SR-Research), and the video (574 by 759) was presented on a 21in CRT monitor (600 x 800; 75 Hz). Prior to the experiment each participant completed a 9-point calibration and eye movements were sampled at 500 Hz. All analyses were conducted by exporting a video file (30 fps), which included the gaze position superimposed on the video. The high sampling resolution of the Eyelink eye tracker was therefore

Fig. 1. (1) Blue cards are taken out of blue box (the box is dropped on floor out of sight) and the cards are shuffled. Six cards are counted and the faces are displayed to the participant (2). Participant is asked to remember one. (3) The cards are closed into a small pack which the magician looks down at. (4) Magician uses misdirection by asking the participant whether he/she remembers the card whilst establishing eye contact, at same time the backs of the cards is changed from blue to red using sleight of hand. (5) Magician looks back at the cards. Five red-backed cards counted to show one vanished. (6) Faces then revealed to show participant’s thought-of card disappeared. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
down-sampled to 30 Hz resulting in the same temporal resolution as the Arrington.

2.2.2. Measures
The questionnaire was designed using a “funnelling” approach to determine if participants detected the colour change without tipping them off that a colour change had occurred. The first question asked participants to name the colour of the backs of the cards. The second question asked “Did you notice anything unusual happen to the cards? If yes, please describe what you noticed”. Ample space was provided for a participant’s free report. Participants were classified as having detected the change if they explicitly stated that they noticed the cards change colour.

2.3. Analysis
The videos recorded by the eyetracker overlaid an eye-movement cursor on top of a video of the participant’s view of the magic trick performance (30 fps). These videos were coded by hand using the frame-by-frame video annotation tool Anvil (Kipp, 2001) to determine where participants looked and whether participants were overtly misdirected during the critical period after the misdirection cues.

2.3.1. Video coding
Two different areas of interest were defined. One of these combined the cards and hands, which we call collectively “cards”. If participants can ignore the misdirection, we would expect them to only fixate this area of interest. The second area was the face and head, which we will refer to as “face”. Eye movements were analysed from the beginning of the magic trick (cards were removed from the packet) until after the colour change. A research assistant, who was blind to the instruction conditions, classified eye positions frame by frame. A 2 cm tolerance was set around each interest area.

Participants were classified as being overtly misdirected if they looked up at the magician’s face during the time of the misdirection. The time period of interest began on the frame when the magician looked down at the cards (Fig. 1.3) and ended after the colour change. A main focus of the trick prior to the misdirection involved the packet of cards, which meant that participants should be attending to the hand and cards prior to the misdirection. Participants who fixated the magician’s face before the misdirection cue were not coded as being overtly misdirected.

3. Results
Eye movement data from seven participants were lost due to poor calibrations or corruptions in the video files.

3.1. Overall viewing times
The first analysis looked at whether the instruction and the context influenced participants’ overall viewing strategies. Fig. 2A shows the percentage of time spent fixating the cards and face interest areas. Although these areas of interest were spatially relatively small, participants’ spent 93% of their time fixating one of the two areas of interest (video M = 92.2%, SD = 0.50; live M = 95.0%, SD = 0.36).

An ANOVA with context (live vs. video) and instruction (naïve vs. warned) as between-group variable and percentage time spent fixating the face as dependant variable, found a significant main effect of context \( F(1,92) = 4.67, p = .033, \eta^2 = .048 \) with participants in the video condition spending more time looking at the face than those in the live situation. There was also a significant main effect of instruction \( F(1,92) = 9.94, p = .002, \eta^2 = .097 \) with instructional warning reducing the time fixating the face. Although participants in the warned condition were explicitly told to keep their eyes on the cards, they spent a significant amount of time fixating the face \([t(48) = 11.2, p < .0005]\). Only less than 5% of the warned participants did not look at the face at all. The context by instruction interaction was not significant \( F(1,92) = .34, p = .53, \eta^2 = .004 \).

The next analysis looked at the time spent fixating the cards. There was a significant main effect of instruction \( F(1,92) = 6.56, p = .012, \eta^2 = .067 \) with participants fixating the cards more when instructed to do so. However, as is apparent in Fig. 2A, the warned participants did not spend all of their time fixating the cards. There was a significant main effect of context \( F(1,92) = 9.77, p = .002, \eta^2 = .096 \), illustrating that participants spent more time fixating the cards in the live face-to-face context than the video condition. There was no significant context by instruction interaction, \( F(1,92) = .79, \eta^2 = .001 \).

3.2. Can participants use top-down control to inhibit overt misdirection?
Participants were classified as being overtly misdirected if they moved their eyes from the cards to the face during the time at which the magician’s gaze moved up and he asked the question. Fig. 2B shows the percentage of participants who were overtly misdirected as a function of context and instruction. The misdirection was clearly effective in triggering participants’ overt eye movements to the face. As the dependant variable was categorical (misdirected vs. not misdirected) the data was analysed using a loglinear analysis with instructions and context as factors [effect sizes are shown as odds ratios, which are only calculated for 2 by 2 contingencies (i.e. not the interactions)]. Participants in the warned condition were significantly less likely to look up to the face than the naïve participants \( \chi^2 = 5.94, p = .015 \), indicating that participants were able to exert some control over the impact of the magician’s social cues on where they looked. Odds ratios revealed that warned participants were 3.19 times less likely to look up. There was no significant effect of context \( \chi^2 = 1.61, p = .20 \) (odds ratio = 1.79) and no significant interaction \( \chi^2 = 0.85, p = .355 \). The overt misdirection therefore did not rely on a face-to-face interaction.

3.3. Does warning improve detection?
Fig. 2C shows the percentage of participants who did not report seeing the colour change, and thus were covertly misdirected, as a function of context and instruction. A log linear analysis found no effect of instruction \( \chi^2 = .012, p = .91 \) reflecting that the warning did not improve detection (odds ratio = 1.05). There was no significant interaction by context interaction \( \chi^2 = 2.01, p = .16 \), but participants in the video condition were less likely to detect the change than in the live face-to-face context \( \chi^2 = 10.6, p = .001 \) (odds ratio = 4.73).

3.4. Relationship between eye movements and change detection
To determine if overt misdirection predicted success of the magic trick, we compared the detection rate of the card colour change for those who were looked at the magician’s face against those who were not misdirected. The results indicated that overt misdirection did not affect the magic trick. Specifically, participants who were overtly misdirected were no more likely to miss the change (75.3%), compared to those who were not overtly misdirected (71.4%) \( \chi^2 = 0.16, p = .67 \) (odds ratio = 1.22).
4. Discussion

Past research has shown that when presented with pictures and videos containing people, participants prioritize faces in general and the eyes in particular (Birmingham et al., 2008; Fletcher-Watson, Findlay, Leekam, & Benson, 2008; Yarbus, 1967). The present study found that while participants spent a substantial time looking at the face of the magician whether he was live or on video, participants spent significantly less time looking at the face in the live face-to-face condition. This finding, that looks to the face of another person may be reduced in a live versus video situation, replicates the recent finding by Laidlaw et al. (2011). Nevertheless, it is noteworthy that there can be exceptions, e.g., when a live situation is more social and conversational (Freeth et al., 2013).

In the present study we exploited a well-documented social misdirection cue that has been reported to automatically orient attention towards a person’s face, namely the act of posing a question (Tamariz, 2007). Consistent with this work, we found that 82% of the naive participants looked at the magician’s face when a question was posed.

Our first aim was to investigate whether top-down control over this social attention effect was possible, and if so, would it be influenced by situation context, i.e., live vs. video. Half of our participants were explicitly instructed to keep their eyes on the magician’s cards. The percentage of participants who now looked at the magician’s face declined from 82% to 59%. This decline in the participants who were explicitly told to keep their eyes on the cards indicates that participants do have some top-down control over whether they are overtly misdirected or not. Nevertheless, more than half of the warned participants looked at the face. Thus these data also agree with results from a recent scene viewing study showing that although people can inhibit the extent to which they fixate eyes even when they are explicitly instructed not to do so (Laidlaw et al., 2012). Finally, while participants spent less time looking at faces in the live compared to the video condition, the extent to which this tendency was modulated by the instructions was independent of the context (i.e. live vs. video). This is revealed by two convergent lines of evidence. First, warning people to keep their eyes on the cards reduced looks to the face by the same degree for live and video situations. Secondly, eye movements that were elicited by the social misdirection cue/question were equivalent for live and video displays. In sum, top-down control over eye movements does not vary with changes in context, although in absolute terms more eye movements are directed to the face of a video version of a person. One potential limitation was that the display sizes between the live and the video conditions were not entirely matched, which could potentially lead to a greater energetic cost of moving the head and the eyes from the cards to the face in the live condition (Solman & Kingstone, 2014). However, as there was no significant difference in the levels of overt misdirection between these two conditions, it is reasonable to conclude that the change in visual angle did not substantially influence the results.

Our second aim was to explore the link between overt and covert attentional orienting. The magician used misdirection to prevent participants from noticing the back of the playing cards changing colour. When attended, the colour change is extremely salient and magicians frequently use these types of colour changes to visibly transform one card into another (Giobbi, 1994). The misdirection employed here resulted in 70% of the participants failing to notice this visible colour change thus demonstrating that they were successfully misdirected. This type of change blindness provides a valuable measure of attentional orienting that is independent of where people look (i.e. covert orienting). Unlike overt orienting, which was affected profoundly by warning, change detection remained steady whether participants were warned about the misdirection or not. In other words, the data suggest that
covert attentional orienting in response to social misdirection is not modulated by top-down control; this is disparate from, and despite significant changes in, top-down control in overt orienting. Convergent with this conclusion is our finding that change detection was insensitive to warning despite baseline differences in the effectiveness of the card trick overall. Indeed, these findings dovetail with previous work on misdirection which has shown that detection of an event is independent of where people look, regardless of whether detection involves a transient event such as the dropping of a lighter (Kuhn & Findlay, 2010; Kuhn & Tatler, 2005; Kuhn et al., 2008), or noticing how one coin changes into another (Smith et al., 2012). Similarly here participants who were overtly misdirected were no more likely to miss the change (75.3%), compared to those who were not overtly misdirected (71.4%). The covert attentional mechanisms responsible for visual awareness seem to be functionally independent of where people look, and critically, unaffected by top-down modulation of the overt attention system.

5. Conclusion

Our eye movements serve a dual function; on the one hand they are used to prioritize relevant information and as such form a pivotal role in selective attention. However, they also play an important role in non-verbal communication, such as signalling desires and intentions or disambiguating verbal content. For example, during a social dinner, people use gaze to signal an upcoming bite (Wu et al., 2014), or gaze can also be used to indicate relations to social status (Gobel, Kim, & Richardson, 2015). We suspect that people look up in response to a question to acknowledge that they heard it, or to be polite when giving their answer. These eye movements are like a paralinguistic act that is meant to communicate that the listener understands their position in the interaction (Bavelas, Coates, & Johnson, 2002) in the same way that nodding communicates that one is listening and understanding (Cummins, 2012), or saying fillers like “err” show that one still has something to say (Bavelas & Gerwning, 2011). Verbal and non-verbal communication requires top-down control, as failing to control our oculomotor behaviour can easily lead to miscommunications. We therefore must have some control over where we look and this top-down control is particularly important when we engage in real social interactions. Covert attentional processes are by nature concealed and play no role in signalling information. Within a social context an ill-placed covert shift of attention is far less disastrous than an inappropriate glance, and thus there is less need to continuously control covert orientive orienting. This framework of attention predicts that we have more top-down control over where we look than where we attend covertly, and our results support this view. Moreover, within this framework we suggest that differences found in eye movements between the lab and the real world may directly result from differences in the way our eyes are used to signal information.

Author contributions

G.K and A.K. designed the research, R.T. and N.T. conducted the research and coded the data, G.K analysed the data, G.K and A.K. wrote the paper.

References


Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). The Premotor theory of attention: Time to move or to be polite when giving their answer. These eye movements are like a paralinguistic act that is meant to communicate that the listener understands their position in the interaction (Bavelas, Coates, & Johnson, 2002) in the same way that nodding communicates that one is listening and understanding (Cummins, 2012), or saying fillers like “err” show that one still has something to say (Bavelas & Gerwning, 2011).


