Attention Capture by Direct Gaze is Robust to Context and Task Demands

Adam Palanica and Roxane J. Itier
Department of Psychology, University of Waterloo, Waterloo, ON N2L3G1, Canada
Adam Palanica: apalianc@uwaterloo.ca

Abstract

Eye-tracking was used to investigate whether gaze direction would influence the visual scanning of faces, when presented in the context of a full character, in different social settings, and with different task demands. Participants viewed individual computer agents against either a blank background or a bar scene setting, during both a free-viewing task and an attractiveness rating task for each character. Faces with a direct gaze were viewed longer than faces with an averted gaze regardless of body context, social settings, and task demands. Additionally, participants evaluated characters with a direct gaze as more attractive than characters with an averted gaze. These results, obtained with pictures of computer agents rather than real people, suggest that direct gaze is a powerful attention grabbing stimulus that is robust to background context or task demands.

Keywords

Gaze direction; Eye-tracking; Body scanning

Introduction

Human faces convey important social information such as a person’s identity, gender, age, race, or emotions, and due to their social and biological significance, they can capture and engage attention in a powerful way (Palermo and Rhodes 2007). One major social cue is eye contact, also called direct or mutual gaze, which can be used to control social interactions and convey information such as friendliness, intimacy, or behavioral intent (Kleinke 1986). Even newborns prefer to look at direct than averted gaze faces (Farroni et al. 2002), and the existence of an innate neurocognitive mechanism specialized in detecting eyes and their gaze direction has been proposed. The function of this Eye Direction Detector (EDD, Baron-Cohen 1995) would be to determine whether one is the focus of the viewer’s attention.

Research has shown that direct gaze can capture and engage attention better than averted gaze (e.g., Senju and Hasegawa 2005; von Grünau and Anston 1995). For example, in Posner-like attention paradigms in which a face cue is centrally presented prior to the onset of a lateral target, target detection (as measured by response time) is longer when the gaze of
the face is direct rather than averted (Senju and Hasegawa 2005; Vuilleumier 2002). This suggests that direct gaze attracts and retains attention at the cue location so that its orienting towards target location is delayed. In contrast, averted gaze usually orient attention toward the direction of gaze as measured by faster response times to targets at gazed-at than at non-gazed at locations (Driver et al. 1999; Friesen and Kingstone 1998; Langton and Bruce 1999; Ricciardelli et al. 2009; for a review, see Frischen et al. 2007). Visual search paradigms have shown that direct gaze is also detected faster and more accurately than averted gaze in a crowd of opposite-gaze distractors (i.e., a direct gaze target among several averted gaze distractors, or vice versa; Conty et al. 2006; Doi and Ueda 2007; Doi et al. 2009; Senju et al. 2005; von Grünau and Anston 1995). Recently, this “stare-in-the-crowd effect” has been replicated using eye-tracking, showing that direct gaze targets are visually located faster than averted gaze targets, although only in the left peripheral visual field (Palanica and Itier 2011). Other eye-tracking research has shown that participants spend more time viewing virtual character faces with a direct rather than an averted gaze (Mojzisch et al. 2006; Wieser et al. 2009). Taken together, these studies suggest that direct gaze is attention grabbing.

Recent studies also suggest that gaze direction modulates cognitive processes related to face and person perception, such as recognizing identity, gender, or emotions in complex ways (for a review, see Itier and Batty 2009). Gaze direction also plays a role in attractiveness evaluations, with direct gaze faces being judged more favorably than averted gaze faces (Conway et al. 2008; Ewing et al. 2010; Kampe et al. 2001; Main et al. 2010). Additionally, attractive facial cues can modulate gaze categorization with gaze direction being judged faster for opposite-sex faces that have exaggerated sex-typical shape cues (e.g., females with large eyes or males with thick eye brows; Jones et al. 2010a). However, whether gaze direction modulates attention differently depending on task demands is unknown. Attention allocation is often studied by monitoring eye movements, in particular, fixation location, onset time, and duration (for a review, see Henderson 2006). Free-viewing allows for the investigation of naturalistic gaze patterns under no specific task constraints, such as during typical one-on-one social interactions. However, eye movement patterns can change dramatically under task constraints (Yarbus 1967), yet the effects of specific task demands on gaze patterns (and thus on attention allocation) are not well understood. In the present study, we investigated whether the attention grabbing effect of direct gaze would vary when participants were freely exploring the stimuli compared to when they were rating the attractiveness of the stimuli.

Similarly, whether the social context in which direct gaze is seen modulates its attention grabbing effect remains unstudied. To the best of our knowledge, all of the research examining the effects of mutual gaze on attention has used isolated face stimuli. Real life social interactions, however, generally involve the perception of an entire person in a particular social context. It is currently unknown whether the attention grabbing effect of mutual gaze could be modulated by the context of a full character (rather than just the face) and the background scene in which it is presented. Viewing characters within a social setting context might be more realistic than against a blank background setting and this may have an impact on the perception of gaze and its attention grabbing effect, with a potentially stronger attention signal in the social context situation.
In the present study, we tested whether the attention grabbing effect of direct gaze would still be seen in the presence of full characters, and whether it would be modulated by task demands and context. We presented participants with full virtual characters who were looking either straight ahead, as if looking at the participant (direct gaze), or away (averted gaze). We also monitored participants’ eye movements while they were both visually exploring the pictures without any explicit task (free-viewing session) and while they were rating each character on physical attractiveness (attractiveness rating session). Scene background context was manipulated as a between subject factor using either a blank white screen (no context condition) or a bar scene (context condition). If direct gaze was as robust as suggested by previous studies using face stimuli, we expected that participants would move their eyes from the central fixation towards faces faster when the face was looking straight than away (e.g., von Grünau and Anston 1995), and would also fixate longer on direct than averted gaze faces as reported earlier (Mojzisch et al. 2006; Wieser et al. 2009). In other words, we expected that the rest of the character’s body would not diminish the attention grabbing effect of direct gaze. We also predicted that direct gaze would have a greater impact in the context condition than in the no context condition because viewing a set of characters within a social environment may seem more realistic compared to viewing them against a blank background, and the effect of direct gaze may convey a stronger signal to the observer. Lastly, we predicted that direct gaze would also have a greater impact in the attractiveness rating session than in the free-viewing session. Since it has been shown that direct gaze faces are evaluated more favorably than averted gaze faces in attractiveness ratings (Conway et al. 2008; Ewing et al. 2010; Kampe et al. 2001; Main et al. 2010), participants may be more inclined to view direct gaze faces longer in the attractiveness rating session because more attention would be allocated to judging the more favorable stimuli (i.e., direct gaze).

Method

Participants

Data were collected from a total of 56 undergraduate students (14 female and 14 male in each context condition, 18–25 years, M age = 19.9) from the University of Waterloo (UW), who participated in the study for course credit. All participants had normal or corrected-to-normal vision, and all were of heterosexual orientation. They all signed informed written consent letters and the study received full ethics clearance from the UW Research Ethics Board.

Stimuli

A set of 30 virtual humans (half male), or computer agents, were obtained from Vizard 3.15 software. Agents were viewed individually against either a white background (no context condition) or a bar scene (context condition; see Fig. 1). Other than the background settings, all other details were identical in both context conditions. The scenes were displayed at a resolution of 1,152 × 864 pixels and subtended 29.2° × 22.2° of visual angle at the viewing distance of 0.70 m. Each of the agent’s full body (including the head) subtended visual angles of 6.5° horizontally by 19.5° vertically (male and female bodies were of the same overall dimensions). Each of the agent’s faces subtended 2.0° horizontally by 2.7° vertically,
which was situated at 8.6° visual angle from the center of the monitor (i.e., the fixation cross). Male and female faces were of the same size.

In a pilot study, 20 independent undergraduates (11 female) of the same age as participants of the present study rated these 30 characters on physical attractiveness. Using a 9-point Likert scale, the results showed an average rating of 5.8 (SD = 0.8) for female characters and 5.2 (SD = 1.4) for male characters, indicating that the characters were of “moderate” attractiveness. This was important to avoid any biased responses that may be associated with either extremely attractive or extremely unattractive individuals.

In order to control for gaze direction and body position, all of the individual characters were viewed four times in either context condition in a counterbalanced fashion; i.e., each character was viewed with a direct eye gaze (as if looking at the participant) and an averted eye gaze (as if looking away from the participant) facing ¾ left, and then with a direct and an averted eye gaze facing ¾ right. All faces were of neutral expression. The gaze of each agent was manipulated using Adobe Photoshop 11.0 and each picture was mirror-reversed to avoid any bias between the right and left sides of gaze or body position. This gave a total of 120 stimuli that were presented to participants in a semi-randomized order; male and female characters were counterbalanced such that the same gender was never viewed more than once in a row (i.e., presentation was male, female, male, female, etc.).

Apparatus

The stimuli were presented on a Viewsonic PS790 CRT 19-inch color monitor driven by an Intel Corel 2 Quad CPU Q6700 with a refresh rate of 60 Hz. Eye movements were recorded using a remote EyeLink 1,000 eye-tracker from SR Research with a sampling rate of 1,000 Hz. The eye-tracker was calibrated to each participant’s dominant eye, but viewing was binocular. Calibration was done using a nine-point calibration accuracy test. Calibration was repeated if the error at any point was more than 1°, or if the average error for all points was greater than 0.5°. Chin and forehead rests maintained participants’ viewing position and distance.

Materials and Procedure

Each participant was randomly assigned to either the no context or context condition. In each context condition, participants performed two sessions: first, a free-viewing session, and then an attractiveness rating session. Before the start of the study, participants were lead to believe that the purpose of the eye-tracking equipment was to investigate their pupillary reactions in response to the different luminosity of the pictures and to assess the dilations of the pupil in response to the stimuli (as in Hewig et al. 2008). This procedure ensured that participants were not aware that their eye movements were recorded. In the free-viewing session, participants were instructed to “look at the pictures as you normally would.” Pictures of characters were individually presented in the center of the monitor for 4 s. A white screen was presented between each stimulus with a black centered fixation cross. After 1.5 s, the fixation cross was replaced by a fixation trigger that participants must have focused on for 300 ms to activate the next trial. Thus, the starting point of eye movements for each trial was always in the center of the screen close to the hip region of the characters.
Once all 120 stimuli were viewed, participants were then instructed to rate the physical attractiveness for each of the characters. In the attractiveness rating session, participants viewed each character again in a new semi-randomized order. However, after each stimulus, they were presented with a Likert scale (1–9 steps) asking “How physically attractive is this character?” Participants used the computer mouse to rate the Likert scale, and were not able to view the next character until they answered the question. Since we did not want to influence the natural scanning of the characters before the attractiveness rating session, the two experimental sessions were not counterbalanced so that all participants performed the free-viewing session before the attractiveness rating session. This was to ensure that any explicit task for scanning was given after freely viewing the characters. For both experimental sessions, participants were allowed to take a self-paced break after each block of 60 trials. At the beginning of each block, participants’ eye movements were recalibrated. After the experiment, all participants were thoroughly debriefed about the real purpose of the experiment and asked whether they became suspicious about any details of the eye-tracking recordings. They all denied having any suspicion. The entire experiment lasted approximately 60 min.

**Data Analysis**

Any fixation of less than 80 ms duration was removed from the analyses to ensure that incomplete fixations were not counted as part of participants’ gaze patterns. For each agent, there were two main regions of interest (ROI), including the face and the whole body (Fig. 1). The background of the characters comprised less than 10% of the total viewing time across tasks and context conditions and will thus not be reported or discussed; there were also no differences in viewing patterns for the background between context conditions. The dependent variables for each participant and each picture consisted of: (1) average viewing time within each ROI, (2) average number of fixations within each ROI, (3) average onset time of first fixation landing in each ROI for the first time, and (4) average duration of the first fixation landing in each ROI for the first time. The average onset time of first fixation landing in each ROI was the average time taken to move the eyes from the fixation cross to the various ROI for the very first time in each trial. Data for the average viewing time and number of fixations were normalized to the average size of the body between male and female characters (face ROI were identical for male and female characters). This was computed by dividing each variable by the average pixel area of the body ROI across characters.

Eye movements were compared across context conditions and tasks using a 2 (gaze direction) × 2 (ROI) × 2 (tasks) × 2 (context condition) mixed ANOVA where gaze, ROI, and tasks were within subject factors and context was a between subject factor. No main effects or interactions with context condition were found for any of the dependent variables, thus not supporting the idea that context can modulate the attention grabbing effect of direct gaze. Consequently, all 56 participants from both context conditions were collapsed together and the remainder of the paper will focus on the effects of tasks. All data were analyzed using a 2 (gaze direction) × 2 (ROI) × 2 (task) repeated measures ANOVA to compare eye movements between the two tasks. The Greenhouse-Geisser degrees of freedom correction was used whenever necessary. Preliminary analyses revealed no significant effects of
participant gender or character gender on eye movements for the effect of gaze direction, so these variables were excluded from further analyses.

**Results**

**Eye-Tracking Data**

Our results showed that, on average, about 55% of all first fixations went directly to the face; and, about 81% of these first fixations on the face remained on the face for a second fixation (i.e., only 19% of the second fixations after the face were made to the body regions). Thus, if participants did not fixate the face on their first eye movement, they did so on their second or third fixation.

For the average viewing time per ROI, a main effect of task was found, $F(1, 55) = 5.53, p < .05$, $\eta^2 = .09$, indicating longer viewing times in the attractiveness rating session than the free-viewing session. A main effect of gaze, $F(1, 55) = 8.56, p < .01$, $\eta^2 = .14$, showed longer viewing times for characters with a direct gaze than an averted gaze. A main effect of ROI, $F(1, 55) = 29.43, p < .001$, $\eta^2 = .35$, indicated significantly longer viewing times for the face as compared to the body. On average, about 62% of the entire viewing time was within the face ROI. There was also a significant interaction between gaze and ROI, $F(1, 55) = 11.04, p < .005$, $\eta^2 = .17$. Paired $t$ tests showed that participants viewed faces with a direct gaze longer overall than faces with an averted gaze, $t(55) = 4.05, p < .001$, and viewed bodies with an averted gaze longer than bodies with a direct gaze, $t(55) = 2.20, p < .05$ (Fig. 2). Lastly, a significant interaction between task and ROI was found, $F(1, 55) = 20.12, p < .001$, $\eta^2 = .27$. Paired $t$-tests showed that participants spent a longer time viewing faces in the attractiveness rating session ($M_s = 2,123.8$ ms vs. $1,785.8$ ms), but spent a longer time viewing the body ($M_s = 1,318.0$ ms vs. $1,057.8$ ms) in the free-viewing session (all $p < .001$). There were no significant interactions between gaze and task. Thus, direct gaze faces were viewed longer than averted gaze faces in both the free-viewing and attractiveness rating sessions. This was confirmed by further analyzing each session individually (Gaze × ROI interaction for each task, $p < .05$).

For the average number of fixations per ROI, a main effect of gaze was found, $F(1, 55) = 4.48, p < .05$, $\eta^2 = .08$, indicating more fixations overall for characters with an averted gaze than a direct gaze. A main effect of ROI approached significance, $F(1, 55) = 3.35, p = .073$, $\eta^2 = .06$, due to marginally more fixations to the body than to the face (Fig. 3). Lastly, a significant interaction between task and ROI was found, $F(1, 55) = 21.66, p < .001$, $\eta^2 = .28$. Paired $t$-tests showed that participants made more fixations to faces in the attractiveness rating session ($M_s = 4.2$ vs. $3.3; p < .001$), but made more fixations to the body ($M_s = 4.6$ vs. $4.0; p < .05$) in the free-viewing session. No other effects were found.

For the average onset time of first fixation landing in each ROI, only a main effect of ROI was found, $F(1, 6) = 113.53, p < .001$, $\eta^2 = .95$, indicating significantly earlier fixations to the face as compared to the body (Fig. 4). No effect of gaze was found.
Similarly, for the average duration of first fixation landing in each ROI, only a main effect of ROI was found, $F(1, 6) = 12.17, p < .05, \eta^2 = .67$, indicating that first fixations to the face were significantly longer than to the body (Fig. 5). No effect of gaze was found.

**Attractiveness Rating Data**

Attractiveness ratings were analyzed using a one-way ANOVA with gaze as a within subject factor. A main effect of gaze, $F(1, 54) = 16.86, p < .001, \eta^2 = .24$, indicated higher ratings for characters with a direct gaze ($M = 5.5, SD = 1.2$) than an averted gaze ($M = 5.4, SD = 1.1$). Preliminary analyses also revealed no significant effects of context condition on attractiveness ratings.

**Discussion**

The present study used eye-tracking to investigate whether a character’s gaze direction influenced the way observers looked at the face in the context of an entire human body. We also investigated whether social setting context and task demands modulated the attention grabbing effects of direct gaze. It was predicted that direct gaze would capture participants’ initial attention faster than averted gaze, and that participants would spend longer times viewing direct gaze than averted gaze faces. We also expected that direct gaze would have a greater impact when viewed in a social setting context, and also when observers were asked to rate the attractiveness of the characters. Overall, gaze direction did not seem to affect participants’ initial eye movements, however, it did influence overall viewing times, with direct gaze faces being viewed significantly longer than averted gaze faces, and this was true regardless of social setting context or task demands. These effects and their implications are discussed below.

Our study did not support the prediction that the background setting of the characters would influence participants’ gaze patterns. It was hypothesized that direct gaze would have a greater influence in the context condition compared to the no context condition because direct gaze may convey a stronger signal in a more realistic setting. However, the omnibus ANOVA revealed no main effect or interactions with context for any of the dependent variables. Perhaps context did not influence participants’ gaze patterns significantly in this study because the background contained irrelevant information for the tasks at hand (i.e., either free-viewing or rating attractiveness).

In general, faces were fixated earlier and longer than the bodies across experimental tasks, which supports other eye-tracking research indicating that faces are attention grabbing even in the context of a full character (Hewig et al. 2008). However, the prediction that characters’ gaze direction would modulate the initial onset time of face scanning was not supported. It was predicted that direct gaze faces would attract initial vigilance faster than averted gaze faces given the literature on the stare-in-the-crowd effect (Conty et al. 2006; Doi and Ueda 2007; Doi et al. 2009; Senju et al. 2005, von Grünau and Anston 1995), but this was not found. It is possible that visual field may modulate initial eye movements to faces with direct versus averted gaze. For example, a recent eye-tracking study involving the detection of direct or averted gaze characters amongst three opposite-gaze distractors showed that direct gaze targets elicited faster reaction times and earlier onset times of first
fixation compared to averted gaze targets, but only for characters situated in the periphery (Palanica and Itier 2011). No gaze difference was found for centrally-situated characters. The lack of gaze effect on the initial onset time of first fixation to the face in the current study may thus be due to the central location of the character.

Even though gaze direction did not influence the initial onset times of first fixation, overall, gaze direction did influence average viewing times. Characters with an averted gaze had longer viewing times toward the body than characters with a direct gaze. Additionally, faces with a direct gaze had longer average viewing times than faces with an averted gaze across experimental tasks. This means that characters with a direct gaze engaged more attention toward the face, and consequently engaged less attention toward the body. This finding was not a matter of fixation number, as there were not more fixations to direct than averted gaze faces. In fact, more fixations overall were found for averted gaze characters than direct gaze characters. Thus, direct gaze may have delayed attention disengagement (Senju and Hasegawa 2005; Vuilleumier 2002) throughout stimulus presentation. It is likely that the gaze difference in average viewing times was diluted across the fixations landing on the face (3–4 fixations on average per trial). This gaze effect is relatively large (~56 ms longer viewing time for direct than averted gaze faces) and replicates other eye-tracking research reporting longer viewing times for direct than averted gaze virtual characters (Mojzisch et al. 2006; Wieser et al. 2009).

It may also be argued that the relatively shorter viewing times for averted gaze faces was a result of the averted gaze direction orienting the attention of the viewer away from the face. Attention orienting by gaze, or gaze-cueing effect, triggers faster target detection for peripheral targets that are congruent, rather than incongruent, with the gaze direction (Driver et al. 1999; Friesen and Kingstone 1998; Frischen et al. 2007; Langton and Bruce 1999; Ricciardelli et al. 2009), and a stronger gaze-cueing effect for dominant-, rather than subordinate-looking faces has recently been reported (Jones et al. 2010b). Recent studies have also shown that eye movements follow the direction signalled by gaze (Castelhano et al. 2007; Itier et al. 2007; Palanica and Itier 2011; Zwickel and Võ 2010). However, attention orienting by averted gaze does not fully explain why longer viewing times to the body were found for averted gaze characters, since gaze was directed to the left or right, not downwards toward the body.

Overall, the current study suggests that direct gaze can be a powerful stimulus for attracting attention during visual scanning despite the presence of a character’s body. Furthermore, static pictures of computer agents were used, unlike Mojzisch et al. (2006) and Wieser et al. (2009) who used dynamic video stimuli of faces. This also suggests that the mere sight of direct gaze can attract attention, aside from the effects of face or eye motion. The lack of gaze by task interaction also suggests that direct gaze can influence visual exploration regardless of task demands and is thus a powerful attention grabber. These findings lend support to the possible Eye Direction Detector mechanism specialized in detecting the presence of eyes in the environment and the direction of gaze (Baron-Cohen 1995). It may also be argued that because participants were presented with entire bodies of individuals, scanning someone’s body when the actor is engaged in mutual gaze with the observer is a social faux pas. That is, according to typical social norms in everyday society, it is generally
considered impolite to scan or stare at someone’s body regions when they are looking directly at you, especially when those body regions are considered signals of sexual attractiveness (e.g., breasts, hips, legs). Thus, this could also account for why participants viewed direct gaze faces longer, but viewed the body longer when the characters had an averted gaze.

It was also predicted, however, that direct gaze would have a greater impact on eye movements in the attractiveness rating session with participants being more inclined to view direct gaze faces longer than in the free-viewing session. However, there were no significant interactions between gaze and task sessions. Perhaps direct gaze, compared to averted gaze, triggers a reward effect when assessing attractive faces (Kampe et al. 2001) without any difference being allocated to viewing time between direct versus averted gaze. That is, participants may intrinsically evaluate direct gaze as more attractive than averted gaze without needing any extra time in making the judgment. Nevertheless, direct gaze played a significant role in physical attractiveness ratings, with direct gaze characters being rated as more attractive than averted gaze characters, supporting previous research (Conway et al. 2008; Ewing et al. 2010; Kampe et al. 2001; Main et al. 2010).

Participants also fixated the face longer and more frequently in the attractiveness rating session than in the free-viewing session (by an average difference of 338 ms), supporting other research suggesting that judging the face is a stronger predictor of attractiveness ratings than judging the body (Alicke et al. 1986; Currie and Little 2009; Furnham and Reeves 2006; Mueser et al. 1984; Peters et al. 2007). However, it should be noted that the order of the free-viewing and attractiveness rating tasks were not counterbalanced to decrease the influence of naturalistic scanning behavior during free-viewing. Consequently, the stronger preference for face viewing in the attractiveness rating session could possibly stem from something other than task effects. This is a methodological limitation of the current study which should be addressed by future research.

Overall, the current study supports the importance of direct gaze during person visual exploration. Even when presented with entire bodies of others, observers scanned faces with a direct gaze longer than faces with an averted gaze, and also evaluated characters with a direct gaze as more attractive than characters with an averted gaze. As these results occurred in the presence of static computer agents, and not real people, this suggests that direct gaze is a powerful attention grabbing stimulus which has a robust effect despite context settings and task demands.

Acknowledgments

This research was supported by funding from the Canadian Institutes for Health Research (CIHR MOP-89822), the Canada Foundation for Innovation (CFI), the Ontario Research Fund (ORF) and the Canada Research Chair (CRC) program to RJI. We thank Dr. Colin Ellard for providing us with the Vizard software stimuli.

References


Fig. 1.
Example of the stimuli used; all images were shown in full color. Left panels examples of agents in the no context condition; areas used for the ROI are outlined in red, which include the face and body. Middle panels examples of agents in the context condition. Note that these agents only represent stimuli with direct gaze; averted gaze agents comprised the same identities and body positions, but with differing eye directions. Right panels upper image is a close-up of an agent with direct gaze; lower image is a close-up of an agent with averted gaze. Left and middle displays represent the entire monitor size.
**Fig. 2.**
Average viewing time as a function of ROI and gaze direction. Direct gaze versus averted gaze paired comparison: *p < .05
Fig. 3.
Average number of fixations as a function of ROI and gaze direction
Fig. 4.
Average onset time of first fixation as a function of ROI and gaze direction
Fig. 5.
Average duration of first fixation as a function of ROI and gaze direction