



Investigating social gaze as an action-perception online performance

Ouriel Grynszpan^{1*}, Jérôme Simonin², Jean-Claude Martin³ and Jacqueline Nadel¹

¹ Centre National de la Recherche Scientifique, CNRS-USR 3246, Université Pierre et Marie Curie, Hôpital de La Salpêtrière, Paris, France

² Laboratoire Lorrain de Recherche en Informatique et ses Applications, LORIA-CNRS-INRIA, Vandœuvre-lès-Nancy, France

³ Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, LIMSI-CNRS, Université Paris Sud, Orsay, France

Edited by:

Ulrich Pfeiffer, University Hospital
Cologne, Germany

Reviewed by:

Rick Dale, University of California,
USA

Cristina Becchio, Università degli
Studi di Torino, Italy

Elizabeth Redcay, University of
Maryland, USA

*Correspondence:

Ouriel Grynszpan, Centre National
de la Recherche Scientifique-USR
3246, Université Pierre et Marie
Curie, Hôpital de La Salpêtrière,
47 Boulevard de l'Hôpital,
F-75013 Paris, France.
e-mail: ouriel.grynszpan@upmc.fr

Gaze represents a major non-verbal communication channel in social interactions. In this respect, when facing another person, one's gaze should not be examined as a purely perceptive process but also as an **action-perception** online performance. However, little is known about processes involved in the real-time self-regulation of social gaze. The present study investigates the impact of a gaze-contingent viewing window on fixation patterns and the **awareness of being the agent** moving the window. In face-to-face scenarios played by a virtual human character, the task for the 18 adult participants was to interpret an equivocal sentence which could be disambiguated by examining the emotional expressions of the character speaking. The virtual character was embedded in naturalistic backgrounds to enhance realism. Eye-tracking data showed that the viewing window induced changes in gaze behavior, notably longer visual fixations. Notwithstanding, only half of the participants ascribed the window displacements to their eye movements. These participants also spent more time looking at the eyes and mouth regions of the virtual human character. The outcomes of the study highlight the dissociation between non-volitional gaze adaptation and the self-ascription of agency. Such dissociation provides support for a two-step account of the **sense of agency** composed of pre-noetic monitoring mechanisms and reflexive processes, linked by bottom-up and top-down processes. We comment upon these results, which illustrate the relevance of our method for studying online social cognition, in particular concerning autism spectrum disorders (ASD) where the poor pragmatic understanding of oral speech is considered linked to visual peculiarities that impede facial exploration.

Keywords: agency, gaze control, social cognition, self-monitoring, self-awareness, eye-tracking, virtual reality

INTRODUCTION

When looking at someone else's face, we tend to scan preferentially and consistently the eye and mouth regions (Mertens et al., 1993) although we are not aware of doing so. Nevertheless, two limiting factors interfere with our automatic tendency to stare at one another's eyes: cultural rules and individual differences in emotional sensitivity to eye-contact. When engaged in conversations with others, visual behavior conforms to implicit social display rules requiring fast and subtle adjustments. For instance, making frequent eye-contact can be experienced by the responder as threatening or intrusive and increases her/his emotional state (Senju and Johnson, 2009). This leads interacting partners in many cultures not to stare too long at each other, or even, as in the Wolof tradition, not to look at all at the partner while speaking (Meyer and Girke, 2011). Whether a face-to-face interaction feels comfortable thus depends on how each partaker self-monitors her/his gaze. While acknowledging the socio-cultural determinants that regulate eye movements, the variety of individual differences in the attention to the other's gaze within the normal population should be noted (Frischen et al., 2007). Frischen et al. (2007) suggest that the acute emotional response triggered

by the eye-gaze of others could lead some individuals to learn to avoid looking directly at the eye region of others, thus developing a voluntary control of gaze processing. In this respect, individual differences are interpreted as stemming from the degree of control an individual exerts over her/his reflex of following another person's gaze (Bayliss et al., 2005) and consequently over her/his general gazing at eyes behavior. Strikingly, although we modulate our gaze according to social rules and our own sensitivity to eye-contact, we are usually unaware of controlling it.

This dissociation between control and awareness of control of our ocular movements is an issue that has been largely overlooked until now, despite being a part of the general question regarding the self-monitoring and awareness of action. Blakemore et al. (2002) suggested that parts of the motor system could function in the absence of awareness, especially the motor commands responsible for predicting the fine trajectory adjustment parameters of a movement. They mention as an example the study of Goodale et al. (1994) showing that the displacement of a target during a saccade remained unnoticed by participants although they adjusted their hand to its new position (Goodale et al., 1994). In contrast, "forward models," which are conceived as

internal neural processes predicting the sensory consequences of a movement based on an efference copy of motor signals, are considered instrumental in bringing about the awareness of action (Wolpert et al., 1995; Wolpert and Miall, 1996; Blakemore et al., 2002). Accordingly, forward models come into play most notably when the intentions and goals are clearly stated or in the case of a clear-cut mismatch between sensory prediction and feedback (Slachevsky et al., 2001).

Evidence supports the existence of forward models and efference copies of eye motor signals. For instance, the perceptual invariance of the world despite the visual flow induced by the eye movements on the retina is classically explained by a neuronal mechanism that compensates for the retinal image displacement, by predicting the visual consequences of the eye movements on the basis of an efference copy (Holst and Mittelstaedt, 1950; Sperry, 1950). As postulated by the forward model hypothesis, some visual neurons in the macaque brain have been shown to predict the visual consequences of a saccade by remapping their receptive field before the saccade so that it accounts for the shift in space caused by the saccade (Colby and Goldberg, 1999). This remapping process is coupled with an efference copy, also called the corollary discharge, which runs along a neural pathway that has been at least partly identified (Wurtz, 2008). Functional Magnetic Resonance Imaging (MRI) lends support for the existence of similar visual remapping neuronal activity in the human brain (Merriam et al., 2003, 2007). Emerging evidence indicates that this predictive mechanism could mainly sub-serve visuomotor control (Bays and Husain, 2007).

Whether and in which circumstances visuomotor control may be subject to awareness are questions that could be addressed in the light of the multifactorial two-step model of agency proposed by Synofzik et al. (2008). Indeed, this model posits two levels that would function with relative independence, namely the feeling of agency and the judgment of agency. As in the classical comparator model (Blakemore et al., 1998), they describe the feeling of agency as stemming from a low-level pre-conceptual mechanism that monitors motor outputs and sensory inputs. Yet, such feeling is not sufficient alone to ascribe self-agency. Many examples driven from de-afferented patients (Fournier et al., 2002) or parietal lobe damaged patients (Sirigu et al., 1999) converge in showing that an efference copy cannot explain on its own how we decide about agency attribution. Rather, Synofzik et al. (2008) reason that the judgment of agency results from the congruency between intention and effects, independently from any comparator output. They propose to conceive the judgment of agency as a high-level interpretative process that attempts to find the best plausible cause for an action based on contextual information and personal beliefs. Several recent studies provide support to this perspective. For instance, Spengler et al. (2009) have demonstrated that participants, who were trained to expect a given consequence for their action, subsequently experience an increased sense of agency in case of congruence between their experimentally induced expectation and the actual consequence of their action. Voss and colleagues (Voss et al., 2010) argue that the combination of reduced prediction and excessive self-agency attribution observed in schizophrenia cannot be explained by the comparator model. They distinguish between a predictive

and a retrospective mechanism of agency attribution, showing an exaggerated reliance on retrospection in schizophrenic patients who associate their actions with unrelated external events, while typical adults rely on probabilistic estimations for binding their actions with corresponding effects.

The present study aimed at investigating the relationship between the self-monitoring of gaze and judgment of agency using a newly developed method and experimental platform with expressive virtual characters and real-time gaze-contingent technology based on a viewing window. Baugh and Marotta (2007) advocate using the viewing window paradigm to examine the interactions between perception and action. In this classical paradigm (McConkie and Rayner, 1975), participants are presented with a degraded visual stimulus on which they can control an area with normal clarity (the viewing window). Similar limited-viewing paradigms have previously been exploited to examine the visual exploration of faces in disorders affecting social cognition (Spezio et al., 2007). In our experiment, the entire display was blurred except for a gaze-contingent moving window. The viewing window was meant to stimulate the ascription of self-agency along a bottom-up pathway starting from simple action-perception coupling. To favor bottom-up processes, the experimental task required participants to be naïve about the purpose of the viewing window. Consequently, the participants were left uninformed about the viewing window's functioning and we chose to evaluate their judgment of agency using an open question asked at the end of the experiment. Although methods for disentangling the feeling and judgment of agency often employ instructions asking participants to rate agency on a scale (Kühn et al., 2011), our evaluation method differed in that we sought to measure the emergence of awareness. Therefore, we opted for an interviewing technique that left the participants uninformed of the matter under scrutiny, as implemented in previous studies of agency (Nielsen, 1963; Jeannerod, 2003).

We examined the effects of the window on fixation patterns in order to assess the self-monitoring of gaze and recorded the participants' answers about what controlled the window to assess their judgment of agency. Although the method used here for examining agency could also be relevant for non-social stimuli, the focus of interest in the present study specifically addresses online face-to-face interaction that involves high-order gaze control guided by social considerations. We employed stimuli depicting animated virtual characters rather than real actors to have precise control over the design of their facial expressions, the intonation of their voice and the synchrony between speech and facial movements. Finally, we sought to enhance the ecological validity of the experimental apparatus by creating virtual characters with realistic physical human features and embedding them in videos of real-life everyday environments.

METHODS

PARTICIPANTS

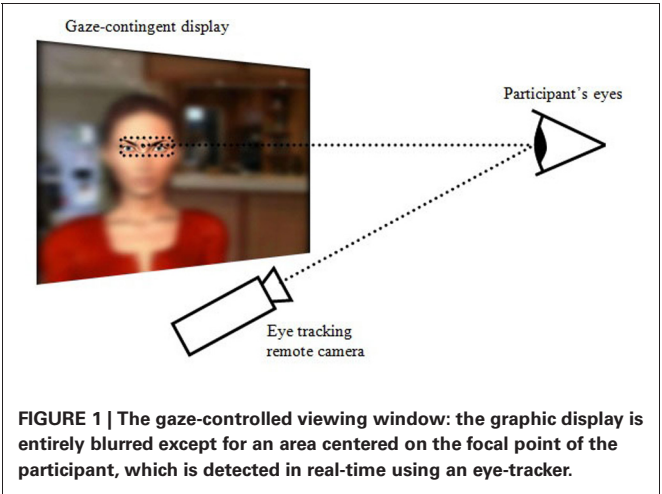
Eighteen adolescents and adults participated in the experiment. They were free of any known psychiatric or neurologic symptoms, non-corrected visual or auditory deficits and recent use of any substance that could impede concentration. Their age ranged from 17 to 40 years with a mean of 28.5 (SD = 6.74). There

were 10 males and eight females. This research was reviewed and approved by the regional ethics committee of Tours, France. Informed consent was obtained from each participant.

MATERIALS

Participants were seated in front of a computer screen located above an eye-tracker (model D6-HS from Applied Science Laboratories with a sampling rate of 50 Hz) that remotely detected their eye orientation, without constraining their head movements or requiring them to wear a helmet. The graphic display was gaze-contingent in that it was entirely blurred, except for an area centered around the current focal point of the participant (**Figure 1**). The apparatus thus simulated a gaze-controlled viewing window, providing real-time visual feedback on the location of the user's gaze. Participants were placed approximately 57 cm away from the screen. The screen's size was 19 inches ($377 \times 302 \text{ mm}^2$) with a resolution of 1280×1024 pixels (approximately $37^\circ \times 30^\circ$ in visual angles). The viewing window was a rectangle with rounded angles measuring 233×106 pixels (approximately $7^\circ \times 3^\circ$ in visual angles, thus encompassing fovea vision). The latency of the gaze-contingent display was of approximately 100 ms between the eye orientation detection and the repositioning of the viewing window. This delay was expected to induce a behavior whereby participants would stabilize their gaze to stay in sync with the viewing window. The gaze-contingency could be switched on or off. In the latter, the eye-tracker would still record the direction of the gaze.

The task presented to the participants was designed to simulate a face-to-face situation involving verbal and non-verbal communication. In each trial, a virtual character told the participants about a recent experience she/he was involved in. Acknowledging the current limitations of technology in creating believable interactions with virtual humans (Groom et al., 2009), the trials were short (less than 25 s) and merely simulated the beginning phase of a social encounter occurring before the participant's turn to talk. The verbal message of the virtual character included an equivocal sentence that implicitly required looking at the facial expressions of the virtual character. For instance, the virtual character would say, "I'm so lucky" while displaying a sad facial expression.



The participants had to answer two close-choice questions that assessed their attention to the facial expressions. The first question was about the character's feelings and the second was about what caused those feelings. There were three possible choices for each question: the correct interpretation; an interpretation coherent with the equivocal sentence left alone, but incoherent with the facial expressions; an interpretation incoherent with both the verbal utterances and facial expressions. **Table 1** provides an example illustrating the task.

The gaze-contingent viewing window was large enough to see the virtual character's two eyes with eyebrows, or the virtual character's mouth (**Figure 1**), which are both highly expressive facial features (Baron-Cohen et al., 1997). A male and a female character were designed with Poser Pro software (Smith Micro Software, Inc.), which enables creating highly realistic virtual humans with animated facial expressions that include creases in the skin. The male and female characters each appeared in half of the trial. They kept gazing at the user throughout the animation. Facial expressions associated with the equivocal sentences were chosen among five basic emotions (disgust, joy, fear, anger, and sadness), which were designed using Ekman's specifications (Ekman and Friesen, 1975). The synchronization of dynamic facial expressions with speech had been studied in a preliminary experiment that included 23 participants (Buisine et al., 2010). Based on its outcomes in terms of recognition performances and perceived realism, we adjusted the facial movements associated with each utterance so that the facial expression would unfold during the entire duration of the utterance and be maintained for 1 sec after the utterance. The participants were to rely predominantly on the facial expressions; thus the characters' emotional intonation was minimized by using synthesized speech, created with Virtual Speaker software version 2.2 for French, from the

Table 1 | Example of an animation presented to the participants.

Animated scene's script (the virtual character is named John)	
Utterance	Facial expression
I was waiting for the bus with Sandra	Joy
But, Franck passed by just at that moment	Surprise
He offered to drive her back home	Sadness
<i>How nice of him!</i>	Anger
CHOICE OF ANSWERS FOR THE FIRST QUESTION, I.E., "HOW DOES JOHN FEEL?"	
<i>He is jealous</i>	
He is anxious	
He is happy	
CHOICE OF ANSWERS FOR THE SECOND QUESTION, I.E., "HOW DO YOU KNOW THAT?"	
<i>Instead of John, Franck gets to be with Sandra on the way back home.</i>	
It was nice of Franck to offer to drive Sandra back home.	
Franck is a poor driver and often has accidents.	

The upper part of the table presents each utterance of the virtual character associated with its simultaneous facial expression. The key sentence inducing ambiguity is in italic font. The possible choices for the two subsequent questions are shown beneath. The correct answers are in italic font.

Acapela Group Company (www.acapela-group.com). Using virtual characters rather than real actors enabled matching the usage of synthesized speech. The virtual characters' lip movements were synchronized with their speech using Poser Pro software. The virtual characters were embedded in videos of real life settings that provided a naturalistic context.

PROCEDURE

The study utilized an ABA design: in the baseline condition, the graphic display was entirely clear; during the experimental condition that followed, the entire display was blurred, and the gaze-controlled viewing window was set on; in the final condition, the gaze-contingent viewing window was set off, and the entire graphic display became clear again. The participants had to perform the same task in all three conditions. There were 20 trials per condition, which totaled 60 different animated scenes. The animated scenes lasted 18 s on average, ranging from 13 to 24 s. The order of the scenes and the condition in which they appeared were randomly counterbalanced across the participants and complied with the constraint that the total duration of the scenes had to be the same in every condition. The experimentation started with a standard calibrating procedure for the eye-tracker. In a first demo animation, both virtual characters introduced themselves and provided the instructions for the task. Just before the experimental condition, a written instruction appeared on the screen that explicitly encouraged the participants to look at the facial expressions ("Think about looking at the characters' faces to understand what they feel"). The purpose of this instruction was to prepare the participants for the gaze-contingent display and exhort them to behave consistently even though their vision would be constrained. The experimental and final baseline conditions also started with a short demo animation showing both virtual characters cheering the participant. The purpose was to give some time for the participants to adapt to the new condition and reward their efforts for continuing the experiment.

The participants were told before they started the experiment that some parts of the visual display would sometimes remain clear, whereas others would be blurred. Yet, they were left uninformed about the fact that the display was gaze-contingent, although they knew that their gaze was being continuously measured by the eye-tracker. At the end of the experiment, they were

asked a question to examine whether they had noticed that they were controlling the viewing window. This question translates into English as the following: "You noticed that in some videos, there were blurred areas and clear areas. What causes the clear areas?" Their answers were recorded and analyzed by two independent judges, as either showing awareness or not. The kappa computation showed a complete agreement between the judges.

DATA ANALYZES

The software we developed automatically recorded eye-tracking data, the scores on the task and the response time. Given that there were 20 scenes in each condition, each of the two closed-choice questions following the animated scenes yielded scores ranging from 0 to 20. Fixations were detected, using a proprietary algorithm of Applied Science Laboratories (the provider of the eye-tracker), on the basis of a cluster of Point-Of-Gaze (POG) that remained in 1° of visual angle for at least 100 ms. The number of POGs collected during a fixation provides a measure of the fixation's duration, which is equal to the duration in seconds divided by the sampling rate. We analyzed the gaze data with a software prototype, developed for the present research, which could handle eye-tracking data on dynamic visual displays. It enabled the aggregation of gaze data on pre-defined rectangular Areas of Interest (AOI). We defined two sets of AOI: first, an AOI that was circumscribed around the head (**Figure 2A**) of the virtual characters and another AOI, named "no-face," which encompassed the rest of the screen; second, an "eyes" AOI that surrounded the eye region and included the eyebrows, and another AOI with the same dimensions for the mouth (**Figure 2B**).

Analyses were performed with the Statistical Analysis Software (SAS www.sas.com). Our hypothesis stated that the gaze would stabilize in the experimental condition. In other words, the number of fixations would decrease, the average duration of fixations would increase and the average distance between two consecutive fixations would decrease. As the experimental design involved repeated measures, we conducted analyses of variance (ANOVA) for these variables using a mixed-design with an unstructured residual covariance matrix. The within-subjects factor was the condition. We also used the bimodal variable derived from the last question of the experiment as an adjustment factor that accounted for the awareness of controlling the viewing window.

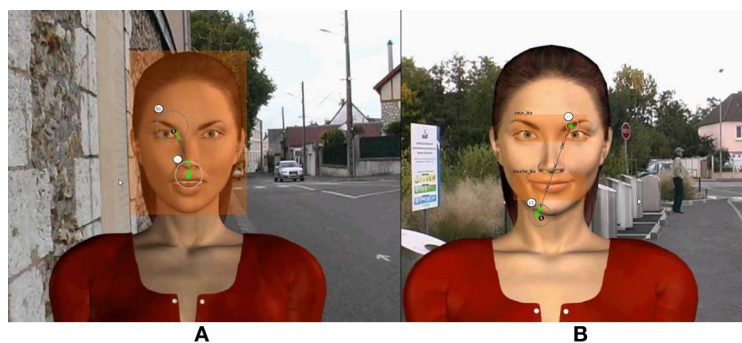


FIGURE 2 | The rectangular Area-of-Interest (AOI) used for analyzing gaze fixations. The circles represent consecutive fixations that are linked by the visual path. (A) the "face" AOI used; (B) the "eyes" and "mouth" AOI.

This bimodal variable divided the participants into two groups: those who gained awareness and those who did not. *Post hoc* *t*-tests were performed using the Tukey adjustment procedure; the *p*-values provided hereunder are adjusted values. We also calculated the corresponding effect sizes, using the commonly accepted threshold fixed at 0.2, 0.5, and 0.8 for small, medium, and large magnitudes, respectively (Zakzanis, 2001). We checked for possible influences of age and gender on the answers given to the last question by performing a *t*-test for age and the Fisher's exact test for gender. To verify that compliance with the task was comparable across conditions, we analyzed the participants' performances based on their scores, response time, and total fixation time on faces. We first checked for possible influences of age and gender on these variables during the baseline condition, using Pearson correlation coefficients for age and the *t*-test (female versus male) for gender. The scores and response time were then processed with the same ANOVA as before. The total fixation time (i.e., the sum of fixation durations) was analyzed using an additional "face/no-face" ("face" AOI vs. "no-face" AOI) within-subject factor. We did not expect the experimental condition to impair the participants' gazing strategies, and thus we assumed that there would not be a condition \times face/no face interaction in this analysis. Finally, we sought to investigate whether the viewing window would influence the visual scanning of the eyes and mouth regions. We thus performed a second ANOVA on the total fixation time using the within-subject factors that follow: the condition, the answer to the last question and an "eyes-mouth" bimodal variable with two modalities: the "eyes" AOI versus the "mouth" AOI.

RESULTS

The answers to the question "What causes the clear areas?" revealed that 9 out of 18 participants noticed that their gaze was responsible for the movements of the viewing window. The other nine participants showed no such awareness. These two groups of participants did not differ significantly in age or gender. Three participants of the latter group acknowledged they were looking at the clear areas, despite not noticing that they controlled them. Three other participants judged that the clear areas were purposely placed on the face by the computer.

The ANOVA showed a significant effect of condition for every dependent variable employed to assess gaze stabilization (Table 2). The number of fixations was significantly lower in the experimental condition compared to the baseline [$t(16) = 4.27$; $p = 0.0016$; $d = 1.20$] and final conditions [$t(16) = 4.14$; $p = 0.0021$; $d = 1.05$]. The average duration of fixations was significantly higher in the experimental condition compared to the baseline [$t(16) = 3.86$; $p = 0.0037$; $d = 1.08$] and final conditions [$t(16) = 3.77$; $p = 0.0045$; $d = 1.00$]. The average distance between consecutive fixations decreased significantly during the experimental condition compared to the baseline [$t(16) = 3.91$; $p = 0.0034$; $d = 0.91$], and final conditions, [$t(16) = 2.85$; $p = 0.0297$; $d = 0.59$].

Performance variables (scores, response time, total fixation time on faces) were not significantly correlated with age, and the *t*-tests comparing female and male did not yield significant differences. Main effects of the condition were also found for each of these variables (Table 2). The scores were significantly higher in the final condition compared to the baseline for both questions, [first question: $t(16) = 3.69$; $p = 0.0053$; $d = 0.81$; second question: $t(16) = 3.79$; $p = 0.0044$; $d = 0.83$]. The scores in the experimental condition were significantly lower than in the final condition for the first question [$t(16) = 2.87$; $p = 0.0284$; $d = 0.81$]. The response times for the first question were significantly longer during the baseline condition than during the experimental condition [$t(16) = 2.68$; $p = 0.041$; $d = 0.57$] and final condition, [$t(16) = 4.93$; $p = 0.0004$; $d = 1.15$]. They also decreased from the experimental to the final condition [$t(16) = 3.13$; $p = 0.0168$; $d = 0.47$]. Although a main effect of the condition was observed regarding response times for the second question, *post hoc* *t*-tests did not reveal any significant effect. The analysis of the total fixation time using the "face/no-face" factor showed that this variable decreased between the experimental condition and the final condition, although the effect size was small, [$t(16) = 6.32$; $p < 0.0001$; $d = 0.16$]. There was an effect of the "face/no-face" factor [$F(1, 16) = 134.97$; $p < 0.0001$], which showed that the participants watched the "face" AOI for a longer period (*Least Square M* = 612 SE = 50) than the "no-face" AOI (*Least Square M* = 51 SE = 9). The condition \times "face/no-face" interaction was not significant.

Table 2 | Means and standard deviations of dependant variables in the three sequential conditions.

	Condition			<i>F</i> (2, 16)	<i>p</i>
	Baseline	Experimental	Final		
Number of fixations	32 (9)	22 (7)	30 (8)	9.30**	0.0021
Average duration of fixations (in number of POG ^a)	22 (8)	36 (17)	23 (8)	7.48**	0.0051
Average distance between consecutive fixations (in pixels)	1204 (399)	938 (269)	1094 (266)	7.86**	0.0042
Total fixation time (in number of POG ^a)	338 (314)	342 (345)	324 (317)	20.65***	<0.0001
First question scores	16.8 (2.9)	17.3 (2.0)	18.6 (1.3)	10.26**	0.0014
Second question scores	16.2 (3.1)	17.1 (2.2)	18.2 (1.8)	12.16***	0.0006
Response time for the first question (in milliseconds)	6964 (2196)	5725 (2220)	4886 (1521)	14.74***	0.0002
Response time for the second question (in milliseconds)	9387 (8857)	6493 (1849)	6351 (2293)	6.09*	0.0108

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

^aPOG: Point-of-Gaze at a sample rate of 50 Hz.

Note should be taken that in none of the above analyses did we observe an interaction between the condition and the answer to the last question. In particular, the variables used to measure gaze stabilization yielded the following interaction statistics: $F(2, 16) = 0.65$ $p = 0.5352$ for the number of fixations; $F(2, 16) = 0.26$; $p = 0.7777$ for the average duration of fixations; $F(2, 16) = 1.23$; $p = 0.3179$ for the average distance between consecutive fixations. The effect of the condition thus remained unchanged whether or not the participants noticed that their gaze controlled the viewing window. The ANOVA on the total fixation time using the “eyes-mouth” factor did, however, yield a significant interaction between the condition and the answer to the last question [$F(2, 16) = 3.74$; $p = 0.0466$]. *Post hoc t*-tests revealed that participants who showed an awareness of controlling the viewing window were those for whom the total fixation time increased between the baseline condition (*Least Square M* = 145 *SE* = 37) and the experimental condition (*Least Square M* = 238 *SE* = 37).

DISCUSSION

The study presented here explores human abilities to monitor and gain awareness of self-generated gaze movements in social contexts. The results show that visual biofeedback of eye movements can be used to monitor one's own gaze behavior, even without the self-ascription of agency. The analyses of eye-tracking variables converge in showing that when the gaze-controlled viewing window was set on, the number of fixations decreased, the average duration of fixations increased, and the average distance between consecutive fixations decreased. In other words, the gaze was stabilized in the experimental condition. The medium-to-large effect sizes of these variations signal important changes in gaze behavior. Notwithstanding, the analyses indicate that this stabilization of gaze was independent from the declared self-ascription of the viewing window's movements. As expected, the results showed that the participants focused predominantly on the face of the virtual characters and that the gaze-controlled viewing window did not impair their compliance with the task. Consistently, the scores and response times did not show an increase or decrease specific to the experimental condition; but rather, they indicated that the performances improved over time.

Although the participants adapted their gaze behavior to the viewing window feedback, only half of them realized that they were controlling it. Literature echoes this incomplete awareness of one's own gaze behavior. For instance, experiments on anti-saccade tasks, where the participants are instructed to glance in the opposite direction of an impending cue, reveal that participants are unaware of half of their errors (Hutton, 2008). Poletti et al. (2010) have reported an experiment using eye-tracking where, paradoxically, the participants accurately tracked a moving dot, which they reported to be stationary and accurately fixated a stationary dot, which they reported to be moving. Hsiao and Cottrell (2008) incidentally observed that participants were not aware that the display used in their experiment was gaze-contingent. Noteworthy, several participants in our experiment explicitly declared that they were looking at the clear areas during the experimental condition and still failed to take responsibility for the appearance of these areas, attributing them instead to

the computer. These participants seem to have operated a time reversal between their action and the ensuing visual sensation, as reported in experiments on motor-sensory recalibration (Stetson et al., 2006).

A pending issue pertains to the fact that the last question regarding awareness divided the group into two equal halves. A possible explanation for this dichotomy involves the fact that the feedback latency of the viewing window could represent a threshold for awareness. This interpretation is based on an analogy with Nielsen's paradigm (Nielsen, 1963) that has been extensively used to investigate the sense of agency (Jeannerod, 2003). In this experimental paradigm, the participants are asked to draw a straight line without having their own hand in sight. They are simultaneously presented with a visual feedback of their trajectory, in which, unbeknownst to them, a deviation is introduced. The participants automatically adjust their hand movement to compensate for the deviation. Yet, they become aware of the conflict only when the deviation exceeds a particular threshold angle, which is common to all healthy adults (Slachevsky et al., 2001). Similarly, we suspect the 100 ms latency between the movements of the eye and the repositioning of the viewing window to represent such a threshold for awareness. The feedback provided in our experiment could have yielded an ambiguous experience of agency. Such an ambiguous experience is likely to disrupt the capacity to adequately distinguish between externally versus internally produced actions, as suggested by Moore et al. (2010). It should be noted that external causation attribution involves different brain mechanisms than those involved in the self-attribution of agency (Sperduti et al., 2011).

The dissociation found in the present study between the effective changes in gaze behavior induced by the gaze-contingent viewing window and failures in the judgment of agency argues in favor of the two-step account of agency proposed by Synofzik et al. (2008). Conceivably, these two levels may function with relative independence in the case of gaze behaviors, at least in the social context that was tested here. Nonetheless, the outcomes of the present study do not rule out possible bottom-up processes that influence awareness and originate from sensorimotor signals, as supported by Kühn et al. (2011). They do, however, support the existence of a top-down pathway demonstrated by the fact that the participants who gained awareness of controlling the viewing window modified their gazing strategy. Indeed, contrasting with the unaware participants, these participants focused more on the eyes and mouth regions when the gaze-contingent viewing window was set on, thus enforcing a usually spontaneous visual scanning behavior that is specifically relevant to social interactions. Based on this observation, if the participants gain an awareness of controlling the viewing window, either on their own or because they were told so, our setup should drive them to pay more attention to the emotionally meaningful facial features.

Overall, these outcomes support the propensity of human beings to monitor and adapt to the visual consequences of their own gaze, as evidenced here using social face-to-face stimuli. Although we acknowledge that gaze stabilization induced by the viewing window could also occur with non-social stimuli (an issue that we plan to investigate in a separate experiment), the self-monitoring of gaze appears particularly relevant for social

interactions due to the eyes' active role in human communication (Emery, 2000; Frischen et al., 2007). Indeed, eye-contact plays an important role in social intercourse and is known to influence social cognitive processes, such as gender discrimination (Macrae et al., 2002), facial recognition (Hood et al., 2003), and visual search for faces (Senju et al., 2008). The present project raises issues that are relevant for a line of research that studies gaze not only as an unidirectional channel enabling the transmission of emotional information from one person to another, but also as a bi-directional channel where interaction is possible (Redcay et al., 2010; Schilbach et al., 2010; Wilms et al., 2010; Pfeiffer et al., 2011). Our study adds on previous knowledge by showing that self-monitoring can function relatively independently from the judgment of agency in the case of gaze. Such a finding questions the role of awareness in social gaze as it suggests that at least some gazing behaviors could occur with an inadequate judgment of agency. Two tentative explanations not mutually exclusive can be speculated. First, the swiftness of gaze could be inappropriate for the time scale of conscious processing in the context of social face-to-face interactions. Second, an explicit consciousness of what originates from the self and what does not, may be less critical for the "language of the eyes" than for other communication channels such as verbal language.

This study raises several questions that require further investigation. First, as mentioned earlier, the neural substrates responsible for the self-monitoring of gaze are still unknown, although potential candidates have been identified. Second, the latency of the visual feedback could have a determining influence on the sense of agency and needs to be more precisely examined. Third, the present study focused on social gaze and the experimental material was exclusively based on social stimuli. In the future, we plan to conduct a similar experiment using non-social stimuli to verify whether the findings are specific to social contexts. The outcomes of the new methodological approach presented here seem relevant for a range of different scientific domains. For instance, in the field of computer sciences, researchers working

on gaze-contingent graphic displays should take into account the fact that gaze-controlled visual feedback can induce behaviors, which the user may not necessarily be aware of. Our research also highlights issues pertaining to social gaze control, which could prove particularly informative in psychopathology. Indeed, peculiar patterns of gaze behavior in social context are frequently associated with autism spectrum disorders (ASD) (Klin et al., 2002; Noris et al., 2011) and gaze control impairments have been demonstrated in schizophrenia and affective disorders (Lencer et al., 2010). Interesting experimental designs for these populations could be derived from the fact that normal participants who were aware of controlling the viewing window tended to focus more on the eyes and mouth (Grynszpan et al., 2011). Using the experimental method and platform described here, we showed in another article (Grynszpan et al., in press), that high functioning autism spectrum disorder (HFASD) is associated with alterations in the self-monitoring of gaze and judgment of agency. Additionally, we found that the gaze-contingent viewing window induced gaze behaviors for which social disambiguating scores on the task were linked to the time participants with HFASD spent looking at the virtual characters' faces. Thus, training based on our platform could conceivably be used to foster relevant visual explorations of faces during social interactions in individuals with HFASD. Further research on social gaze control may eventually generate new knowledge regarding these disorders, which in turn could be used to enhance treatment approaches.

ACKNOWLEDGMENTS

This work was supported by a grant provided by a partnership between the two following foundations: Fondation de France and Fondation Adrienne et Pierre Sommer (Project #2007 005874). We are thankful to Pauline Bailleul and Daniel Gepner for their assistance during parts of the experimental study. We are particularly thoughtful of Noëlle Carbonell, one of the project's initiators, who passed away before the project was completed.

REFERENCES

- Baron-Cohen, S., Wheelwright, S., and Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger Syndrome. *Vis. Cogn.* 4, 311–331.
- Baugh, L. A., and Marotta, J. J. (2007). A new window into the interactions between perception and action. *J. Neurosci. Methods* 160, 128–134.
- Bayliss, A. P., di Pellegrino, G., and Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *Q. J. Exp. Psychol. A* 58, 631–650.
- Bays, P. M., and Husain, M. (2007). Spatial remapping of the visual world across saccades. *Neuroreport* 18, 1207–1213.
- Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.
- Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends Cogn. Sci.* 6, 237–242.
- Buisine, S., Wang, Y., and Grynszpan, O. (2010). Empirical investigation of the temporal relations between speech and facial expressions of emotion. *J. Multimodal User Interf.* 3, 263–270.
- Colby, C. L., and Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349.
- Ekman, P., and Friesen, W. V. (1975). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Englewood Cliffs, NJ: Prentice-Hall Inc.
- Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* 24, 581–604.
- Fournier, P., Paillard, J., Lamarre, Y., Cole, J., and Jeannerod, M. (2002). Lack of conscious recognition of one's own actions in a haptically deafferented patient. *Neuroreport* 13, 541–547.
- Frischen, A., Bayliss, A. P., and Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychol. Bull.* 133, 694–724.
- Goodale, M. A., Jakobson, L. S., Milner, A. D., Perrett, D. I., Benson, P. J., and Hietanen, J. K. (1994). The nature and limits of orientation and pattern processing supporting visuomotor control in a visual form agnostic. *J. Cogn. Neurosci.* 6, 46–56.
- Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., and Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *Int. J. Hum. Comput. Stud.* 67, 842–849.
- Grynszpan, O., Nadel, J., Constant, J., Le Barillier, F., Carbonell, N., Simonin, J., Martin, J. C., and Courgeon, M. (2011). A new virtual environment paradigm for high-functioning autism intended to help attentional disengagement in a social context. *J. Phys. Ther. Educ.* 25, 42–47.
- Grynszpan, O., Nadel, J., Martin, J. C., Simonin, J., Bailleul, P., Wang, Y., Gepner, D., Le Barillier, F., and Constant, J. (in press). Self-monitoring of gaze in high

- functioning autism. *J. Autism Dev. Disord.* doi: 10.1007/s10803-011-1404-9. [Epub ahead of print].
- Holst, E., and Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Die Naturwissenschaften* 37, 464–476.
- Hood, B. M., Macrae, C. N., Cole-Davies, V., and Dias, M. (2003). Eye remember you: the effects of gaze direction on face recognition in children and adults. *Dev. Sci.* 6, 67–71.
- Hsiao, J. H., and Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychol. Sci.* 19, 998–1006.
- Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain Cogn.* 68, 327–340.
- Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behav. Brain Res.* 142, 1–15.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* 59, 809–816.
- Kühn, S., Nenchev, I., Haggard, P., Brass, M., Gallinat, J., and Voss, M. (2011). Whodunnit? Electrophysiological correlates of agency judgements. *PLoS One* 6: e28657. doi: 10.1371/journal.pone.0028657
- Lencer, R., Reilly, J. L., Harris, M. S., Sprenger, A., Keshavan, M. S., and Sweeney, J. A. (2010). Sensorimotor transformation deficits for smooth pursuit in first-episode affective psychoses and schizophrenia. *Biol. Psychiatry* 67, 217–223.
- Macrae, C. N., Hood, B. M., Milne, A. B., Rowe, A. C., and Mason, M. F. (2002). Are you looking at me? Eye gaze and person perception. *Psychol. Sci.* 13, 460–464.
- McConkie, G. W., and Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* 17, 578–586.
- Merriam, E. P., Genovese, C. R., and Colby, C. L. (2003). Spatial updating in human parietal cortex. *Neuron* 39, 361–373.
- Merriam, E. P., Genovese, C. R., and Colby, C. L. (2007). Remapping in human visual cortex. *J. Neurophysiol.* 97, 1738–1755.
- Mertens, I., Siegmund, H., and Grüsser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia* 31, 989–998.
- Meyer, C., and Girke, F. (2011). *The Rhetorical Emergence of Culture*. Oxford: Berghahn.
- Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., and Haggard, P. (2010). Disrupting the experience of control in the human brain: pre-supplementary motor area contributes to the sense of agency. *Proc. Biol. Sci.* 277, 2503–2509.
- Nielsen, T. I. (1963). Volition: a new experimental approach. *Scand. J. Psychol.* 4, 225–230.
- Noris, B., Barker, M., Hentsch, F., Ansermet, F., Nadel, J., and Billard, A. (2011). “Measuring gaze of children with autism spectrum disorders in naturalistic interactions,” in *Proceedings of Engineering in Medicine and Biology Society (EMBC '11)*, 5356–5359.
- Pfeiffer, U. J., Timmermans, B., Bente, G., Vogeley, K., and Schilbach, L. (2011). A non-verbal Turing test: differentiating mind from machine in gaze-based social interaction. *PLoS One* 6: e27591. doi: 10.1371/journal.pone.0027591
- Poletti, M., Listorti, C., and Rucci, M. (2010). Stability of the visual world during eye drift. *J. Neurosci.* 30, 11143–11150.
- Redcay, E., Dodel-Feder, D., Pearrow, M. J., Mavros, P. L., Kleiner, M., Gabrieli, J. D. E., and Saxe, R. (2010). Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. *Neuroimage* 50, 1639–1647.
- Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., Shah, N. J., Fink, G. R., and Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.
- Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn. Sci.* 13, 127–134.
- Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., and Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain Cogn.* 67, 127–139.
- Sirigu, A., Daprati, E., Pradat-Diehl, P., Franck, N., and Jeannerod, M. (1999). Perception of self-generated movement following left parietal lesion. *Brain* 122, 1867–1874.
- Slachevsky, A., Pillon, B., Fournier, P., Pradat-Diehl, P., Jeannerod, M., and Dubois, B. (2001). Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *J. Cogn. Neurosci.* 13, 332–340.
- Spengler, S., von Cramon, D. Y., and Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *Neuroimage* 46, 290–298.
- Sperduti, M., Delaveau, P., Fossati, P., and Nadel, J. (2011). Different brain structures related to self- and external-agency attribution: a brief review and meta-analysis. *Brain Struct. Funct.* 216, 151–157.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychol.* 43, 482–489.
- Spezio, M. L., Adolphs, R., Hurley, R., and Piven, J. (2007). Abnormal use of facial information in high-functioning autism. *J. Autism Dev. Disord.* 37, 929–939.
- Stetson, C., Cui, X., Montague, P. R., and Eagleman, D. M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. *Neuron* 51, 651–659.
- Synofzik, M., Vosgerau, G., and Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Conscious. Cogn.* 17, 219–239.
- Voss, M., Moore, J., Hauser, M., Gallinat, J., Heinz, A., and Haggard, P. (2010). Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* 133, 3104–3112.
- Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., and Vogeley, K. (2010). It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98.
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.
- Wolpert, D. M., and Miall, R. C. (1996). Forward models for physiological motor control. *Neural Netw.* 9, 1265–1279.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Res.* 48, 2070–2089.
- Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Arch. Clin. Neuropsychol.* 16, 653–667.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 December 2011; accepted: 02 April 2012; published online: 20 April 2012.

Citation: Grynszpan O, Simonin J, Martin J-C and Nadel J (2012) Investigating social gaze as an action-perception online performance. *Front. Hum. Neurosci.* 6:94. doi: 10.3389/fnhum.2012.00094

Copyright © 2012 Grynszpan, Simonin, Martin and Nadel. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.