Eye to I: Males Recognize Own Eye Movements, Females Inhibit Recognition

Sanjay Chandrasekharan, Geetanjali Date, Prajakt Pande, Jeenath Rahaman, Rafikh Shaikh, Anveshna Srivastava, Nisheeth Srivastava, Harshit Agrawal

Abstract

Studies show that people can recognize their own movements, such as their own walking (presented in silhouette using point lights), their own drawing (presented as a moving point light), own clapping, and their own piano playing. We extend this result to proprioceptive control, showing that people can recognize their own eye movements, when presented as just a point moving against a black background. Eye movements were recorded using a wearable eye tracking glass, while participants executed four tasks. A week later, participants were shown these videos, alongside another person's videos, for each task, and asked to recognize their own movements. Males recognized their own eye movements significantly above chance, but only for tasks with large and familiar body movements. Females performed below chance in these tasks. We argue that the standard common coding/motor simulation model does not account for this result, and propose an extension where eye movements and body movements are strongly coupled. In this model, eye movements automatically trigger covert motor activation, and thus participate directly in motor planning, simulations and the sense of agency.

Keywords: Self-recognition, Eye movements, Common coding, Motor simulation, Oculo-motor coupling, Agency

Introduction

The ability to recognize oneself is a central component of self-awareness. Many studies have examined the evolutionary and developmental origins of self-recognition, particularly the ability to recognize oneself in a mirror, which has been studied in the case of different animals (Gallup, Anderson & Shillito, 2002) as well as human babies (Bertenthal & Fischer, 1978; Lewis & Brooks-Gunn, 1979). Another approach to understanding self-recognition involves studying the way people recognize their own faces (Tsakiris, 2008), particularly the neural mechanisms involved in this process (Deveu et al., 2007).

A third approach to study self-recognition is based on recording people's movements, and presenting sparse versions of these movements, to examine whether people can recognize their own movements, when presented next to others' movements (Loula et al., 2005). An influential experiment (Johansson, 1973) created 'point light walkers' by attaching lights to participants' joints, and filming their walking in a dark room. When presented sets of such point light videos, with one encoding their own movements and another encoding someone else's movements, participants could recognize their own movements. Extending this approach, Knoblich and Prinz (2001) showed that people can recognize their own handwriting traced by a moving dot of light, and their own clapping from a set of recordings of clapping (Flach, Knoblich & Prinz, 2004). Similarly, pianists can pick out their own rendition of a piece from a set of recordings of the same piece (Repp & Knoblich, 2004). People can also recognize their own manipulation of a puppet (Mazalek et al., 2009), as well as virtual avatars that encode their own movements (Mazalek et al., 2010).

This type of self-recognition is explained by the theory of common coding (Prinz, 1992; 1997; Hommel et al., 2001), which postulates that execution, perception and imagination of movements share a common code at the neural level. This code leads to the automatic, but covert, activation of the motor system when perceiving biological movements. This covert activation of the motor system (or simulation) allows the participant to judge which encountered movement is more familiar, and this familiar movement is then identified as one's own movement.

In the study reported here, we extend this line of research in two ways. One, we investigated whether the self-recognition effect holds for proprioceptive control (Donaldson, 2000), by examining recognition of own eye-movements when presented in a format similar to the point light walker, where the eye movements made during tasks is displayed using a red dot moving in a dark background. Results show that people can recognize their own eye movements, but the recognition response is different for males and females. Second, we argue that the common coding/simulation account is insufficient to explain our results, and propose a related model, where overt eye movements trigger covert body movements.

Experiment Design

Briefly, we recorded eye movements of participants using a wearable eye tracking glass (Tobii), while they executed four actions. Two of the actions were familiar (walking, climbing) and involved systematic eye movements in relation to whole body movements. The other two actions (walking with one leg tied to another person's leg, shading different sized circles in a sequence) were chosen to minimize the systematic connection between eye movement and whole-body movement. These tasks were chosen based on pilot testing, where participants were shown actual scene videos (i.e. the world as seen by the wearer of the eye tracking glass) generated by the tracker software. The eye movements were superimposed in this video scene as a moving red dot. Participants could identify their own eye movements, as well as others' movements, in these scene
videos. Interviews suggested that this was achieved by identifying the walking style of the participant in these videos, particularly the head movement patterns, which are encoded in the way the scene moves ('bounces') in these videos, as the external world moves in tandem with the head movements. To remove these body movement cues, we superimposed the eye movements on a black background, thus removing all scene movement information. These videos minimize head movement cues. However, it was felt that some systematic body movement may still be embedded in the eye movement patterns, and this implicit body movement could be used as a cue for recognizing one's own movement. Two tasks (3-leg, drawing) were developed as controls to address this issue, as these tasks disrupt/minimize the connection to full body movement.

A week after recording the videos using the eye tracking glass, participants did a 2-alternative forced choice task, where two black background videos were displayed side by side in each trial (one showing their own eye movements, the other showing another person's eye movements). Participants had to identify their own eye movements.

Materials and Methods

The design of the study followed the standard format of the earlier self-recognition studies (Knoblich & Prinz, 2001, Mazalek et al., 2009; 2010). All participants individually completed two separate protocols, a recording block and a test block, with an intervening interval of 7-12 days. In the recording block, they completed a set of four actions:

1. walking in a corridor (walk)
2. climbing four flights of stairs (climb)
3. walking in a corridor, with one leg tied to an experimenter's leg (3-leg)
4. shading differentially sized circles on an A3 sheet with a pencil (draw)

The primary recording was done without any instruction on the details of the experiment. After completing the four actions once without any instruction, participants were selected randomly to receive one of two instructions:

A) be aware of how your eyes are moving as you perform these tasks

B) next week, we will ask you to try and pick out your eye movements from two sequences of eye movements

Participants given instruction (B) were also shown a demo of the recognition task performed in the test block (Figure 1). After receiving one of the instructions, participants were asked to complete all four tasks again, remaining mindful of the instructions they had received.

This condition explored the role of instruction, if any, in identifying one's own eye movements. Knowing about the recognition task in advance provided participants the option of laying down eye movement markers if they wished, and then do the recognition explicitly, based on these markers. If such an explicit strategy is used, and it is effective, accuracy in self-recognition would be very high for the videos recorded with instruction. However, given the absence of scene elements in the black background videos, it would be very difficult for participants to refer to and track any markers to identify their own eye movements.

In the test block, run after a week, each participant was first shown a demo, where two videos were shown, and the experimenter showed how to select a video using keyboard input (Q for left video, P for right video). Participants were instructed that their task was to select the video that showed their own movement. Once a participant indicated understanding of the setup, we showed them 8 recordings (4 tasks x 2 instruction conditions) of their eye movements alongside those of another participant randomly selected from our participant pool (a new contrast participant picked for each of the 8 trials). The relative position of the videos was selected as the outcome of a Bernoulli trial (p = 0.5). Videos recorded without instruction were shown first in the block, followed by videos recorded with instruction.

The videos participants saw were generated with a C program, using screen coordinates of gaze-points detected by the tracker. The program determined where on the screen a circle sized 20 pixels would be drawn across a series of frames. We sampled the frame rate of the videos to synchronize with the 30 fps rate of the tracker, to ensure that the eye movements retained their original timing in the video. Videos were looped indefinitely until the participant was ready to make a choice. (see videos at this link: http://gnowlke.org/~sanjay/LSRCogsci_2014/)

The study was run in two phases, an exploratory phase, and a testing phase. In the exploratory phase, we ran 20 participants (10 males, 10 females), and analysed the data. This analysis identified a clear gender difference in self-recognition. The testing phase, with another 22 participants (11 males, 11 females), was run to test the robustness of this effect. We report the combined data from the two phases, as the results were similar for both phases. We also combine the instruction and no-instruction data, as there was no significant difference between the two conditions.

Participants

Across the two phases, we recruited 21 males (mean age 23.6, S.D. 6.5 years) and 21 females (mean age 22.7, S.D. 4.9 years), with uncorrected normal vision. Informed consent was obtained from all participants. During testing,
one of the female participants reported physical discomfort and withdrew consent for participating in the experiment. Another participated in part 1 of the experiment (video recording), but could not participate in the second part (recognition). Thus, our final sample contained 40 participants (20 males, 20 females). The response time data for two participants was lost due to a computer problem, so this data is reported only for 38 participants. All participants were volunteers, and were students or staff members of our institute. All were originally naive as to the purpose of the study, and with no previous experience with eye tracking.

**Apparatus**

For recording eye-tracking data, we used the Tobii Glasses system, which is a lightweight (75 gms) wearable eye-tracker (Figure 2). Participants performed the recognition task on a custom application designed using UNIX shell scripts (Figure 1) running on a 15” screen laptop. Data collection and analysis was done using GNU Octave.

![Figure 2: Tobii Eyetracking Glasses, used to record eye movements in the study (Source: Tobii.com)](image)

**Data collection and analysis**

We recorded accuracy and response time data for each of the eight recognition trials. To ensure reliability, response time was recorded using UNIX system calls.

Statistical significance for testing whether a particular sample performed better than chance on any subset of the tasks was assessed against a binomial distribution generated using an identical number of trials as the sample size. Chi-square proportion testing was used to differentiate performance between pairs of samples.

**Results**

Our overall sample performs almost exactly as by chance (accuracy mean % = 48.1). However, the histogram of performance quality is not binomially distributed as one would expect from a random outcome (see Figure 3).

The true distribution of outcomes we obtained suggested a bimodal generative process, viz., there might be two subpopulations in our sample, one of which is able to recognize their gaze data, and the other not. Such an inference, however, is susceptible to the possibility of over-fitting the data. One way to disambiguate would be to identify discriminative features of these hypothesized sub-populations. When we tried to do so, we immediately encountered a strong gender effect in our data.

![Figure 3: Deviation of the performance of our participant sample from random behavior predicted by the null hypothesis. We plot a histogram of the number of participants getting different number of trials right (max = 8) against a plot showing the baseline binomial distribution we would obtain from 1000 repetitions of 40 Bernoulli trials. Error bars are 2 SD across.](image)

As Table 1 shows, males recognize themselves in the walk and climb tasks (actions involving a close connection between eye movements and whole-body movement) at rates significantly better than chance. Females perform below chance, significantly for the climb and 3-leg task, and at trend level in the walk task.

Males took more time than females overall in making the recognition decisions (p=0.03), but this difference was not significant for individual tasks.

![Figure 4: The gender effect in the recognition task. Males clearly outperform females (p<0.005), and a random baseline, in identifying themselves from prior recordings of their eye movements. Error bars are 2 SD across.](image)
As Figure 4 shows, males significantly outperform females (p<0.005) in recognizing their own eye movements. The difference in these two populations accounts largely for the deviation from chance behavior seen in Figure 3.

Table 1: Average accuracy for male and female cohorts in the recognition task

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk</td>
<td>65.00%*</td>
<td>#37.50%</td>
</tr>
<tr>
<td>Climb</td>
<td>67.50%**</td>
<td>*30.00%</td>
</tr>
<tr>
<td>3-leg</td>
<td>42.50%</td>
<td>**30.00%</td>
</tr>
<tr>
<td>Draw</td>
<td>47.50%</td>
<td>55.00%</td>
</tr>
<tr>
<td>* p=0.013</td>
<td>* p=0.04</td>
<td></td>
</tr>
<tr>
<td>** p=0.0008</td>
<td>** p=0.03</td>
<td></td>
</tr>
<tr>
<td># p=0.14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Discussion

Our results show that:

1) Males recognize their own eye movements significantly above chance in the walk and climb tasks.

2) The same participants cannot recognize their own eye movements in the 3-leg and drawing tasks.

3) Females perform below chance, significantly in the climb and 3-leg tasks, and at trend level in the walk task.

4) Females take less time than males to make a decision.

We will first discuss results 1 and 2. One possible account for this pattern of results would be a strictly localist model, where the eye is treated as a standalone movement system, and proprioception of the eye muscles (Donaldson, 2000) is the possible mechanism involved in the self/other judgment. In this view, the eye moves in specific patterns while executing the actions during the recording session. When the gaze-point data is played back as a video in the choice task, similar eye movement patterns are overtly activated while watching each gaze point video, as watching the gaze point move recruits smooth pursuit. One of the overt eye movement patterns activated by the gaze point videos appear familiar to the participant, based on previous proprioceptive experience. This video is then identified as one's own movement. Note that this account does not assume common coding, and the associated covert activation of the motor system, as eye movements are overtly executed during the choice task, and the familiarity judgment is based on this overt movement.

However, in this account, participants would be expected to recognize their own eye movements in the three-leg and the drawing tasks as well, because the eye would move in familiar patterns for their own video, for all the tasks. Since our results show recognition above chance for males only in the walking and climbing tasks, and not in the 3-leg and drawing tasks, this account does not explain our results.

Remember that the latter two tasks (3-leg, drawing) minimize systematic head movement patterns, and were developed as controls to address the issue of embedded head movements in the eye movement videos. The inability of male participants to identify themselves in these tasks suggests that information about body movement, particularly head movement, is used, implicitly, while recognizing one's own eye movements in the walk and climb conditions. This means a mechanism that can access body movements from eye movements is required to account for our results.

A good candidate mechanism is common coding (Prinz, 1992; 1997; Hommel et al., 2001), which postulates a common representation at the neural level, connecting execution, perception and imagination of movements. Self-recognition effects are explained by the covert activation of familiar (one's own) motor patterns, which are triggered, via common coding, by the perception of movement.

In all the cases of self-recognition where common coding is provided as the explanation (Knoblich & Prinz, 2001; Flach, Knoblich, & Prinz, 2004; Repp & Knoblich, 2004; Mazalek et al., 2009; 2010), the choice task involves perceiving a biological movement. In our case, the perception also leads to the actuator (the eye) moving overtly, in patterns similar to the movements originally executed during the recording phase. The motor activation is overt, but this overt replay of executed eye movements is not enough for males to identify one's own eye movement in every case, as the recognition happens only in the walking and climbing tasks, where the eye movements occur with systematic body movements. This suggests full body movements are accessed, and they may even be required, to identify one's own eye movements.

How could full body movements be accessed via the overt activation of eye movements? Common coding theory does not provide an account of such a mechanism. To extend common coding theory to include such a mechanism, we propose that this is achieved by a two-way oculo-motor coupling, where overt eye movements trigger covert motor plans, and, in the other direction, planned motor movements trigger compatible eye movements. In this view, body movements in response to dynamic environmental stimuli (such as catching a suddenly thrown ball) are eye movements 'writ large', so to speak, as the pattern of eye movement (such as smooth pursuit) generated by the dynamic stimuli provides real-time, precise, often scaled, information for the motor plan. In the other direction, planned motor movements (such as inserting a door key) lead to 'orienting' eye movements, which can act as forward models that help plan and execute fine motor movements.

A crude analogy for this two-way coupling could be a pantograph, which allows a small figure traced using a pen to be automatically converted to a large figure traced using another pen, coupled to the first pen using a parallelogram structure (Figure 5). The pantograph can convert large drawings to small ones as well. This system only provides scaling, and is thus not a good analogy for complex control. A more sophisticated analogy for the coupling between body movements and eye movements would be Watt's Centrifugal Governor (van Gelder, 1997), a dynamic control
system that mechanically regulates the speed of the steam engine. van Gelder (1997) proposed the Watt Governor as a model of the mind, arguing for a dynamic systems approach to cognition. Our proposal is inspired by this model, but combines it with the imagery and representation possibilities of common coding, which are based on covert and off-line activation of the common code. The overt eye movements thus work as an embodied emulator (Grush, 2004).

Figure 5: A pantograph (Source: Wikipedia)

In our proposal, the eye functions as a physical micro-simulator that is coupled to the external world in real-time, similar to the Watt Governor. The state of the world activates the eye overtly, and this movement provides precise and real-time parameters to the motor system for environment-driven motor plans. For intended actions, i.e. actions driven by the self and not by the environment, the eye again moves overtly, but in micro-simulator mode, providing forward models (Wolpert & Kawato, 1998).

These overt forward models allow more precise and detailed predictions than imagined models of movement. It is possible that some of the parameters of the imagined movement are set using the parameters of the overt eye movements. Other parameters could be set by optical flow (Gibson, 1950), which is also modulated by eye movements.

The eye-as-micro-simulator model extends the common coding proposal, by outlining one specific mechanism, where visual pursuit of movement triggers motor activation. In the other direction, the model predicts that imagination of movement, required for intended actions, would be accompanied by eye movement patterns similar to execution/perception of movements. There is significant evidence for eye movement during imagination (Johansson, Holsanova & Holmqvist, 2006; Johansson & Johansson, 2014). Performance in insight problem-solving improves when participants are made to implicitly generate eye movements related to a solution (Thomas & Lleras, 2007). Related work shows that making eye movements influence aesthetic judgments (Topolinski, 2010), and the eye pupil adjusts to imagined light (Laeng & Sulutvedt, 2014).

The eye-as-micro-simulator account explains our results well. In the walking and climbing tasks, watching the videos trigger overt eye movement patterns, which automatically generate covert body movements. As the covert motor activations proceed over time, the body movement triggered by one of the videos appear familiar to participants, and this video is chosen as one's own eye movement.

In the three-leg task as well, motor activation is generated by both the videos. But as the activations proceed, the familiar motor pattern doesn't rise up consistently, as the eye movements in this case are patchy, as they are influenced by the other person's movements. This lowers familiarity for the own video, leading to chance performance.

Why do females perform below chance, significantly in the climb and 3-leg tasks, and at trend level in the walk task? In the eye-as-micro-simulator account, they would be expected to do well in the walk and climb cases, as the proposed two-way oculo-motor coupling is a basic psycho-physical process, and would be similar in females.

We propose that the automatic covert activation of the motor system by eye movements happens for females as well, for both videos. However, only for the self video, further activation of the covert process is blocked, and not allowed to proceed, by an inhibition signal. This inhibition of the covert body movement leads to zero motor system response for the self video, but some from the other video, which leads to the other video being consistently chosen. This explains the consistent below chance performance in the 3-leg case: here also both the videos lead to covert motor activation, and the patchy covert activation from one's own video (modulated by another person's movement) is blocked, leading to zero motor activation for this video. This leads to the other video being chosen consistently. As the comparison between no-motor-response and some-motor-response would be faster, this mechanism account also explains the lower response time for females.

Why do females block the motor activation for one's own movements from proceeding further? We propose that this is because of a strong bias to control full-fledged mimicry (Wang & Hamilton, 2012), and other signals of affiliation -- i.e. a bias towards not letting the body display signs of interest. The activation of the motor system by one's own eye movement possibly moves the motor activation too close to overt display, and this could be one reason why motor activation is blocked. A related reason could be eye movements activating the self (Baltazar et al., 2014), which could independently lead to the blocking, as situations where the self is activated would require caution. A parallel case is the lack of emotional arousal for recognized faces, which underlies the Capgras delusion (Ramachandran & Blakeslee, 1999), where a close relative or friend is replaced by an imposter. Interestingly, Capgras delusion occurs more frequently in females (Todd, Dewhurst & Wallis, 1981).

This account leads to a specific prediction about neural activation in our task: a continuing motor area activation for males, and a short activation and then inhibition of the motor area for females. This prediction is best tested using
an electro-physiology study using the same recognition paradigm. We are currently developing such an experiment.

It is possible that the inhibition response we report for females is specific to our study population, as the study was done in a cultural milieu where overt physical signaling by females is discouraged. Cross-cultural studies are needed to test whether the effect is specific to cultural environments where females are cautious about overt physical responses. If yes, this specificity would provide a window to explore how the oculo-motor coupling is tuned by cultural norms.

Acknowledgments

We gratefully acknowledge the support of Department of Science and Technology, Government of India, through a Cognitive Science Initiative grant (SR/CSI/186/2012). We thank our participants for volunteering for the study.

References


