Proceedings of the 2007 IEEE
International Conference on Robotics and Biomimetics
December 15 -18, 2007, Sanya, China

Experimental Analysis and Computational Simulation of the Attribution of Own Actions by the Multiple Forward Models

Mihoko Otake†, Kohei Arai†, Motoichiro Kato‡, Takaki Maeda‡ , Yusuke Ikemoto†
Kuniaki Kawabata*, Toshihisa Takagi† and Hajime Asama†

†Research into Artifacts, Center for Engineering, the University of Tokyo
5-1-5 Kashiwa-no-ha, Kashiwa-shi, 277-8568, Japan
E-mail: otake 'at' race.u-tokyo.ac.jp
‡Department of Neuropsychiatry, School of Medicine, Keio University
35 Shinanomachi, Shinjuku-ku, 160-0016, Japan
Distributed Adaptive Robotics Research Unit, RIKEN
2-1 Hirosawa, Wako, 351-0198, Japan

***Abstract*— Human cognitive mechanisms have been studied for designing user-friendly interface. One of the key issues is the attribution of own actions to the intention of self or others. It is known that patients with schizophrenia who sometimes attribute their own actions to the intentions of others may perceive themselves as causing events which they do not in fact control, when they feel they are in voluntary movement. In this study, we administered similar experiments to normal subjects. We also conducted computational simulations through extending the multiple forward models, which successfully described the experimental results.**

***Index Terms*— Forward model, schizophrenia, attribution of own actions, simulations, experimental analysis**

I. Introduction

Information systems are ubiquitous in the modern information society. People spend a long time in front of the personal computer using the typical input and output devices: keyboard, mouse, and display. It is very important to analyze how we adopt to this human-machine interface in order to understand cognitive mechanism of human, because such a mechanism provides us design principle of daily used human-machine interface. In general, people become comfortable using machines when they feel the machines as if they are part of their bodies. Let us imagine driving the car. We feel like cars as if they are part of our body when we get used to driving them. Drivers can estimate speeds or angles of the cars when they push the accelerators or rotate the handles. The drivers feel that they are controlling the cars when they can attribute motions of the cars to the intentions of themselves. Likewise, people feel comfortable using computers when they can attribute the motion of objects in the display, which are caused by their own key or mouse inputs, to their own intentions. Such feeling is called "sense of agency". To date, usability of typical human-machine interface with key board, mouse, and display have not been analyzed in terms of the sense of agency. In this study, we experimentally analyze the sense of agency. Studies on sense of agency are overviewed, and experimental results are simulated by the multiple forward models. Our model is validated through computational simulation.

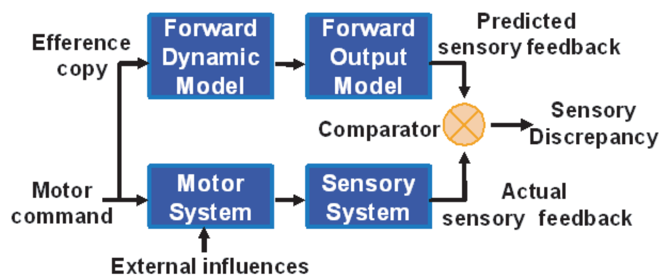

Fig. 1. The forward model of motor control: forward model contains a forward dynamic model which predicts the consequences of motor commands and a forward output model which makes a prediction of the sensory consequences of motor commands, which is compared with the actual consequences of movement from sensory system. [1]

II. Sense of Agency and Schizophrenia

Sense of agency has been studied in schizophrenia research, because several psychotic symptoms with schizophrenia could result from dysfunction of the awareness of one's own action as well as recognition of actions performed by others. It is possible that the elucidation of such symptoms offer a framework for studying the determinants of agency. In particular, paranoid schizophrenia has a variety of delusions or hallucinations. Delusions are bizarre and false beliefs and hallucinations are unreal perceptions of the environment, which can be auditory, visual, olfactory, and tactile. Delusions and hallucinations can be described as dysfunction of the sense of agency. For example, patients with delusions of alien control, or passivity experiences may misattribute self-generated actions to an external source. In line with this phenomenon, patients with auditory hallucinations misattribute self-generated internal speech to external others' voice. Frank et al. investigated the possibility that delusions of influence

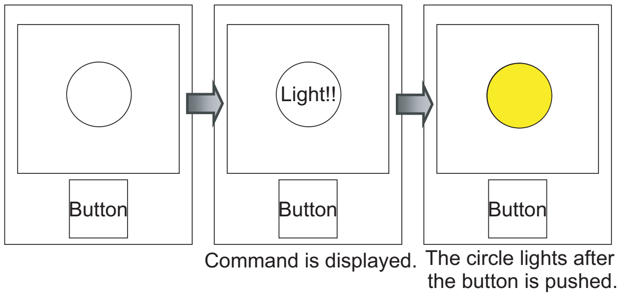

Fig. 2. A graphical user interface for the sense of agency experiment

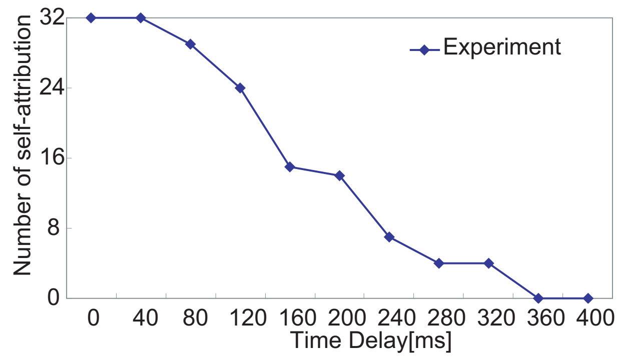

Fig. 3. Relationship between the number of self attribution and time delay: sum of 8 subjects

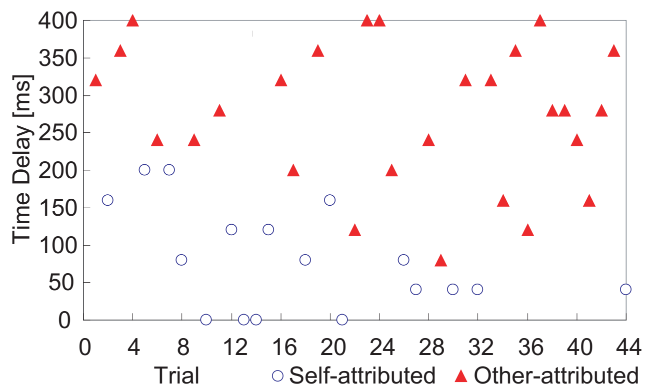

Fig. 4. Attribution of one subject on each trial: x-coordinate is a number of trials; y-coordinate is time delay. Circle dots are self-attributions while the triangle dots are others-attributions.

could be related to abnormal recognition of one's own actions in individuals with schizophrenia [2]. Subjects executed discrete movements in different directions holding joystick. The image of a virtual hand holding a joystick was presented on a screen with angular biases and time delays. Patients with schizophrenia made significantly more recognition errors in trials with time delays, compared with normal subjects. In the experiment by Daprati et al., hallucinating and deluded patients with schizophrenia tended to over-attribute the alien hand to themselves [3]. The two brain areas, the inferior part of the parietal lobe and the insula, appeared to be modulated by the degree of discrepancy, which has been investigated through brain imaging studies [4][5]. Diverse models such as Forward model [1] and Who system [6] have been proposed which describe the mechanism of the sense of agency. In this study, we focus on normal subjects rather than patients with schizophrenia in order to understand the sense of agency characteristics of normal subjects for designing efficient human-machine interface, because previous studies have been focusing on patients with schizophrenia, and the performances of normal subjects were not tested in detail.

III. Forward Model

Diverse hypotheses proposed to describe the sense of agency. In the present study, we hypothesize a neural mechanism of forward model in the brain. Forward model could qualitatively describe the sense of agency. Forward model contains a dynamic model which predicts the consequences of motor commands and an output model which makes a prediction of the sensory consequences of motor commands, which is compared with the actual consequences of movement from sensory system. Blakemore et al. [1] proposed that an underlying dysfunction in the forward model may lead to the lack of the sense of agency, and consequently the development of delusions of control in patients with schizophrenia (Fig. 1). It was not clear, however, how the forward model predicts the consequences of actions which are necessary for attribution of actions to the intention of self or others. A purpose of this study is to propose a quantitative forward model which can simulate the experimental results of sense of agency, in order to understand its mechanism. We conduct the psychological experiments for normal subjects which originally have been given for patients with schizophrenia. In this paper, the characteristics of the attribution of actions to self or others are explored, and the novel forward models are developed whose simulation results are comparable to experimental results.

IV. Experiment

A. Method

1) Subjects: Eight normal subjects (6 men and 2 women; mean age= 22.9 years, SD=16.8) participated in the study. Subjects were naive about the purpose of the study procedure. After a complete instruction of the procedures, all participants provided informed consent.

2) Materials: During the study, the image of a circle and a button were presented to the subjects on a computer screen with a high refresh rate. After the button is pushed by the subjects, the circle lights and the color of the circle turns yellow (Fig. 2). Time delays could be introduced in this representation.

3) Procedure: Subjects sat comfortably in front of the computer screen. They held the mouse button with a hand. The task consisted of executing a series of simple movements with the mouse button. Each trial started with a white screen with a white circle and a white button. The instruction

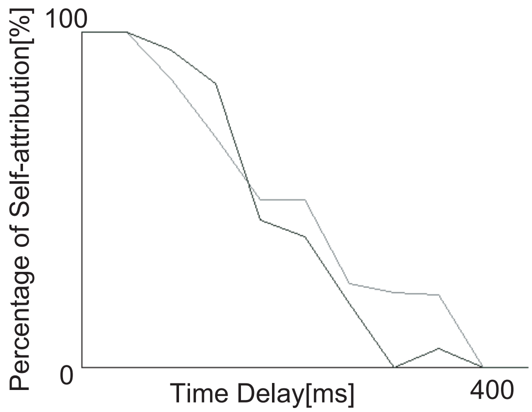

Fig. 5. The percentage of self-attribution on each time delay in the first 22 trials and the last 22 trials (sum of 8 subjects : the first 22 in gray！$the last 22 in black)

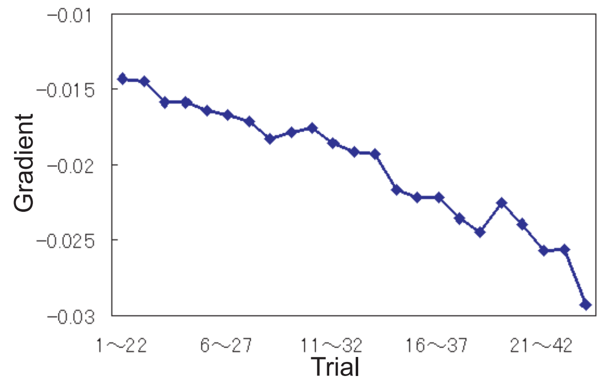

Fig. 6. The gradient of the line which is produced by logistic regression analysis on every 22 trials

”Light!” is displayed at the center of the circle. The subject push the button and the color of the circle become yellow with delays. Immediately after the trial, subjects had to answer the following question with a yes-or-no response: ”Who did switch on the light you saw on the screen?”. The buttons with the term ”self” or ”computer” appear on the screen and the subjects select each of them. Trials with time delays were used, in which the lighting of the circle were delayed by a given time (0, 40, 80, 120, 160, 200, 240, 280, 320, 360, and 400 [ms]). Thus, we prepared eleven kinds of time delays in total. Each trial was run four times for each of the 11 kinds of time delay. Each subject executed a total of 44 trials. The order of presentation of the 44 trials was randomized for each subject. Identical trials could not be presented one after the other.

4) Data Analysis: The responses of the subjects were recorded. The numbers of ”self” responses of subjects were counted by time delay of lighting on the screen. The percentage of ”self” responses from the first 22 trials (1 - 22 trials) to the last 22 trials (23 - 44 trials) for every successive 22 trials were calculated by time delay. The percentages of ”self” responses of every successive 22 trials were analyzed by logistic regression analysis. The gradients of logistic regressions of the percentage of ”self” responses were calculated.

B. Results

1) Descriptive Analysis: Figure 3 shows the number of ”self” responses or ”self-attribution” by subjects, who switched on the light they saw on the screen, as a function of time delay. There were total of 8 subjects and 4 trials for each time delay. Therefore, 32 responses was maximum number of ”self” responses while 0 was minimum number. Subjects gave ”self” responses in nearly all trials without time delay and small delay (40 [ms]). Subjects gave ”computer” responses in nearly all trials with large delay (360 - 400 [ms]). The number of ”self” responses was higher for the smaller time delay and decreased as the delay increased. Subjects showed a clear and sharp reduction in ”self” responses for a relatively small bias (120 - 160 [ms]).

2) Time Series Analysis: Time series of attribution of one representative subject on each trial is shown in Fig. 4. X-coordinate is a number of trial and Y-coordinate is a time delay. Circle dot is plotted with the time delay when the subject answered ”self”. Triangle dot is plotted with the time delay when the subject answered ”computer”, which means other-attribution. The time delay which discriminates self- or other-attribution is around from 200 to 240 [ms] in the early stage of total trials (0 to 16 trials). The time delay which discriminates self or other gradually decreases in time. It is about 80 to 120 [ms] in the end stage (28 to 44 trials). The discrimination criterion changes as the trial proceeds. This characteristic is observed for every subject.

Then, temporal change of the percentage of self-attribution was analyzed in the course of trials for the average of 8 subjects. The percentage of self-attribution was calculated by dividing the number of self-attribution by total number of trials for each time delay. The percentage of self-attribution for each time delay in the first 22 trials (plotted in dotted line) and the last 22 trials (plotted in line) are shown in Fig. 5. The percentage of the first 22 trials showed smooth decrease in ”self” responses while that of the last 22 trials showed sharp decrease in ”self” responses.

3) Logistic Regression Analysis: The percentages of self-attribution in every successive 22 trials which were sampaled by turns in 44 trials, were calculated by bringing out the percentages from i th trial to (i + 21) th trial, and shifted the number i from 1 to 23 recursively. They were analyzed by logistic regression analysis. The curves of the percentages were hypothesized as logit curves. They were converted with the following equation:

$$\mathrm{logit}(p) = \log\left(\frac{p}{1-p}\right). \tag{1}$$

The percentages of every 22 trials were substituted to the variable p. First-order approximations were calculated for each curve of every 22 trials. The gradient of the first-

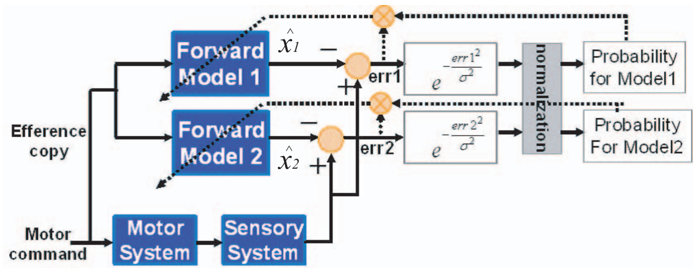

Fig. 7. Multiple forward models for sense of agency based on those for motor control [7]

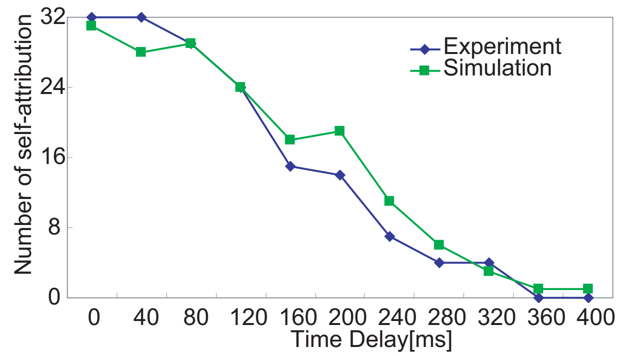

Fig. 8. Relationship between the number of self attribution and time delay: sum of 8 subjects. Square dots are simulation while rhombus dots are experimental results.

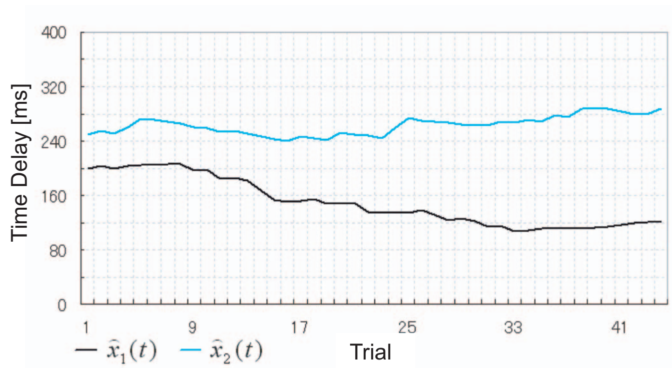

Fig. 9. Transition of predicted time delay: $\hat{x}_1(t)$ is the average time delay of the first forward model, self predictor. $\hat{x}_2(t)$ is the average time delay of the second forward model, others predictor.

order approximation of every 22 trials obtained by logistic regression are plotted in Fig. 6. X-coordinate is a number of the first 22 series of trials and Y-coordinate is a gradient of the first-order approximation of the 22 trials. The absolutes of the gradient of the first-order approximation increase in the course of trials. The discrimination criterion of self-attribution becomes clear in the course of trials.

V. Computer Simulation

A. Multiple forward models

Next, we focus on multiple forward models. The model is proposed for context estimations [8]. Multiple paired forward and inverse models were proposed for motor control in order to generate accurate and appropriate motor behavior under many different and often uncertain environmental conditions [7]. The condition which was presented by Wolpert et al. [8] could be applied to the attribution problem. The example condition in Wolpert's model is in which a teapot to be lifted is either full or empty. If we replace full for self and empty for others, our condition is modeled as almost the same. In our case, the multiple forward models contain two forward models, one predicts the stochastic distribution of time delay of self and another predicts that of other. The multiple forward models that we use contain only predictors, in other words, forward models. Originally, each prediction is pared with controller, in other words, inverse models. We removed the controller from the models because output behavior, button press does not change by the consequences of action in this study. Original unit that accumulates the output of predictors was also removed because which was only used for motor control. States of each predictor change by the errors of prediction and real feedback from the sensory system at each state (Fig. 7).

B. Simulated time delays of self and other attribution

The estimated delay of each forward model was simulated by using the experimental data. Initial average time delay of self-attribution was 200 [ms], and that of others-attribution was 250 [ms]. Variance of each predictor, σ_1 was 5122 for self-attribution and σ_2 was 6796 for others-attribution based on the experimental data of all delays for each decision.

Figure 8 shows the number of "self" responses or "self-attribution" by real and simulated subjects, who switched on the light they saw on the screen, as a function of time delay. There were total of 8 subjects and 4 trials for each time delay. Therefore, 32 responses was maximum number of "self" responses while 0 was minimum number. The number of "self" responses was higher for the smaller time delay and decreased as the delay increased for both simulation and experimental results.

Figure 9 shows the transition of predicted time delay. X-coordinate is a number of trial and y-coordinate is a time delay. $\hat{x}_1(t)$ is the average time delay of the first forward model, self predictor. $\hat{x}_2(t)$ is the average time delay of the second forward model, others predictor. The difference between the average time delay of the first and the second forward model presents clarity of the discrimination of self and others. Discrimination is clear when the difference is large. Discrimination is ambiguous when the difference is small.

Initial and final likelihoods of time delay of self and others are shown in Fig.10 and Fig. 11. Initial likelihoods of time delay of self $e^{-\frac{\epsilon_1(0)}{\sigma^2}}$ and others $e^{-\frac{\epsilon_2(0)}{\sigma^2}}$ before the course of trials were similar (Fig. 10), whereas final likelihoods of time delay of self $e^{-\frac{\epsilon_1(44)}{\sigma^2}}$ and others $e^{-\frac{\epsilon_2(44)}{\sigma^2}}$ after the course of

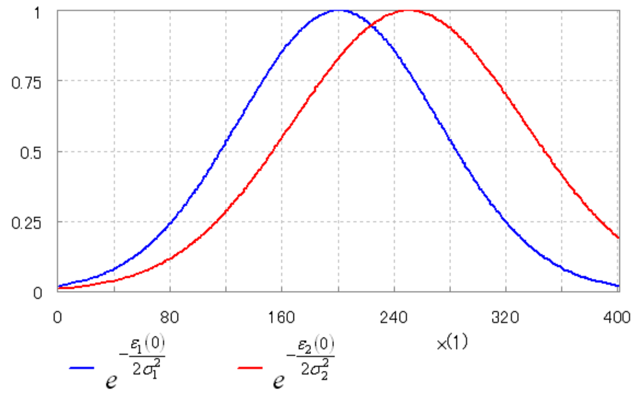

Fig. 10. Initial likelihood of time delay before the course of trials: $e^{-\frac{\epsilon_1(0)}{\sigma^2}}$ is the initial likelihood of self and $e^{-\frac{\epsilon_2(0)}{\sigma^2}}$ is the initial likelihood of others.

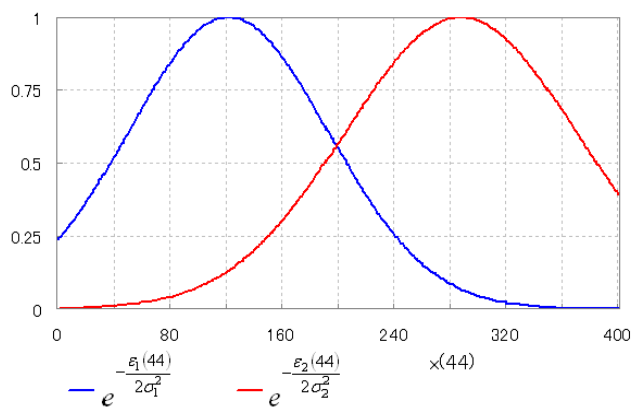

Fig. 11. Final likelihood of time delay after the course of trials: $e^{-\frac{\epsilon_1(44)}{\sigma^2}}$ is the final likelihood of self and $e^{-\frac{\epsilon_2(44)}{\sigma^2}}$ is the final likelihood of others.

trials were different (Fig. 11). Discrimination became clear after the course of trials.

We can see that:

- The discrimination criterion changes in time.
- The discrimination criterion of self-attribution becomes clear in the course of trials.

These characteristics are observed in the experimental results. Simulation results were well correspondent to the characteristics of the experimental results. Therefore, multiple forward models mechanisms were suggested to be valuable to explain the problems of the self- and others-attribution on the sense of agency, through the comparison of simulation and experimental results.

VI. Conclusion

We investigated the sense of agency of normal subjects through psychological experiments and computational simulation. Multiple forward models, which were originally proposed for motor control and social interaction, were introduced for describing the mechanism of attribution of actions. It has been believed that the sense of agency is particular to individuals independent of external input[2], and thought to be not applicable to the cognitive experiments. However, we carefully analyzed the temporal changes of the attribution criteria by logistic regression analysis. Experimental results suggest that the sense of agency is influenced by the series of input. Simulation results also suggest that the predicted average time delay of self-attribution and that of others-attribution changes over time. The multiple forward models were well correspondent to the characteristics of the experimental results. Namely, we found that there is a kind of learning effect by the inputs and attribution process. This implies that the attribution criteria might be navigated by the series of input. Users might feel as if they are controlling real world information systems without delay, even though the systems have time delay, because of the training of users' attribution criteria in advance.

Future work includes conducting experiments alternating the maximum time delay, in order to investigate the validity of the multiple forward models. We will add a unit to the models for representing the symptoms of schizophrenia. We would like to understand the mechanism of the attribution of action in both health subjects and psychiatric patients, which leads to universal design.

Acknowledgement

This work was supported by a Japan Science and Technology Agency Grant for the PRESTO program "Development of the Bilateral Multiscale Neural Simulator" (PI: Mihoko Otake) and a Grant-in-Aid for Scientific Research on priority area Systems Genomics (#014), Mobilligence (#454) and Information Explosion (#456) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

References

[1] S. J. Blakemore, D. A. Oakley, and C. D. Frith, "Dellusions of alien control in the normal brain," *Neuropsychologia*, vol. 41, pp. 1058–1067, 2003.

[2] N. Frank, C. Farrer, N. Georgieff, M. Marie-Cardine, J. Dalery, T. d' Amato, and M. Jeannerod, " Defective recognition of one's own actions in patients with schizophrenia," *American Journal of Psychiatry*, vol. 158, pp. 454–459, 2001.

[3] E. Dapratia, N. Francka, N. Georgieffa, J. Proustc, E. Pacheriec, J. Daleryd, and M. Jeanneroda, " Looking for the agent : an investigation into consciousness of action and self-consciousness in schizophrenic patients," *Cognition*, vol. 65, pp. 71–86, 1997.

[4] C. Farrer and C. D. Frith, "Experiencing oneself vs another person as being the cause of an action : the neural correlates of the experience of agency," *Neuroimage*, vol. 15, pp. 596–603, 2002.

[5] C. Farrer, N. Franck, N. Georgieff, C. D. Frith, J. Decety, and M. Jeannerod, "Modulating the experience of agency: a positron emission tomography study," *Neuroimage*, vol. 18, pp. 324–333, 2003.

[6] M. Jeannerod and N. Georgieff, " Consciousness of external reality: a "who" system for consciousness of action and self-consciousness," *Consciousness and Cognition*, pp. 465–477, 1998.

[7] D. M. Wolpert and M. Kawato, "Multiple paired forward and inverse models for motor control," *Neural Networks*, vol. 11, pp. 1317–1329, 1998.

[8] D. M. Wolpert, K. Doya, and M. Kawato, "A unifying computational framework for motor control and social interaction," *Philosophical Transactions of the Royal Society of London B*, vol. 358, pp. 593–602, 2003.