様式６－３

Topic-Setting Program to Advance Cutting-Edge Humanities and Social Sciences Research

（Area Cultivation）

Progress Report

（Summary of Final Report）

[Internally superimposed selves: Developing a new research in collaboration between philosophy and empirical science]

Core-Researcher: Masanori Takezawa
Institution: Hokkaido University
Academic Unit: Graduate School of Letters
Position: Associate Professor

Research Period : 2017 – 2020

1. Basic information of research project

Research Area	**Cognitive turn" and the transformation of identities**
Project Title	**Internally superimposed selves: Developing a new research in collaboration between philosophy and empirical science**
Institution	Hokkaido University
Core-Researcher (Name, Academic Unit & Position)	Masanori Takezawa, Graduate School of Letters, Associate Professor
Project Period	2017 - 2020
Appropriations Plan (¥)	2017 2,925,000 JPY
	2018 3,770,000 JPY
	2019 3,640,000 JPY
	2020 2,827,500 JPY

2. Purpose of research

We recognize the self as a unique entity that only we can know, and at the same time, we recognize that others have a "self" just like us. But how can we know that others have the same self as ours when we cannot directly experience their minds? This is one of the significant problems in contemporary philosophy, the problem of other minds (Husserl, 1931; Merleau-Ponty, 1945). Problems of self, consciousness, and intersubjectivity have long been the subject of philosophy, not science.

At the end of the 20th century, scientific research about consciousness, self, and intersubjectivity has rapidly progressed. Research on the neuronal correlates of self-awareness, theory of mind, mentalizing (Lombardo et al., 2010; Jankowiak-Siuda et al., 2011), and consciousness (Koch et al., 2016), has progressed. In addition, this trend has led to attempts to formulate subjective experiences such as qualia through the concept of information content (Tononi, 2004) and to understand the functioning of the computational system of the brain in an integrated manner using the free energy principle (Friston, 2010). Rather than simply identifying brain regions correlated with mental activity, researchers applied information theory to explore concepts that have been the subject of phenomenology. This movement on the scientific side has generated reactions from the philosophy side. Fields such as neurophilosophy, neuroethics, and experimental philosophy have emerged. However, this trend is a unidirectional influence from science to philosophy.

The purpose of this study is to develop new research that truly integrates philosophy and empirical science by deriving empirically testable hypotheses from the rich debates in philosophy about the self, consciousness, and the other's minds. We aimed to provide a role model of research that integrates philosophy and empirical science.

3. Outline of research (Including study member)

In response to one of the significant problems in contemporary philosophy, the problem of other's mind, Taguchi, a philosopher, has proposed the hypothesis that the self and others are internally superimposed and argued that what we might call zero personhood takes the form of the primordial state which is neither differentiated to the self or the other (Taguchi, 2018). In this study, we attempted to

develop new research, which empirical scientists alone cannot come up with, based on this philosophical argument. We conducted fMRI experiments and computer simulations with DNN for testing a series of hypotheses derived from Taguchi's considerations. For promoting research that truly integrates philosophy and empirical science, we held a public symposium in Tokyo in 2018, "Adventures on the Self: Beyond the Boundaries of Phenomenology, Robotics, Neuroscience, and Psychiatry," and disclosed the progress of our research to more than 200 participants. In addition, in July 2019, the Center for Human Nature, Artificial Intelligence and Neuroscience (CHAIN) was established as a joint research institution within Hokkaido University. This project aims to develop new research through the fusion of philosophy and empirical science, will continue to grow and expand towards a genesis of new science.

	Names	Affiliations	Roles
Principal Investigator	Masanori Takezawa	Associate Professor, Graduate School of Letter, Hokkaido University	Directing all the research
Co-investigator	Shigeru Taguchi	Professor, Graduate School of Letter, Hokkaido University	Developing research concepts and hypotheses
	Koji Ota	Associate Professor, Department of Humanities, Niigata University	Developing research concepts and hypotheses
	Kenji Ogawa	Associate Professor, Graduate School of Letter, Hokkaido University	Designing and conducting fMRI experiments
	Hiroyuki Iizuka	Associate Professor, Graduate School of Information Science and Technology, Hokkaido University	Designing and conducting DNN simulations
	Taiki Takahashi	Associate Professor, Graduate School of Letter, Hokkaido University	Developing research concepts and hypotheses

4. Research results and outcomes produced

In modern philosophy, the self is considered an object that only the isolated I can perceive. However, Taguchi points out that until the beginning of the 20th century, the self was not regarded as essentially

individual in philosophy. The questions similar to the problem of the other's mind were rarely found. He also points out that the primary state of the self is not such an isolated individual. Still, an overlapping state of self and others, which should be called zero personhoods, and that phenomenologist Husserl's theory can be interpreted in this direction (Taguchi, 2006, 2010).

The possibility of the existence of a state of superimposition of self and other (Taguchi 2019a), which is pre-reflective and not explicitly conscious, has been discussed in the context of imitation and self-consciousness (e.g., Piaget, Gopnik, Meltzoff), mirror neurons, and simulation theory (Gallese & Goldman, 1998; Gallese, 2003, 2005). The idea that pre-reflective processes precede and underlie self-consciousness and awareness of others is consistent with various other findings in neuroscience. In the project, we attempted to test the theoretical hypothesis formulated by the philosopher Taguchi using the methods of empirical science.

First, using fMRI and the multi-voxel pattern analysis (MVPA), a technique inferring mental processes (e.g., what the brain is looking at or thinking about) from brain activity data, we tested Taguchi's hypothesis. His hypothesis states a zero-person state where the self and others' perspectives are superimposed, underlies the basis for the representation and perception. We hypothesized that this state might occur in our daily visual experience of the world and tried to detect such states with the MVPA. Participants were presented with a video that simulated everyday observation of an object, and their brain activity was measured using fMRI while watching the video. FMRI measured the brain activity while the participants watched the video without thinking. Brain activities were also recorded while the participants watching the same videos from a specific perspective, either the self's or the other's perspective. According to Taguchi's hypothesis, when we are experiencing the world without being aware of a specific viewpoint, we are not separated into either the self or the other perspectives. We hypothesized that the brain activity in such a state is a mixture of the two brain states of the self-perspective and the other's perspective".

To test this hypothesis, we first trained a decoder to classify the brain activity when perceiving the world from the self-perspective or perceiving the world from the other's perspective. Then, we used decoders to classify brain activity while watching the world without thinking. Specifically, we calculated the probability that the data in the no specific viewpoint condition would be classified as either the self or other's perspective and calculated the entropy of the classification probabilities. High entropy means that the data can be classified into either viewpoint. The results of the experiment generally support the hypothesis. However, since the low accuracy of the decoders may cause the above results, further analyses have been further conducted.

Second, we tested the other hypothesis that the overlap of self and others is necessary to infer and understand the minds of others using the robotics simulation with deep neural networks (DNN). In advance of the current project, using hierarchical recursive neural networks (RNNs), a type of deep neural network, Iizuka had created a robot that acquires a cognitive map based on its own visual and motor experiences (Noguchi, Iizuka & Yamamoto, 2017). In this project, we added a different robot in the same field. Let's call them the self robot and the other robot. If the self robot has a theory of mind and can infer the mental state of the other robot, the self robot should be able to predict the scenes the other robot should be seeing accurately. We have shown that for such visual perspective-taking to be possible, a module called a superposition of self and others must be embedded in the robot's RNN (Noguchi, Iizuka, Taguchi & Yamamoto, 2019, in prep).

In both studies, we succeeded in developing research that could not have been possible by empirical scientists alone. We thus demonstrated the power of the rich discussions in philosophy for designing a new

study in the empirical sciences. Interaction between philosophy and empirical science is not uncommon. For example, in the field of experimental philosophy, philosophers themselves conduct psychological experiments to shed new light on philosophical debates. However, what we have aimed in this project is the opposite.

In February 2019, we held a public symposium at the University of Tokyo entitled "Adventures on the Self: Beyond the Boundaries of Phenomenology, Robotics, Neuroscience, and Psychiatry" to present the progress of our research to the public as a new role model for research that genuinely integrates philosophy and empirical science. More than 200 people attended the two-day symposium. It became clear that our vision of fruitful collaboration between philosophy and empirical science is getting significant attention from people working in broad areas.

In July 2019, the core member of the current project established a new research and education center at Hokkaido University called the Center for Human Nature, Artificial Intelligence and Neuroscience (CHAIN). In April 2020, CHAIN launched a graduate-level educational program that combines humanities and empirical sciences, and more than 20 Hokkaido University graduate students had participated in the program.

In November 2019, the current project and CHAIN co-hosted an international symposium, "Adventures in the Science of Consciousness: Crossover of Philosophy, Brain Science, AI, and Robotics Research," Hokkaido University. The symposium, which invited researchers actively working at the borders of philosophy and empirical science, attracted more than 200 participants. It is evident that our goal envisaged in the current project has been attracting a great deal of attention, and it is expected to develop further beyond the boundaries of existing disciplines and across generations.