과학철학 17-3(2014) pp. 53-68

Causal Mechanisms in Neurosciences: Toward the Integration of Optogenetics and Brain Imaging Studies

Yuki Sugawara*

Although a central aim of neuroscience is to describe the causal relationship between brain structure and function, it is unclear whether the methods currently used in neuroscience describe such a relationship. My aim was to articulate some of the methods by which neuroscience can describe this causal relationship and provide explanations. Thus, I used optogenetics and functional magnetic resonance imaging (fMRI) studies to analyze this relationship. I then argued that although optogenetics, but not fMRI studies, can describe the causal relationship, one can solve the problem by integrating optogenetics with fMRI studies on basis of the conception of mechanisms.

【Key Words】 Mechanisms, Explanation, Causation, Optogenetics, fMRI, Integration

* Department of Philosophy and History of Science, Kyoto University, sugawara.yuuki@gmail.com.

1. Introduction

Although a central aim of neuroscience is to describe the causal relationship between brain structure and function, it is unclear whether current methods in neuroscience describe this. My goal was to articulate some of the methods by which neuroscience can describe the causal relationship and provide explanations. I used optogenetics and functional magnetic resonance imaging (fMRI) studies to analyze whether methods in neuroscience describe the causal relationship. I then argued that although optogenetics, but not fMRI, can describe the causal relationship, one can overcome this limitation by integrating optogenetics with fMRI studies based on the conception of mechanisms.

To illustrate this point, I analyzed whether optogenetics and fMRI studies meet the requirement of the conception of mechanisms and characterize these methods (sections 2 and 3). After creating the foundation for the integration of optogenetics and fMRI, I argued that concrete and practical problems can be solved upon integrating these methods on the basis of the conception of mechanisms.

2. Causal mechanisms in optogenetics

2.1. Problems in neuroscience

With the development of methods to visualize previously unseen structures, the field of neuroscience has developed. Research on nervous tissue was advanced by Camillo Golgi. He developed *the black reaction* (known as *the Golgi method*), used to visualize nervous tissues under light microscopy. Scientists could not distinguish nervous tissues in fine detail until the Golgi method was developed. Using the Golgi method,

parts of neurons only combine with a black dye (a mixture of potassium dichromate and silver nitrate). Afterward, Santiago Ramón y Cajal learned the Golgi method described the dendrites for the first time and provided *the neuron doctrine*, which stipulates that nerve cells are contiguous and not continuous.

Although the Golgi method facilitated the growth of neuroscience, it has several limitations, which neuroscientists have attempted to overcome. In particular, the Golgi method has four limitations. First, cells obtained from dead animals were used in this method. To understand the function of neural circuits, the neurons of living animals should be visualized. Second, only a few neurons in a sample are stained using this method. Because the neurons are stained randomly, particular cells targeted for observation cannot be visualized. Third, the Golgi method visualizes only parts of neurons. To understand the brain circuit, the network of neurons should be visualized. Fourth, this method cannot reveal the causal relationship between brain structure and function.

Optogenetics has been developed as a method that overcomes the four limitations of the Golgi method. This method clarifies neural activities using particular wavelengths. Optogenetics was selected as *The Method of the Year* by *Nature* in 2010. This award encompassed all natural sciences, and thus, it has a major impact on science.

2.2. Solving problems in neuroscience: Optogenetics

The solution for limitation 1: Neurons of dead animals

The first limitation was solved by the discovery of green fluorescent protein (GFP). Using GFP, scientists gained the ability to examine the neurons of living animals. GFP was isolated from the Pacific Northwest jellyfish *Aequorea victoria*, which emits green light similarly as other jellyfish. In the 1960s, Osamu Shimomura examined the phenomenon of

bioluminescence and discovered two proteins. One of the proteins emits blue light when in contact with calcium salt in seawater, and the other protein emits green light when in contact with the protein that emits blue light. Shimomura named the blue protein *Aequorin* and the green protein *Green Protein* (later, *Green Fluorescent Protein*). In 1992, Douglas Prasher discovered the DNA sequence that encodes GFP and recognized the potential of GFP as a tracer molecule (Prasher et al. 1992). In 1994, Martin Chalfie introduced the *GFP* gene into cells such as colon bacilli and *Caenorhabditis elegans* (Chalfie et al. 1994). These cells emitted green light, and basic structure themselves (e.g., axon, soma, and dendrite) could be illustrated when illuminated by blue light.

The solution for limitation 2: Dying randomly

The second limitation was also overcome with GFP. For example, scientists can introduce the *GFP* gene into particular cells by methods of molecular biology. In this manner, particular cells can be observed instead of groups of cells.

The solution for limitation 3: Visualizing only parts of neurons

The third limitation was resolved by Roger Tsien (Shaner et al. 2004), who altered the *GFP* gene such that it could emit various colors (e.g., yellow, blue, red, etc.). Moreover, Jean Livet developed a method that enables various neurons to emit different colors (Livet et al. 2007), thus making it possible to distinguish and trace these neurons.

The solution for limitation 4: Non-causality

The fourth limitation was solved by the discovery of optically active molecules such as ChR2 and HaloR. Until neuroscientists obtain genetically targeted optical control of neural activity using these molecules (e.g., Boyden et al. 2005), no other method to control specific

neural activities on a millisecond timescale exists. The primary methods for stimulating neural activity were electrical stimulation and chemical injection. In electrical stimulation, relatively specific neural activity can be stimulated by inserting a microelectrode into the targeted brain area and then electrically stimulating the tissue. However, the stimulated neural activity is not specific, because electrical stimulation induces peripheral neural activity. Moreover, the electrical stimulation can induce neural activity but cannot restrain it. On the contrary, regarding chemical injection, agonists (e.g., glutamic acid) and antagonists (e.g., muscimol) have been used by many neuroscientists. Although neural activity can be induced and restrained by chemical injection, specific neural activity cannot be induced. In contrast to these methods, optogenetics uses high-membrane and low-invasive light. Because the optically active molecule can be expressed in specific neurons, they can induce specific neural activity. With time, scientists were able to induce targeted neural activity on a millisecond timescale. Moreover, because ChR2, which induces neural activity, is activated by blue light and HaloR, which restrains neural activity, is activated by blue and orange lights, neuroscientists will be able to switch between activation and restrainment rapidly by altering the wavelength of light, thus permitting the simultaneous activation and restrainment of specific neural activities.

Primary culture neurons are used for the first optical operation. Neural activities are induced by transfecting the *ChR2* gene and providing blue light illumination (Boyden et al. 2005). This is because ChR2 is a non-selective positive ion channel that, after opening, facilitates membrane potential depolarization via sodium ion flow to the inside of cells. In the same manner, in culture neurons that express HaloR, the membrane potential is hyperpolarized by orange light illumination, and the occurrence of action potentials is restrained. HaloR is a chloride ion pump that is activated by orange light. HaloR pumps

chloride ions out of the cell. After scientists confirmed the functions of ChR2 and HaloR *in vitro*, they were applied in the assessment of neural activity and behavior *in vivo*. First, Nagel and colleagues (2005) illustrated that they could control the neural activity and behavior of *C. elegans* using light. Subsequently, the method was applied to various animal models (e.g., flies, primates, etc.). For example, previous research illustrated that the regulation of sleep/wakefulness of mice can be altered via interventions targeting the orexin system (Adamantidis et al. 2007, Tsunematsu et al. 2011) and that memory is stored in the neurons of the hippocampus (Liu et al. 2012). I argue that optogenetics can clarify the causal relationship between brain structure and function by controlling neural activity at a millisecond timescale. Adamantidis et al. also remarked, "this study establishes a causal relationship between frequency-dependent activity of a genetically defined neural cell type and a specific mammalian behavior central to clinical conditions and neurobehavioural physiology"[1].

2.3. Generalizing optogenetics on the basis of the conception of mechanisms

Optogenetics overcomes the four limitations of the Golgi method and is useful for understanding neural activity. However, to what extent can optogenetics can serve as a generalized method for approaching the objects remains unknown. I analyzed the extent to which optogenetics meets the requirements of the conception of mechanisms as defined by Craver and Darden (2013) and arranged the characterizations that optogenetics has as a science.

There are five components and features of mechanisms[2]:

---

1) Adamantidis et al. (2007), p. 420.

2) Craver and Darden (2013), pp. 16-20.

(1) Entities and activities: Mechanisms are composed of both *entities* and *activities*. The entities are the parts with their various properties. Activities are what entities perform[3].
(2) Setup, start, and finish conditions: Ideally, mechanisms can be described as working from the *start* to the *finish*[4].
(3) Productive continuity: Mechanisms are productive, i.e., they are the processes by which an end state, a product, a process, or a change occurs[5].
(4) Regularity: Most interesting biological mechanisms are regular, i.e., they usually work in the same or similar manner under the same or similar conditions[6].
(5) Organization: The entities and activities in a mechanism are organized spatially, temporally, and actively such that they produce the phenomenon. Spatial organization includes locations, sizes, shapes, and orientations of component entities. Temporal organization includes the orders, rates, and durations of the stages. Active organization includes facts about one component affecting the other and the mechanisms by which such alterations occur[7].

It remains controversial whether the requirements of the conception of mechanisms can capture the features of these mechanisms (e.g., Skipper and Millstein 2005). Some philosophers proposed a conception of mechanisms that differ from that proposed by Craver and Darden's definition (e.g., Glennan 1996, 2002; Bechel and Abrahamsen 2005). One may be able to modify these definitions and create a unique definition. For example, Torres (2009) attempted to modify Machamer, Darden, and Craver (2000) and Glennan (2002) definitions and create a single definition of mechanisms. Although many philosophers of science have analyzed the definition of mechanisms, I believe that the

3) Ibid., p. 16.
4) Ibid., p. 18.
5) Ibid., p. 19.
6) Ibid.
7) Ibid., p. 20.

conception of mechanisms defined by Craver and Darden is useful for analyzing the validity of the methods in neuroscience.

Does the explanation of optogenetics meet the requirements of the conception of mechanisms? In this section, I discuss whether the experiment of optical control for the DG of the hippocampus (Liu et al. 2012) meets the requirements. First, in the experiment, the *entity* is the cell in the DG of the hippocampus, and the *activity* is the activity of the cell. Second, the *start* condition is the optical irradiation with the fiber, and the *finish* condition is freezing behavior of mice. Third, there is *productive continuity* while the process between the optical irradiation and appearance of freezing behavior occurs. Fourth, there is *regularity* in term of statics. Fifth, the *spatial organization* is completed by the specification of the location and the size, the *temporal organization* is completed by the specification of the orders and rates, and these organizations are performed *actively*.

The explanation in optogenetics meets the requirements for the conception of mechanisms. It appears that the explanation of optogenetics is valid.

3. Causal mechanisms in brain imaging studies

3.1. Problems in brain imaging studies

Brain imaging has revealed important information about cognitive function in humans. fMRI is used to measure neural activity by detecting the local changes in blood volume and oxygen volume. fMRI is based on MRI, which exploits the fact that magnetism varies with the volume of oxygen in the blood. This magnetism has an effect on the relaxation rate of protons locally. This method has advantages, i.e., it is

non-invasive and safe for humans. It is an important method for measuring human brain activity, which has been researched extensively.

However, the interpretation of fMRI data is not simple. fMRI does not measure neural activity directly. The relationship between neural activity and blood volume is complicated. Measurements of blood volume do not clarify neural activity. Moreover, the resolution of fMRI is relatively low. As a result, each picture element of a dataset is not related to the activity of a single neuron, but the activities of millions of neurons. As such, each function in the picture elements may be different.

The temporal resolution of fMRI is also a problem. Changes in blood volume occur in intervals of seconds, whereas changes in neural activity occur in intervals of milliseconds. Neural activity observed on fMRI represents a part of the total neural activity.

3.2. Generalizing brain imaging studies on the basis of the conception of mechanisms

To what extent does the scientific explanation given by fMRI studies meet the requirements of the conception of mechanisms? First, in fMRI experiments, the *entity* is the blood volume, and the *activity* is the change in the blood volume. Second, because of the low-temporal resonance of fMRI, scientists cannot identify the *start* and *finish* conditions of mechanisms producing the phenomena. They can only describe the difference between when the subject performs and does not perform the experimental task. Third, because there are no start and finish conditions for mechanisms producing the phenomena in fMRI experiments, there is no *productive continuity*. This indicates that it is doubtful whether the processes in fMRI have *causal relation*. Fourth, the *regularity* is clarified statistically. Fifth, because of the low-temporal resolution, the mechanisms described by fMRI are not *organized*

temporally. Moreover, fMRI only measures the changes in the blood volume, and thus, the mechanisms described by fMRI studies are not *organized spatially*.

Can fMRI studies explain the causal relationship between brain structure and function? First, the results of fMRI studies provide statistics but not the blood oxygenation level dependent (BOLD) signal itself. Moreover, it remains unclear what the BOLD signal represents. In addition, the relationship between brain activity and the blood volume is unknown. It is supposed that the BOLD signal is correlative with the change in neural activity because the BOLD signal is reduced by deoxy-Hb; however, it is reported that the BOLD signal is not correlated with neural activity (Mishra et al. 2011). Previous research indicated that the neural activity of glial cells induces the BOLD signal (e.g., Schulz et al. 2012). *Nature Methods* also reported, "it's striking how little we know about the fMRI signal itself"[8] and "the biggest conundrum in fMRI is that, exactly, the technique is measuring"[9].

4. Discussion: Integration of optogenetics with brain imaging studies

Although I confirmed that a correlation between fMRI data and neural activity can only be assumed, some researchers attempted to resolve the limitations of fMRI by combining fMRI with optogenetics (Lee et al. 2010). Baltes et al. succeeded in devising an optogenetic fMRI device for use in mice (Baltes et al. 2011). The developed system enabled whole-brain searches (including the ability to describe the effect from

8) Nature Methods (2012), p. 517.
9) Ibid.

one brain area to another) for mice and enabled the start condition to be set as neural activity and the detection the whole-brain responses. As a result, it is likely that this approach can clarify causal relationships regarding interactions between brain areas. However, the technique requires further developments of new analytical methods for new measurement data. Because research on mice is conducted in detail, it is possible to make good use of the measurement results of fMRI in humans and clarify whether the BOLD signal corresponds to brain activity.

In addition, I argued that optogentics meets all the requirements of the conception of mechanisms, but fMRI studies do not. It appears that optogenetics, but not fMRI, can describe the causal relationship between brain structure and function. However, this does indicate that fMRI is not a useful method. Although fMRI studies have some problems, if fMRI is integrated with optogenetics methodologically, then the limitations of this method can be overcome. For example, the combined method can clarify the causal relationships regarding interactions between brain areas from a macro viewpoint.

If fMRI must be integrated with optogenetics to be recognized as a valid method, then the requirements of the conception of mechanisms can confirm the usefulness of the combined technique as a scientific method. If so, then the requirements of the concept of mechanisms are useful for judging the scientific validity of this method.

5. Conclusion

My goal was to articulate some of the methods by which neuroscience can describe causal relationships and provide explanations. I used optogenetics and studies on fMRI to analyze whether methods in

neuroscience describe causal relationships. I then argued that although optogenetics, but not fMRI, can describe causal relationships, one can overcome the limitations of fMRI by integrating the technique with optogenetics on the basis of the conception of mechanisms. In addition, I argued that the requirements of the concept of mechanisms could be useful for judging the scientific validity of this method.

References

Adamantidis, A. R., Zhang, F., Aravanis, A. M., Deisseroth, K., and de Lecea, L. (2007), "Neural substrates of awakening probed with optogenetic control of hypocretin neurons", *Nature* 450: pp. 420-5.

Baltes, C. (2011), "Increased blood oxygen level-dependent (BOLD) sensitivity in the mouse somatosensory cortex during electrical forepaw stimulation using a cryogenic radiofrequency probe", *NMR Biomedicine* 24(4): pp. 439-46.

Bechtel, W. and Abrahamsen, A. (2005), "Explanation: a mechanist alternative", *Studies in History and Philosophy of Biological and Biomedical Sciences* 36: pp. 421-41.

Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., and Deisseroth, K. (2005), "Millisecond-timescale, genetically targeted optical control of neural activity", *Nature Neuroscience* 8(9): pp. 1263-8.

Chalfie, M. Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994), "Green fluorescent protein as a marker for gene expression" *Science* 263(5148): pp. 802-5.

Craver, C. and Darden, L. (2013), *In search for mechanisms: Discoveries across the life sciences,* The University of Chicago Press.

Glennan, S. (1996), "Mechanisms and the nature of causation", *Erkenntnis* 44: pp. 49-71.

_____ (2002). "Rethinking mechanistic explanation", Philosophy of Science 69: pp. 342-53.

Lee, J. H. (2010). "Global and local fMRI signals driven by neurons defined optogenetically by type and wiring", *Nature* 465(7299): pp. 788-92.

Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A.,

Deisseroth, K., and Tonegawa, S. (2012), "Optogenetic stimulation of a hippocampal engram activates fear memory recall", *Nature* 484: pp. 381-7.

Livet, J., Weissman, T. A., Kang, H., Draft, R. W., Lu, J., Bennis, R. A., Sanes, J. R., and Lichtman, J. W. (2007), "Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system", *Nature* 450: pp. 56-62.

Machamer, P., Darden, L., and Craver, C. (2000), "Thinking about Mechanisms", *Philosophy of Science* 67: pp.1-25.

Mishra, A. M., Ellens, D. J., Schridde, U., Motelow, J. E., Purcaro, M. J., DeSalvo, M. N., Enev, M., Sanganahalli, B. G., Hyder, F., and Blumenfeld, H. (2011), "Where fMRI and electrophysiology agree to disagree: Corticothalamic and striatal activity patterns in the WAG/Rij rat", *The Journal of Neuroscience* 31(42): pp. 15053-64.

Nagel, G., Brauner, M., Liewald, J. F., Adeishvili, N., Bamberg, E., and Gottschaik, A. (2005), "Light activation of channelrhodpsin-2 in excitable cells of caenorhabditis elegans triggers rapid behavioral responses", *Current Biology* 15: pp. 2279-84.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992), "Primary structure of the *Aequorea victoria* green-fluorescent protein" *Gene* 111(2): pp. 229-33.

Schulz, K., Sydekum, E., Krueppel, R., Engelbrecht, C. J., Schlegel, F., Schroter, A., Rudin, M., and Helmchen, F. (2012), "Simultaneous BOLD fMRI and fiber-optic calcium recording in rat neocortex", *Nature Methods* 9: pp. 597-602.

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., and Tsien, R. Y. (2004), "Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. Red fluorescent protein", *Nature Biotechnology* 22(12): pp. 1567-72.

Skipper, R. A. and Millstein, R. L. (2005), “Thinking about evolutionary mechanisms: Natural selection”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 36(2): pp. 327-47.

Nature Methods Editorial (2012), “fMRI: a tree with fuzzy roots”, *Nature Methods* 9(6): p. 517.

Torres, P. J. (2009), “A modified conception mechanisms”, *Erkenntnis* 71: pp. 233-51.

Tsunematsu, T., Kilduff, T. S., Boyden, E. S., Takahashi, S., Tominaga, M., Yamanaka, A. (2011), “Acute optogenetic silencing of orexin/hypocretin neurons induces slow-wave sleep in mice”, *The Journal of Neuroscience* 31(29): pp. 10529-39.

Date of the first draft received	2014. 11. 03
Date of review completed	2014. 11. 17
Date of approval decided	2014. 11. 18

신경과학에서의 인과적 기제들: 광유전학과 뇌영상연구의 융합을 향하여

유키 스가와라

신경과학의 주요 목적이 두뇌 구조와 기능 간의 인과적 관계를 기술하는 것이지만, 그 분야에서 현재 사용되고 있는 방법들이 그런 관계를 제대로 기술하고 있는지는 분명치 않다. 이 글의 목적은 신경과학이 인과적 관계를 기술하고 설명을 제공할 수 있는 방법 중 몇 가지를 제시하는 데 있다. 그 관계를 분석하기 위해 나는 광유전학과 기능성자기공명영상(fMRI) 연구를 사용할 것을 추천한 다음, 기능성자기공명영상 연구는 인과적 관계를 기술할 수 없는 반면에 광유전학이 기술할 수 있다는 문제를 기제라는 개념에 근거를 두어 광유전학을 기능성자기공명영상 연구를 결합시킴으로써 해결할 수 있다고 주장할 것이다.

주요어: 기제, 설명, 광유전학, 기능성자기공명, 융합