Human Brain Project

Co-funded by the European Commission

Project Number:	284941	Project Title:	Human Brain Project

Document Title:	First Report on How Far Brain Simulation can Explain Mechanisms of the Mind
Document Filename[(1)]:	HBP_SP12_MS219_Final.docx
Milestone Number:	MS219
Work Package(s):	WP12.2
Dissemination Level:	PU = Public
Planned Delivery Date:	Month 18 / 31 March 2015
Actual Delivery Date:	Month 18 / 30 March 2015
Authors:	Yadin DUDAI, WIS (P77), Kathinka EVERS UU (P78)
Compiling Editors:	Yadin DUDAI, WIS (P77), Kathinka EVERS UU (P78)
Contributors:	
STO Review:	UHEI (P45): Björn KINDLER, Sabine SCHNEIDER, Martina SCHMALHOLZ
Editorial Review:	EPFL (P1): Richard WALKER, Guy WILLIS, Celia LUTERBACHER
Abstract:	Simulation is a powerful method in science and engineering. In neuroscience, problem-oriented computer simulations of specific systems and functions of the brain are extensively used to test predictions, validate conclusions and models, and to guide hypothesis-driven experiments and new models at various levels of analysis. Rapid advancements in neuroscience and in computing draw increasing attention to large-scale brain simulations. Against this background, we raise the question: 'how far can brain simulation contribute to the explanation the brain and the mind?' We delineate three types of issues that relate to the potential explanatory power of large-scale brain simulations. We note that, whereas some types of issues are expected to be resolved with the advance of neuroscience and computing technology, others pose more profound and long-lasting conceptual obstacles that should be taken into account in managing the expectations from the approach.
S	Simulation, brain, computers, explanation, scale, hypothesis, epistemology

.

Table of Contents

1. Executive Summary

Simulation is extensively used in brain research to test predictions, validate conclusions and models, and guide hypothesis-driven experiments and new models at various levels of analysis. Rapid advancements in neuroscience and in computing draw increasing attention to large-scale brain simulations. But how far can brain simulation contribute to the explanation of the brain and the mind?

We first delineate three types of goals of brain simulation:

- Serve as one of the tools in the rapidly expanding, rich methodological and technological armoury of modern neuroscience;
- Serve as a device in the planning, development and testing of brain-inspired technologies, such as bionic devices and hominid robots; and
- Contribute to the understanding of the brain and the mind.

Our discussion refers only to the latter goal.

We then proceed to delineate three types of issues that relate to the potential explanatory power of large-scale brain simulations:

- Methodological and Technical issues,
- Conceptual and Philosophical issues, and
- Those that refer to the role of consciousness, about which we still know very little and therefore consider it as an Issue of Type Unknown.

Methodological and technical obstacles are likely to be reduced or even resolved as the field advances, but conceptual issues may stay with us irrespective of advancement in science. A major example relates to epistemic opacity, i.e. the cognitive inaccessibility of intermediate steps in a highly complex process or mechanism, irrespective of whether we can simulate the process or mechanism on a computer, or predict their phenomenological outcome. Profound and long-lasting conceptual obstacles, such as the barrier of epistemic opacity, should be taken into account in managing the expectations from improved understanding of the brain and the mind by using brain simulation.

2. Aim of this Document

The aim of this report is to present the advances of WP12.2 in its analyses of the meaning of "simulation", and the authors' stand on how far the scientific method of brain simulation can help explain mechanisms of the mind.

3. Simulations, small and large

Simulation is a powerful method in science and engineering (Dudai & Evers 2014). In neuroscience, problem-oriented computer simulations of specific systems and functions of the brain are extensively used to test predictions, validate conclusions and models, and to guide hypothesis-driven experiments and new models at various levels of analysis (Dayan and Abbott 2005, Schneidman *et al*. 1998, Loewenstein and Sompolinsky 2003, Norman 2010). Rapid advancements in neuroscience and in computing, however, draw increasing attention to ambitious large-scale brain simulations, some of which are based on a massive bottom-up approach, of which the major one is the Human Brain Project (HBP), in which the simulation effort is a sequel of the Blue Brain Project (Markram 2006; see also Kandel *et al*. 2013).

Generally speaking, simulation in the context of our discussion is a set of computational operations performed on virtual agents in a virtual environment, embodied either in general-purpose or special-purpose computers, with the goal of imitating the native computations, algorithms, actions and emergent behaviour of a natural or artificial system. The agent is a distinct entity in the simulation that can perform or undergo a set of interactions (Kubera *et al*. 2008). An environment is the collection of interactions, computational states, contexts and boundary conditions in which the agent exists. A simulation is called 'large scale' if the environment contains a large number of agents with different behaviours and/or a large number of possible actions per population of agent. The value of 'large number' is not strictly defined. In brain research, simulations are commonly considered as large scale if they exceed the testing of hypothesis-driven predictions pertaining to a limited part or function of the nervous system; hence bottom-up simulation of the behaviour of 100 neurons in a cortical minicolumn would be considered large-scale, let alone of the behaviour of interacting minicolumns in a patch of neocortex, let alone of the behaviour of an entire mammalian brain, which is presented as the futuristic goal of HBP. Unless otherwise indicated, in our present discussion we focus on large-scale, mostly bottom-up brain simulation, such as originally envisaged in the HBP. We should also pinpoint at the outset of this discussion the 'elephant in the room', i.e., our assumption that explanation of the brain is bound to yield explanation of the mind; this we consider for the sake of discussion as a given in contemporary mainstream neuro- and cognitive science.

How far should we expect simulations in general, and large-scale brain simulations in particular, to contribute to the understanding of the brain and the mind?

4. Three potential goals

Three types of goals of brain simulation should be considered concerning the potential explanatory power of the approach:

1) Simulation as a methodological heuristic as part of the rapidly expanding, rich armoury of modern neuroscience;

2) Simulation as a device in the planning, developing and testing of brain-inspired technologies, such as bionic devices and hominid robots;

3) Simulation as an aim *per se*, attempting to understand the brain and furthermore to serve as an *in silico* platform to complement *in vivo* neuroscience.

Goal 1) is partially agnostic to the explanatory power of simulation, whereas goal 2) is completely so. It is of note that 1) above is only partially agnostic to the issues of explanatory power, because if simulation serves to verify expectations stemming from hypotheses and models, its role in promoting explanation is evident, yet, it is not *per se* the conceptual and experimental tool that carves explanation, but rather an auxiliary tool. It is goal 3), namely the ambition to explain that is at the crux of the question: 'can simulations spearhead explanations?' Are they obligatory, let alone sufficient, to achieve explanation in the brain sciences?

5. Types of issues concerning explanation

Those who hold an affirmative stance on this follow, sometimes explicitly but mostly implicitly, a maxim coined by the Italian philosopher Giambattista Vico (Vico 1710, 1988), that posits that one can only understand what one is able to build, i.e. that truth is realized through creation, or recreation. In embarking in the discussion of the validity of this belief in the context of attempts to explain the brain by simulating it, several potential issues and caveats should be noted.

It is methodologically convenient to group these issues at the outset into three types:

- Methodological/Technical Issues,
- Conceptual/Philosophical Issues, and
- 'Issues of Type Unknown'.

The Methodological/Technical issues could be resolved relatively quickly, as developments in brain and cognitive sciences are likely to gradually convert them into minor issues or non-issues. In contrast, the Conceptual/Philosophical issues go considerably beyond the methodological ones and may therefore be expected to remain open for much longer. As for 'Issues of Type Unknown', we cannot define them well in the lack of sufficient knowledge about their nature, and their contribution to the ability of simulation to explain the brain and mind is for the future to tell.

6. Methodological/Technical Issues

These issues hinder the accomplishment of brain simulation at the time of writing, even if one believes that simulation may be a scientifically valid approach to explaining the brain and mind.

6.1 Scarcity of knowledge

Many neuroscientists consider contemporary attempts to undertake large-scale brain simulations as premature (Courtland 2014). The argument is that, at this stage, we know too little about the brain, while productive simulations require established theory and data (e.g. Kamerlin *et al*. 2011). Not only do we lack information about identified types of mostly-unknowns, such as neuronal codes, computational goals or wiring diagrams, we may also have not yet even identified other types of unknowns required to model large subsystems of the brain, let alone the whole brain. Some take the stand that, in the absence of well-established top-down theory and recurrent high-level reality-checks (i.e., 'on-the-fly' consideration of global physiological and behavioural output – see below), bottom-up large-scale reconstruction may lead to much wasted effort. Furthermore, the federation of data from different labs has to take into account the fact that that even small variations in methodology and conditions can yield different interpretations of neuronal state and activity, and that different labs seldom, if ever use, exactly the same conditions and protocols. The invariants identified under these conditions may mask important features (see below).

As far as data required for human brain simulation are concerned, it is sufficient to note that cellular physiology data are scarce and obtainable only from patients, and that non-invasive functional neuroimaging (mainly functional magnetic resonance imaging, fMRI) has limited spatiotemporal resolution, which currently constrains its applicability to high-resolution brain simulation, although is useful in obtaining important information on the role of specific brain areas and their functional connectivity in perceptual and cognitive processes. One possibility to bridge the gap from the cellular to the cognitive is to use data from the non-human primate brain, but there is not yet enough of this sort of data for the purpose of large-scale brain simulation. However, this is a methodological and technical problem, which although clearly not trivial, can be expected to be resolved over time.

6.2 Computing power

The computing power required for large-scale simulation of a whole human brain, using current techniques, is not yet available. Exascale-level machines are required, that, if pursued by current technology, will demand daunting amounts of energy (Kogge 2011). However, given the fast pace of progress in computer technology (e.g. Merolla *et al*. 2014), this issue will probably be resolved prior to the resolution of the scarcity of knowledge problem mentioned above.

6.3 Reality checks

Large-scale simulations are expected to involve iterations in which the performance of the simulated systems is evaluated by benchmarks. However, scarcity of knowledge (*see 6.1* above) may raise doubts concerning the suitability of such benchmarks. In most cases, we do not yet know whether our presumed correlation between the activity of a particular circuit with specific physiological or behavioural performance indeed reflects the native function of the circuit. For example, are place cells primarily sensitive to spatial coordinates, or amygdala circuits to fearful stimuli? Uncertainty about such relationships may result in optimizing simulations to deliver performances which do not reflect those in nature. Yet again, this problem is expected to resolve itself gradually as our understanding increases.

7. Conceptual/Philosophical Issues

Conceptual issues arise from and are substantially informed by technological and methodological advances, but raise questions that go beyond empirical data.

7.1 Epistemic opacity

Is the aforementioned Vico maxim, that explanation is achieved through creation, applicable to computer simulation of complex systems? Having fed in the information and let the machine run the computations involving strings of equations and come up with emergent properties, can we really claim to understand the system better when part of the process is epistemically opaque? We define epistemic opacity in the context of this discussion as the cognitive inaccessibility to intermediate steps in a process or mechanism, even when the outcome of the process or mechanism can be predicted. And what is it that creates the opaqueness, given that in computer brain simulation we in fact wrote the equations? Is it the numerical iterations, high dimensionality, nonlinearity, emergence, all combined? This brings us to the meaning of 'understanding'. Some will note that even in daily life, we claim to understand natural phenomena without really mentally grasping their inner working. For example, we predict that if we release a ball from a tower, the ball will fall because of gravity. But is the attraction of physical bodies transparent to us, epistemically? Or is our sense of understanding due to habituation with the phenomenon or our awareness of the existence of the physical law? In the meantime, the acceptable magnitude of epistemic opacity in a computer simulation designed to predict the behaviour of a particular system, is for the individual scientist to decide, and will probably vary with the professional training and the level of description and analysis.

7.2 Collapsing resolutions

Attempts at large-scale brain simulation currently rely on limited sampling and statistical typification (de Garis 2010; Eliasmith and Trojillo 2014). It is one thing to sample phenomena in experiments in search of mechanisms and to classify the data to facilitate understanding, another to rely on limited sampling to faithfully build the system anew. The possibility cannot be excluded that important *in vivo* properties of real-life neurons are concealed or minimized in such a process. Despite the robustness of biological systems, relying on extracted invariants may result, not only in missing data, but also in going beyond the data, because of potentially erroneous generalizations. Such methods may also reduce the ability to rely on the simulation to perform new fine-grained experiments *in silico* ('higher order simulation'), which is advocated as a possible contribution of brain simulation (i.e. to replace *in vivo* or *in vitro* experiments that are complex, time-consuming and/or cause animal suffering). Furthermore, it may result in a situation in which the outcome of an *in silico* experiment will have to be verified *in vivo* after all.

7.3 Representational parsimony

Much of our scientific progress, understanding and intellectual joy stems from the ability to extract and generalize laws of nature. Describing the universe in a minimal number of equations is often equated not only with ultimate understanding but also with beauty (Weinberg 1992). If we aim to reproduce details in simulations, do we still advance in 'understanding' in that respect, or just imitate nature? Besides raising again the issue of epistemic opacity (see 7.2 above), a more practical question comes up: 'should we expect a small set of laws to describe a complex adaptive system like the brain?' Some will say that this depends on the level of description. The brain can be considered as a community of organs with different functions and phylogenetic history, which renders the hope of understanding in detail the operation of each via task-relevant computations doubtful. However, it still leaves open the possibility that some basic principles of brain operation are explainable by a unified theory. But this depends on the level of description. One may claim that we already understand some fundamental principles of brain operation, for example, that spikes encode and transmitters convey information, but this level of description is obviously not what brain scientists have in mind in trying to 'understand' the brain. It is of note that high parsimony in realistic models has the potential to ameliorate epistemic opacity.

8. Consciousness: an unknown issue

Despite attempts to classify consciousness in terms of the brain and cognitive sciences and to identify biological models of consciousness, our scientific knowledge on this topic is at present very limited, to the extent that we consider it as an issue of type unknown. Yet it is expected to impact strongly the potential of brain simulation to explain brain and mind. The possibility should not be excluded, though, that the status of this issue will mutate over time, as we gain experience in brain processes and mechanisms, in bridging the gap between mental phenomenology and its brain underpinning, and in the mental ability of bionic systems and hominid robots.

8.1 How to test if consciousness emerges in a simulated brain?

Can consciousness be simulated at all on a computer? For simplicity, we will not discuss here the multiplicity of conscious states in humans, and refer to one type only, which is intuitively appreciated: conscious awareness involving self awareness that one is conscious (autonoetic consciousness, Tulving 1985). Suppose the answer is no, would we then be still able to explain the brain by simulation (epistemic opacity notwithstanding)?

At present, available evidence justifies only a rather tame hypothetical stance: If consciousness is necessarily an outcome of a certain type of organization or function of biological matter, then brain simulation will never gain consciousness; whereas if consciousness is a matter of organisation alone, e.g. extensive functional interconnectivity in a complex system, then it might arise in simulations *in silico*.

How would we recognise whether a future brain simulation is conscious or not? Two main types of approaches can be envisaged. The first, a Turing-type test for a conscious entity. Yet by itself this is insufficient, because we can easily imagine a computer being able to mimic the expected responses of a conscious entity without experiencing consciousness. The second, provided we assume faithful imitation of the relevant native brain activity, would be to identify activity signatures that reflect conscious awareness in the human brain. This is in principle similar to the way one attempts to identify sleep and dreams objectively, by looking for characteristic brain activity signatures (Nir and Tononi 2009). But on the one hand, we do not yet know such signatures; on the other, even if they are identified, they may not exhaust signatures of conscious awareness in a simulated system. A pragmatic heuristic approach could be combination of two elements, still short of a sufficient condition. One, a Turing-type test; the second, activity signature in the simulated entity that fits the one expected in the original biological brain, and is time locked to the responses taken to reflect conscious behaviour.

8.2 Is realistic human brain simulation possible in the absence of consciousness?

It is possible to consider brain simulation without the question of consciousness arising However, when processes in the brain are simulated that are conscious in the human being (for example, declarative emotion), the question arises: if consciousness is not simulated, how adequate can that simulation be? The epistemological question here needs to be borne in mind throughout the project's development.

9. Can brain simulation help to improve medical treatment?

One of the proposed goals of human brain simulation is to increase our understanding of neurological and mental illnesses, and ultimately develop improved therapies on the basis of this understanding. There is little to argue when the disease in question is characterized mainly by its neurological manifestation, for example, seizures, palsy, tremor and the rigidity manifestations of Parkinson's disease. But the more we shift into diseases that are considered essentially mental, such as chronic anxiety, depression or schizophrenia, the question comes up, how adequate, or informative, can a simulation of these mental disease be if there is no conscious experience in the simulation? This issue is the core of a separate chapter in the research of our WP, Gold & Dudai (*submitted.*) The role of consciousness and the effects of this role on the outcome of simulation of human brain faculties will be important to assess in this context.

We argue that at the current state of knowledge, psychiatrists identify most severe mental disorders as conjoint occurrence of multiple symptoms, in which there is a major role to subjective experiences expressed as verbal self-report. It is still to be proven that these could be translated one-to-one into distinct and uniquely defined sets of biological markers. Unless this happens, the lack of conscious states and the ability to translate them into verbal report is bound to severely limit the explanatory power of simulation of the human complex disease state into objectively observed behaviour of a simulated system.

The potential contribution of large-scale simulation to better understanding, let alone improved treatment, of mental disease, could hence be considered as positioned somewhere in between issues that are expected to be ephemeral and those of type unknown. This is definitely an issue that epitomizes the daunting challenges facing attempts to simulate the brain while expecting in this way to simulate aspects of the mind as well.

10. Conclusion

All in all, the attempt to proceed on large-scale simulation of the human brain is imaginative and visionary. Yet careful contemplation not only of its feasibility at the current state of knowledge, but also of its expected contribution to our knowledge of the brain and the mind, are warranted.

Annex A: Glossary

Term	Explanation
Brain simulation	The attempt to imitate or replicate the functional brain, either in part or *in toto*, outside of the brain. In contemporary neuroscience, this implies modelling on a computer. The ultimate primary aim is to imitate and understand the native computations, algorithms, states, actions and emergent behaviour of the brain, as well as to promote brain-inspired technology. The process involves the application of mathematical models that are preferably constrained by biological information to experimental or mock data.
Epistemic opacity	The cognitive inaccessibility of intermediate steps in a process or mechanism, irrespective of whether the outcome of the process or mechanism can be predicted.
Simulation	The process and product of making something appear or perform like something else.

Annex B: References

Courtland R (2014). http://spectrum.ieee.org/tech-talk/computing/hardware/can-the-human-brain-project-succeed

Dayan P, Abbott LF (2005). Theoretical Neuroscience: Computational and mathematical modeling of neural systems. Cambridge, MA: MIT Press.

de Garis H, Shuo C, Goertzel B, Ruiting L (2010). A world survey of artificial brain projects, Part: Large-scale brain simulations. Neurocomputing 74, 3-29.

Dudai Y, Evers K (2014). To simulate or not to simulate: What are the questions? Neuron 84, 254-261.

Kamerlin S, Vicatos S, Dryga A, Warshel A (2011). Coarse-grained (multiscale) simulations in studies of biophysical and chemical systems. Annu Rev Phys Chem 61, 41-64.

Kandel, ER, Markram, H, Mathews, POM, Yuste, R, and Koch, C (2013). Neuroscience thinks big (and collaboratively). Nature Rev Neurosci. 14, 659-664.

Kogge P (2011). http://spectrum.ieee.org/computing/hardware/nextgeneration-supercomputers.

Kubera Y, Mathieu P, Picault S (2008). Interaction-Oriented Agent Simulations: From Theory to Implementation. In Ghallab, M, *et al.* (eds), Proceedings of the 18th European Conference on Artificial Intelligence, IOS Press, pp.383-387.

Loewenstein Y, Sompolinsky H (2003). Temporal integration by calcium dynamics in a model neuron. Nature Neurosci 6, 961-967.

Markram H (2006) The blue brain project. Nature Rev Neurosci 7, 153-160.

Merolla PA, Arthur JV, Alvarez-Icaza, et al. (2014). A million spiking-neuron integrated circuit with a scalable communication network and interface. Science 345, 668-673.

Norman KA (2010). How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. Hippocampus 20, 1217-1227.

Schneidman E, Freedman B, Segev I (1998). Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. Neural Comput. 10, 1679-1703.

Tulving, E. (1985). Memory and consciousness. Canad. Psychol. 26, 1-12.

Vico, G. (1710/1988). On the most ancient wisdoms of the Italians unearthed from the origins of the Latin language. Ithaca: Cornell University Press.

Weinberg S (1992). Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature. NY: Pantheon.