Bartosz Brożek
Jagiellonian University
Copernicus Center for Interdisciplinary Studies

Philosophy and Neuroscience
Three Modes of Interaction[1]

Introduction

The relationship between philosophy and neuroscience[2] may be – and has been – accounted for in various ways. The first option is to claim that philosophy is firmly isolated from neuroscience – that they are mutually independent, or even that philosophical considerations are, in some way or another, *a priori* to neuroscientific endeavours. The second view is that of replacement: it may be argued that science in general – and neuroscience in relation to some philosophical issues – answers all the important questions of philosophy, but

---

[1] This paper was written within the research grant "The Limits of Scientific Explanation" sponsored by the John Templeton Foundation. It is largely based on, and further develops the ideas expressed in my previous publications: B. Brożek, *Philosophy in Neuroscience*, [in:] *Philosophy in Science*, eds. B. Brożek, J. Mączka, W. Grygiel, Copernicus Center Press, Kraków 2011, pp. 163–188; B. Brożek, A. Olszewski, *Logika zapętleń*, [in:] *Oblicza racjonalności*, eds. B. Brożek, J. Mączka, W. Grygiel, M. Hohol, Copernicus Center Press, Kraków 2011, pp. 33–50; and B. Brożek, *Rule-following. From Imitation to the Normative Mind*, Copernicus Center Press, Kraków 2013.

[2] The term "neuroscience" is understood very broadly here, and refers to, *inter alia*, behavioural neuroscience, cellular neuroscience, clinical neuroscience, cognitive neuroscience, computational neuroscience, cultural neuroscience, developmental neuroscience, molecular neuroscience, neuroimaging, neuroengineering, neuroinformatics, neurolinguistics, social neuroscience and systems neuroscience. However, I believe that my conclusions are applicable, mutatis mutandis, to any refinement of the term, i.e. to any reasonable rendering of 'neuroscience *sensu stricto*'.

does so in a better way – not through mere speculation, but on the basis of solid empirical data. On this account, scientific scrutiny has already replaced much of what used to be the domain of philosophical argument. I believe that both these stances are erroneous; however, it is interesting – and highly relevant – to ask, why this is the case. The analysis provided below suggests that both mistakes – of isolationism and of replacement – can be traced back to the same source: the assumption that either philosophy or neuroscience provides us with foundational, unshakable knowledge. I posit that only the rejection of this assumption – i.e., the appreciation of the fact that in philosophy, as well as in neuroscience, one has to do with non-foundational argumentation – opens the way for a truthful description of the interplay between the two disciplines and to the claim that they can *enrich* one another in many ways.

1. Isolation

The isolationist projects are based on the assumption that the findings of the natural sciences, neuroscience included, do not influence directly the practice of philosophy, as at least some aspects of philosophical reflection are independent of empirical facts investigated by scientists. The doctrine of isolation takes various shapes and forms, from some incarnations of the classical philosophy of Aristotle and Aquinas, through Kant's critical project and post-Kantian philosophies, such as phenomenology, to some versions of analytic philosophy. Let us have a look at two instructive examples.

The defenders of the contemporary versions of Thomism underscore the *autonomy* of philosophical thinking:

> The autonomy of Thomism boils down to the fact that its point of departure, as well as justification criteria, are independent of the truth of revelation as well as the findings of the natural sciences. The results of those disciplines can only (and often do) constitute the source of inspi-

> ration for new philosophical questions and determine new issues for metaphysical reflection. The maximalism (of Thomism) is connected to the fact that the goal of philosophizing is to uncover the fist and ultimate causes of the entire reality, including the cause of all causes – the Absolute, which makes the world intelligible and frees philosophical explanation from absurdity.[3]

Thus, the representatives of Thomism stress repeatedly that the autonomy of philosophy hangs together with its specific object and method: while the sciences consider only the so-called proximate causes of things, philosophy is capable of uncovering the ultimate causes of reality. Because of that, no empirical finding can falsify – or serve as a means for the rejection of – a philosophical theory. One should rather speak of two separate *planes of reflection*, the philosophical and the scientific, and if there is any relationship between them, it is the philosophical method that represents a higher, more profound mode of cognition.

This is an example of foundationalism in philosophy. Thomists believe that there exists only one true view of reality, captured by the Aristotelian-Thomistic conceptual scheme and penetrable by the Aristotelian-Thomistic method. All three dimensions of this foundationalism – the ontological, conceptual and methodological – prevent the findings of neuroscience from having any bearing on philosophical discourse: the sciences investigate only the manifestations of substances, utilize a method which cannot account for beings *qua* beings, and hence take advantage of a conceptual scheme which is not translatable into the metaphysical conceptual scheme and is inferior to it. The problem is that any kind of foundationalism in philosophy leads to daring consequences, when the relationship between philosophy and science is considered. As Michael Heller puts it:

---

3 A. Maryniarczyk, *Tomizm*, [in:] *Powszechna encyklopedia filozofii*, http://www.ptta.pl/pef/.

> Today, after 300 years of the dynamic development [of the natural sciences], the employment of the strategy [of isolation] leads to two different kinds of danger. Firstly, some deep questions of obvious philosophical character (Did life originate from inanimate matter with no external factor at play? Is human brain only a perfect calculator?) may be rejected as no genuine philosophical issues (as they cannot be formulated within a given philosophical system). Secondly, artificial and highly confusing problems arise when one tries to speak of nature using a language which is inadequate for this purpose (i.e., a language of a certain philosophical system).[4]

What Heller stresses here is that the faith in a philosophical system – in unshakable ontological or conceptual foundations – may easily lead to dispensing with real problems and to devoting time and effort to pseudo-problems. A closed, isolated philosophical system, one that provides answers to any questions, but only such that can be formulated within its conceptual framework, generates neither truth nor understanding, and hence becomes a caricature of what philosophical reflection should be.

A more elusive, although equally destructive, is the isolationist stance defended by Bennett and Hacker in their celebrated book *Philosophical Foundations of Neuroscience*.[5] Bennett and Hacker insist, first, that one should clearly distinguish between two types of questions, conceptual and empirical:

> Distinguishing conceptual questions from empirical ones is of the first importance. (…) Conceptual questions antecede matters of truth and falsehood. They are questions concerning our forms of representation, not questions concerning the truth or falsehood of empirical statements. These forms are presupposed by true (and false) scientific

[4] M. Heller, *Nauki przyrodnicze a filozofia przyrody*, [in:] M. Heller, *Filozofia i Wszechświat*, Universitas, Kraków 2006, p. 28.

[5] M.R. Bennett, P.M.S. Hacker, *Philosophical Foundations of Neuroscience*, Wiley, Blackwell, Malden, Oxford 2003.

> statements and by correct (and incorrect) scientific theories. They determine not what is empirically true or false, but rather what does and what does not make sense. Hence conceptual questions are not amenable to scientific investigation and experimentation or to scientific theorizing. For the concepts and conceptual relationships in question are presupposed by any such investigations and theorizing.[6]

The final sentence of the quoted passage is of special interest. Bennett and Hacker claim that concepts are *a priori* to any scientific investigation. They believe further that the failure to notice this fact often leads to serious errors, and in particular the so-called mereological fallacy, common – as they stress – in the contemporary cognitive neuroscience. It consists in referring to the brain or its parts with concepts which are correctly applicable only to a person as a whole. They observe:

> [talking] of the brain's perceiving, thinking, guessing or believing, or of one hemisphere of the brain's knowing things of which the other hemisphere is ignorant, is widespread among contemporary neuroscientists. This is sometimes defended as being no more than a trivial *façon de parler*. But that is quite mistaken. For the characteristic form of explanation in contemporary cognitive neuroscience consists in ascribing psychological attributes to the brain and its parts in order to explain the possession of psychological attributes and the exercise (and deficiencies in the exercise) of cognitive powers by human beings.[7]

One may ask whether the problem Bennett and Hacker identify is a real one. It may be argued, for example, that such claims as 'the brain thinks' or 'the right hemisphere is responsible for decision-making' are not to be taken literally. Some fundamental linguistic intuitions and the basic knowledge of language are enough to realize that

[6] M. Bennett, D. Dennett, P. Hacker, J. Searle, *Neuroscience and Philosophy: Brain, Mind, and Language*, Columbia University Press, New York 2007, p. 4.
[7] *Ibid.*, p. 7.

such an utilization of the words ‘think’ or ‘decide’ is metaphorical or analogical. Bennett and Hacker are fully aware of this strategy to defend the existing neuroscientific idiom. They elaborate it further claiming that the strategy may be used in four different ways. First, one may insist that psychological concepts used by neuroscientists, e.g. ‘to think’, have a different, derivative meaning to the meaning of the terms in the ordinary language. Second, the ‘neuroscientific meaning’ of a term may be analogical or constitute some other extension of the meaning of the corresponding ordinary language concept. Third, it may also be treated as a homonym: ‘to think’ or ‘to decide’ in a description of brain functions may have an altogether different meaning than the corresponding ordinary language terms. Finally, such concepts in neuroscience may be treated as metaphorical expressions.

Bennett and Hacker believe, however, that the above described strategies fail. They put forward a number of arguments to back this claim; the most important and which is applicable to all four strategies is the following. According to Bennett and Hacker there exists a criterion that suffices to show that the use of psychological terms in neurobiology is neither a case of taking advantage of derivative meaning, nor of analogical, homonymous or metaphorical. The criterion in question is the analysis of *conclusions* that the neuroscientists draw from the claims such as ‘the brain thinks’. Let us have a look at an example. Colin Blakemore notes:

> We seem driven to say that such neurons [as they respond in a highly specific manner to, e.g., line orientation] have knowledge. They have intelligence, for they are able to estimate the probability of outside events – events that are important to the animal in question. And the brain gains its knowledge by a process analogous to the inductive reasoning of the classical scientific method. Neurons present arguments to the brain based on the specific features that they detect, arguments on which the brain constructs its hypothesis of perception.[8]

[8] *Ibid.*, p. 16.

In this passage Blakemore claims that ‘neurons possess knowledge’. However, he does not end here; on the basis of this observation he constructs a complex conception of the interaction between neurons and the brain, which utilizes almost exclusively psychological terminology (‘intelligence’, ‘inductive reasoning’, ‘construction of a perceptual hypothesis’). But why can’t we consider this *entire* passage as one complex metaphor? Blakemore, in a different context, observes:

> Faced with such overwhelming evidence for topographic patterns of activity in the brain it is hardly surprising that neurophysiologists and neuroanatomists have come to speak of the brain having maps, which are thought to play an essential part in the representation and interpretation of the world by the brain, just as the maps of an atlas do for the reader of them. (...) But is there a danger in the metaphorical use of such terms as ‘language’, ‘grammar’, and ‘map’ to describe the properties of the brain? (...) I cannot believe that any neurophysiologist believes that there is a ghostly cartographer browsing through the cerebral atlas. Nor do I think that the employment of common language words (such as map, representation, code, information and even language) is a conceptual blunder of the kind [imagined]. Such metaphorical imagery is a mixture of empirical description, poetic license and inadequate vocabulary.[9]

Here, however, Bennett and Hacker launch their counter-attack. They ask how one should understand such claims as ‘the brain interprets the world’. They suggest that Blakemore’s use of ‘metaphorical’ expressions such as ‘a map’ leads directly to the utilization of inadequate terminology in the entire argumentation. This shows, as they believe, that there is no metaphor here; rather, Blakemore commits the mereological fallacy. In addition, they observe:

[9] *Ibid.*, p. 32.

> whatever sense we can give to Blakemore's claim that 'brain-maps' (which are not actually maps) play an essential part in the brain's 'representation and interpretation of the world', it cannot be 'just as the maps of an atlas do for the reader of them'. For a map is a pictorial representation, made in accordance with conventions of mapping and rules of projection. Someone who can read an atlas must know and understand these conventions, and read off, from the maps, the features of what is represented. But the 'maps' in the brain are not maps, in this sense, at all. The brain is not akin to the reader of a map, since it cannot be said to know any conventions of representations or methods of projection or to read anything off the topographical arrangement of firing cells in accordance with a set of conventions. For the cells are not arranged in accordance with conventions at all, and the correlation between their firing and features of the perceptual field is not a conventional but a causal one.[10]

Bennett and Hacker's position is that the evidence that neuroscientists commit the mereological fallacy does not lie in the fact that on occasions they use 'inadequate' psychological terms to describe the functioning of the brain, which may easily count as taking advantage of analogy, metaphor, homonym or using a concept with a derivative meaning. The mereological fallacy results when neuroscientists transfer entire complexes of concepts from 'psychological discourse' to the 'neuroscientific' one, and – on the basis of such inadequate attributions – they draw conclusions.

Is Bennett and Hacker's argument tenable? I believe not, and the reason is analogous to the case of the Thomistic isolationist project: their foundationalism. There are two interpretations of Bennett and Hacker's foundationalism. The stronger interpretation, attributed to them by John Searle, is that they believe natural language to determine the only acceptable ontology. Searle says that they commit a fallacy:

---

[10] *Ibid.*, p.33.

> the fallacy, in short, is one of confusing the rules for using the words with the ontology. Just as old-time behaviorism confused the evidence for mental states with the ontology of the mental states, so this Wittgensteinian criterial behaviorism construes the grounds for making the attribution with the fact that is attributed. It is a fallacy to say that the conditions for the successful operation of the language game are conditions for the existence of the phenomena in question.[11]

This reading finds some textual evidence. Interestingly, while elaborating the doctrine of the mereological fallacy, Bennett and Hacker quote Aristotle as one of those who first condemned this erroneous mode of thinking. He observed that "to say that the soul is angry is as if one remarked that the soul weaves or builds, for it is surely better not to say that the soul pities, learns or thinks, but that a man does these with his soul". One needs to remember, however, there is a certain metaphysical view standing behind his claim. Aristotle's metaphysics is essentialist: he believes that every entity belongs to some natural category, one determined by the entity's essence (form); moreover, he believes that the essences may be captured by the so-called essential definitions.[12] Thus, the incorrect or metaphorical use of words is not a mere mistake – it is an error that may effectively ruin our attempts to construct the foundations of knowledge, captured by the essential definitions. This doctrine is, of course, far from actual scientific practice. The history of science shows clearly that no such foundations should be assumed as they are most likely to hinder scientific progress. But if so, the same holds for Bennett and Hacker's view: if they indeed believe that the conceptual scheme of the ordinary language determines 'the only' ontology, their conception is hopelessly flawed.

It is also possible to read Bennett and Hacker in a more moderate manner; this weaker interpretation is that they only underscore that

---

[11] *Ibid.*, p.105.

[12] See K. Popper, *The Open Society and Its Enemies*, vol. II, Princeton University Press, Princeton 1966, p. 30.

the conceptual scheme which constitutes the framework for the ordinary language does not determine any unique ontology, but nevertheless is independent of any scientific practice, in the sense that in order to communicate any scientific discovery one needs to employ concepts according to some pre-existing criteria. If one does not do so, one risks following wrong paths and uttering nonsensical statements: the incorrect use of language can lead us astray. The conceptual scheme of ordinary language constitutes, at the very least, the foundation for communicating scientific theories.

This incarnation of foundationalism is equally troublesome as its Aristotelian-Thomistic predecessor. Firstly – and less importantly – Bennett and Hacker are mistaken when they claim that an excessive use of metaphors, and in particular – clusters of metaphors – is destructive for any neuroscientific endeavour. Certainly, it may lead to blind alleys yet there is little danger that the consequences of such a way of expression will be daring. The reason is that neuroscience, as with any other science, has some built-in corrective mechanisms that ultimately help us to distinguish progress and fruitful hypotheses from mere mistakes and useless conjecture. That this mechanism is present is evident once one considers the recent successes of neuroscience. A science which overuses metaphors and leads to no serious predictions or explanations is simply a bad science; the mere fact of committing or omitting the mereological fallacy is of no significance here.

Secondly, we should consider the bigger picture, which is encapsulated in Bennett and Hacker's claim that the conceptual scheme of ordinary language is *a priori* relative to scientific practice. It is particularly troublesome with respect to neuroscience. It must be realized that the psychological idiom, characteristic of the ordinary language, is not only shaped by our inner experience, but also by the theories developed throughout history which aimed at conceptually capturing mental phenomena. The problem is that the conceptual scheme of ordinary language is characterized by some inertia: it takes a considerable amount of time for current scientific conceptions to 'infiltrate' our ordinary conceptual scheme. It is safe, therefore, to assume that today's ordinary language 'embraces' some psychological theories of yesterday, or better even: a blend of those theories and common-sense ideas, often referred to as the folk psychology. Now, to say that ordinary concepts are a priori relative to neuroscience amounts to saying that folk psychology is *a priori* to neuroscientific theories, which is outright nonsense: it is one of the main goals of the contemporary neuroscience, one that it fulfils vigorously and with much success, to revise our old, common-sense psychological notions.

This clearly shows that Bennett and Hacker's conceptual foundationalism is faithful neither to the mechanisms of scientific practice, nor the way our conceptual schemes evolve: they are never final, or independent of the theories we develop. This point is quite general and pertains to any foundational philosophical project: the sources of philosophical reflection are always, at least partially, based on *some* scientific conceptions, although often on outdated ones. In the case of Thomism, the Aristotelian view of the world – or the Aristotelian science – constitutes the foundations of the conceptual scheme. Similarly, in the case of those philosophies that find confirmations or disconfirmations in the workings of the ordinary language, it is the knowledge encapsulated there (e.g., a kind of folk psychology being a blend of the common sense observations and some old psychological theories) that ultimately determines the philosophical doctrines of the followers of Austin and Strawson. In other foundational projects, such as phenomenology, the scientific knowledge internalized by any given person crucially shapes this person's experiences, and so her philosophical views. All in all, there is no source of philosophical knowledge that would be independent of some kind of science, and the key point is that this 'hidden science' may be at odds with what the contemporary science has to say. The splendid isolation of philosophers is an illusion: there is no escape from the confrontation with the barbarians from the other side of the Channel.

2. Replacement

The second stance towards the relationship between philosophy and neuroscience is one that sees an inherent conflict between the two. This seems to be mainly the perspective adopted by neuroscientists who believe that the findings of neuroscience resolve philosophical problems and, what follows, scientific method and scientific theories should *replace* philosophical reflection. Thus, there is a conflict here, but one in which there may be only one survivor. The problem is that – similarly to the case of isolationism – those who believe in the replacement strategy base their convictions on a foundational view of knowledge.

In order to illustrate how neuroscientists tend to approach philosophical problems, let us have a look at the neuroscience of mathematics. This field of knowledge has made impressive progress during the past twenty years. Although there is no single, commonly accepted theory of the neural basis of mathematical skills, a general picture that emerges from the findings of neuroscience may be deemed the 3E view of mathematics; mathematics which is embrained, embodied and embedded. The general elements of this view are the following. Firstly, the current research implicates two separate brain systems as responsible for the basic numerical capacities: the object tracking system (OTS) and the approximate number system (ANS). OTS is a system that enables the tracking of multiple individuals (up to 3 or 4). It is based on the principles of cohesion (moving objects are recognized as bounded wholes), continuity (objects move on unobstructed paths) and contact (objects do not interact at a distance).[13] The existence of the OTS system is confirmed by a number of tests, including visual short-term memory tasks, multiple-objects tracking tasks, or enumeration tasks. The last kind of tests confirms human

[13] M.Piazza, *Neurocognitive Start-Up Tools for Symbolic Number Representations*, [in:] *Space, Time and Number in the Brain*, eds. S. Dehaene, E. Brannon, Academic Press, London 2011, p. 270.

ability of *subitizing*, i.e. of an instant and highly accurate determination of a number of objects in small collections (3–4), even presented very briefly.[14] Furthermore, it is speculated that the posterior parietal and occipital regions of the brain play the crucial role in the performance of such tasks, which suggests that these regions are the location of OTS.[15]

ANS, on the other hand, is a system for representing the approximate number of items in sets. It works according to Weber's Law: the threshold of discrimination between two stimuli increases linearly with stimulus intensity. In the case of ANS, the Weber fraction, or the smallest variation to a quantity that can be readily perceived, changes over human development. For newborns it is 1:3, for 6-month-old babies it is 1:2, for 1-year-old children it is 2:3, for 4-year-olds it is 3:4, for 7-year-olds it is 4:5, while for 20-year-olds it is 7:8. It means that a newborn can discriminate between 1 and 3, or 2 and 6, or 10 and 30, but not 1 and 2, 2 and 5, or 10 and 27. Four-year-old children can tell that there is a difference in numerosity between sets consisting of 6 and 8 or 12 and 16 elements, but not 7 and 8 or 12 and 15. Adults' ANS system is even more 'sensitive': they can discriminate (without counting) between sets consisting of 14 and 16 elements or 70 and 80 elements, but not 70 and 78 elements. It is quite well established that ANS is located in the mid-intreparietal sulcus.[16]

OTS and ANS constitute innate or 'embrained' arithmetic skills and are clearly quite limited. The question is, how people move beyond these limited inborn abilities and acquire 'full-blooded' mathematical skills. There are a number of hypotheses explaining this process, but most of them point out to the key role of language in both phylogenetic and ontogenetic development of arithmetical ability. The hypothesis that the development of counting skills is conditioned and mediated by the acquisition of language is supported

[14] *Ibid.*, p. 271.
[15] *Ibid.*, p. 270.
[16] *Ibid.*, p. 268–269.

by the following evidence. First, both children and adults in remote cultures, whose languages have no words for numbers, when dealing with numbers larger than three only recognize their equivalence approximately. Second, an interesting line of evidence comes from the study of the mathematical abilities of deaf people. Deaf persons living in numerate cultures but not exposed to the deaf community use a gestural system called homesign; they use fingers to communicate numbers, but only with approximate accuracy. Similarly, they perform matching tasks with approximate accuracy.[17] Moreover, language continues to play an important role in mathematical cognition even after the mathematical skills are mastered. For instance, educated adults who suffer language impairments have problems with exact, but not approximate numerical reasoning. Similarly, when doing exact (but not approximate!) tasks, adults spend more time with numbers that are difficult to pronounce, even if they are presented in Arabic notation. But "if language merely scaffolded the acquisition of natural number concepts and abilities, and then was replaceable by other symbol systems, one would not expect adults to translate Arabic symbols into words for purposes of exact computation".[18] Finally, bilingual adults who are taught some new mathematical facts in one of their languages have difficulties in the smooth production of exact number facts in the other language.[19]

These observations lead to an important question: if the acquisition of language plays such a crucial role in the ontogenetic development of mathematical skills, what exactly is the language-dependent mechanism that enables to move beyond embrained mathematics? One of the most intriguing, albeit controversial, answers to this question was formulated Lakoff and Núñez, within the so-called 'embodied mind' paradigm. The very idea of embodiment boils down to the thesis that human mind and human cognition are decisively shaped

[17] E.S. Spelke, *Natural Number and Natural Geometry*, [in:] *Space, Time and Number in the Brain*, eds. S. Dehaene, E. Brannon, Academic Press, London 2011, p. 306–307.
[18] *Ibid.*, p. 307.
[19] *Ibid.*, p. 307.

by the experiences of our bodies. This is a vague claim that only underscores the rejection of other paradigms, such as Cartesian mind-body dualism or computationalism (the rough idea that human brain is hardware, and the mind is software implemented in the brain). However, other claims of the representatives of the embodied mind approach are more informative. They believe that the human mind is a powerful conceptual system shaped by our bodies' experiences during their interactions with the environment. The most basic mental concepts or schemas, probably derived from the neural motor-control programs, express spatial relations (such as the Source – Path – Goal schema). Since such "image schemas are conceptual in nature, they can form complex composites. For example, the word 'into' has a meaning – the *Into* schema – that is the composite of an *In* schema and *To* schema".[20] Furthermore, this mental machinery is capable of producing abstract concepts with the use of concrete ones through the use of metaphors. In the 'embodied paradigm' metaphors are understood as the means for "understanding and experiencing one kind of thing in terms of another".[21] And so, importance is conceptualized in terms of size ("This is a big issue", "It's a small issue; we can ignore it"), difficulties are conceptualized as burdens ("He is overburdened", "I've got a light load this semester"), etc.[22] "Each such conceptual metaphor has the same structure. Each is a unidirectional mapping from entities in one conceptual domain to corresponding entities in another conceptual domain. As such, conceptual metaphors are part of our system of thought. Their primary function is to allow us to reason about relatively abstract domains using the inferential structure of relatively concrete domains".[23]

Lakoff and Núñez claim further that it is the mechanism of conceptual metaphorization that enables the construction of complex and

[20] G. Lakoff, R. Núñez, *Where Mathematics Comes From*, Basic Books, New York 2000, p. 39.
[21] *Ibid.*, p. 5.
[22] *Ibid.*, p. 41.
[23] *Ibid.*, p. 42.

precise mathematical concepts. In the case or arithmetic, they postulate the existence of four basic or grounding metaphors: the Arithmetic as Object Collection (where the source domain concept of collections of objects of the same size is mapped to the concept of numbers, the size of the collection is mapped to the size of the number, the smallest collection is mapped to the concept of the unit, while putting collections together is mapped to the process of addition); the Arithmetic as Object Construction (where the source domain concept of objects consisting of ultimate parts of unit size is mapped to the concept of numbers or the act of object construction is mapped to the concept of arithmetic operations); the Measuring Stick (where physical segments are understood as numbers, the basic physical segment as one, and the length of the physical segment as the size of the number); and the Arithmetic as Motion Along a Path (where the act of moving along the path is understood as representing mathematical operations, point-locations on the path are understood as numbers, etc.).[24]

Lakoff and Núñez claim that those four grounding metaphors give rise to the development of basic arithmetic. One begins with innate capacities to 'deal' with small numbers (up to 4). In addition, one has primary experiences with object collections, object construction, physical segmentation and moving along a path. "In functioning in the world, each of those primary experiences is correlated with subitizing, innate arithmetic, and simple counting".[25] Those two domains are combined through the four metaphors in such a way that the primary experiences become sources of the metaphors and the domain of numbers is the target of the metaphors. "The next step is the conflation among the primary experiences: object construction always involves object collection. Placing physical segments end to end is a form of object construction (…). From a neural perspective, [such conflations] involve co-activations of those brain areas that characterize each of the experiences. (…) As a consequence, an isomorphic

[24] *Ibid.*
[25] *Ibid.*, p. 93.

structure emerges across the source domains (…), which is independent of numbers themselves and lends stability to arithmetic".[26] The ability of subitizing, found in all normal human beings, leads to precise and stable results regarding small numbers; when extended with the four grounding metaphors, the precision and stability extends to all natural numbers. Finally, "the laws of arithmetic (commutativity, associativity and distributivity) emerge first as properties of the four source domains, then as properties of numbers *via* those metaphors, since the metaphors are inference-preserving conceptual mappings".[27]

If this picture is roughly correct, mathematics – although based on some simple innate or embrained skills – is also embodied and embedded. It is largely derived – through the process of metaphorization – from concrete concepts shaped by our bodies' interactions with the environment, as well as sustained and further developed through the interactions with others by imitating their behaviour. The embedding of mathematical practice in social interactions contributes to the stability of mathematical knowledge.

There are many philosophical problems connected to mathematics. Two of the most fundamental are the problem of the necessity in mathematics, and the problem of the mathematicity of the universe. Let us begin with the former. There is a dimension of mathematical and logical research that traditionally poses a challenge to any naturalistic accounts of the ontology of mathematical or logical objects. It is well captured in the following observation by Jan Łukasiewicz:

> Whenever I deal with the smallest logical problems, I always have the feeling that I am facing some powerful, incredibly coherent and enormously resistant structure. I cannot make any changes within it, I create nothing, but working hard I uncover new details, gaining eternal truths.[28]

[26] *Ibid.*, p. 95–96.
[27] *Ibid.*, p. 96.
[28] J. Łukasiewicz, *W obronie logistyki. Myśl katolicka wobec logiki współczesnej*, "Studia Gnesnensia" 1937, no. 15, p. 14.

Such views as the one expressed by Łukasiewicz give rise to the development of mathematical Platonism (realism), a view that "mathematics is the scientific study of objectively existing mathematical entities just as physics is the study of physical entities. The statements of mathematics are true or false depending on the properties of those entities, independent of our ability, or lack thereof, to determine which".[29]

There are many forms of mathematical Platonism. In particular, one should distinguish between ontological Platonism (a view pertaining to the existence of mathematical objects) and semantic Platonism (an epistemological view that mathematical statements are true or false). Ontological Platonism is a stronger theory – it implies the semantic one, but the opposite implication does not hold. Thus, in what follows I shall concentrate on the stronger claim. Arguably, ontological Platonism in mathematics, though it comes in various incarnations, embraces the following three theses:

> (The existence thesis) Mathematical objects (or structures) exist.
> (The abstractness thesis) Mathematical objects are abstract, non-spatio-temporal entities.
> (The independence thesis) Mathematical objects are independent of any rational or irrational activities of the human mind. In particular, mathematical objects are not our constructions.[30]

The key question is how the above formulated theses are justified. With no pretence with regards comprehensiveness, I posit that there are three kinds of arguments backing mathematical Platonism in its ontological version. The first one is the *semantic argument*, well captured by Balaguer, but formulated earlier by Frege[31]:

---

[29] P. Maddy, *Realism in Mathematics*, Oxford University Press, Oxford 1990, p. 21.
[30] Cf. Ø. Linnebo, *Platonism in the Philosophy of Mathematics*, [in:] *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), ed. E.N. Zalta, URL = <http://plato.stanford.edu/archives/fall2011/entries/ platonism-mathematics/>.
[31] Cf. *Ibid.*

(a) Mathematical propositions are true.
(b) Mathematical propositions should be taken at their face value. In other words, there is no reason to believe that mathematical propositions, as they appear, are not what they really are, or that there is a deep structure of mathematical propositions which differs from their surface structure, of what they seem at their face.
(c) By Quine's criterion, we are ontologically committed to the existence of objects which are values of the variables in the propositions we consider true.
(d) We are ontologically committed to the existence of mathematical objects.
(e) Therefore, there are such things as mathematical objects, and our theories provide true descriptions of these things. In other words, mathematical Platonism is true.

The second argument defending mathematical Platonism is *the indispensability argument*, or the Quine/Putnam argument. Muddy summarizes it in the following way: "we are committed to the existence of mathematical objects because they are indispensable to our best theory of the world and we accept that theory".[32] And in Putnam's own words: "mathematics and physics are integrated in such a way that it is not possible to be a realist with respect to physical theory and a nominalist with respect to mathematical theory".[33] A reconstruction of this argument may appear as follows:

(a) By Quine's criterion, we are committed to the existence of objects which our best physical theories speak of.
(b) Our best physical theories are expressed with the use of the language of mathematics.

[32] P. Maddy, *Realism…*, *op. cit.*, p. 30.
[33] H. Putnam, *What is mathematical truth?*, [in:] H. Putnam, *Mathematics, Matter and Method*, Cambridge University Press, Cambridge 1979, p. 74.

(c) Therefore, we are committed to the existence of mathematical objects.
(d) When one is a realist with respect to physical theories, one must also be a realist with respect to mathematics.
(e) Therefore, mathematical Platonism is true.

Finally, Gödel's *intuition-based argument* may be reconstructed in the following way:

(a) The most elementary axioms of set theory are obvious; as Gödel puts it, they "force themselves upon us as being true"[34].
(b) In order to explain (a), one needs to posit the existence of mathematical intuition, a faculty analogous to the sense of perception in the physical sciences.
(c) Not all mathematical objects are intuitable; but our belief in the 'unobservable mathematical facts' is justified by the consequences they bring in the sphere controllable by intuition and through their connections to already established mathematical truths. As Gödel says, "even disregarding the [intuitiveness] of some new axiom, and even in case it has no [intuitiveness] at all, a probable decision about its truth is possible also in another way, namely, inductively by studying its 'success' (…). There might exist axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving problems (...) that, no matter whether or not they are [intuitive], they would have to be accepted at least in the same sense as any well-established physical theory".[35]

Let us consider now, whether the 3E account of mathematics sketched above has any bearing on the arguments favouring mathe-

[34] K. Gödel, *What is Cantor's continuum problem?*, [in:] *Philosophy of Mathematics*, eds. P. Benacerraf, H. Putnam, Cambridge University Press, Cambridge 1983, p. 484.
[35] *Ibid.*, p. 477.

matical Platonism. Lakoff and Núñez believe that the conception of the embodied mathematics puts mathematical Platonism to eternal rest. For them, mathematical Platonism is "the romance of mathematics", a "story that many people *want* to be true"[36]; a story that mathematical objects are real, and mathematical truth is universal, absolute, and certain. They succinctly reject this view:

> The only access that human beings have to any mathematics at all, either transcendent or otherwise, is through concepts in our minds that are shaped by our bodies and brains and realized physically in our neural systems. For human beings – or any other embodied beings – mathematics *is* embodied mathematics. The only mathematics we can know is mathematics that our bodies and brains allow us to know. For this reason, the *theory of embodied mathematics* (…) is anything but innocuous. As a theory of the only mathematics we know or can know, it is a theory of what mathematics *is* – what it really is![37]

As I read them, Lakoff and Núñez emphasise two things. First, they put forward an epistemological claim that we have no cognitive access to independent abstract objects, since the only way of practicing mathematics is through the concepts "shaped by our bodies and brains". The problem of the cognition of abstract objects has been a subject of controversy since the beginnings of philosophy. Painting with a broad brush, one may claim that two solutions have been defended in this context, both already present in Plato: that there exists a rational intuition enabling us to contemplate abstract objects or that our access to the abstract sphere is discursive, mediated by language. Lakoff and Núñez seem to consider only the first option, and dismiss it on the basis of the recent findings in neuroscience.

Second, they seem to embrace a version of Quine's criterion: we are committed to the existence of only those things which our best

[36] Cf. G. Lakoff, R. Núñez, *Where…*, *op. cit*.

[37] *Ibid.*, p. 346.

scientific theories speak of. They add that the best – or rather: the only – theory of mathematical cognition we have is the theory of embodied mathematics, and since it does not speak of independent abstract objects, we have no grounds for postulating their existence. The problem is that Quine's criterion – applied to other theories, not necessarily accounting for the nature of mathematics, e.g. to our best physical theories – brings a different outcome: that we are indeed committed to the existence of abstract mathematical objects.

A similar line of argument is developed by Stanislas Deheane. Here is a longer passage that encapsulates his view well:

> For an epistemologist, a neurobiologist, or a neuropsychologist, the Platonist position seems hard to defend — as unacceptable, in fact, as Cartesian dualism is unacceptable as a scientific theory of the brain. Just as the dualist hypothesis faces insurmountable difficulties in explaining how an immaterial soul can interact with a physical body, Platonism leaves in the dark how a mathematician in the flesh could ever explore the abstract realm of mathematical objects. If these objects are real but immaterial, in what extrasensory ways does a mathematician perceive them? This objection seems fatal to the Platonist view of mathematics. Even if mathematicians' introspection convinces them of the tangible reality of the objects they study, this feeling cannot be more than an illusion. Presumably, one can become a mathematical genius only if one has an outstanding capacity for forming vivid mental representations of abstract mathematical concepts — mental images that soon turn into an illusion, eclipsing the human origins of mathematical objects and endowing them with the semblance of an independent existence.[38]

Thus, Deheane also stresses the epistemological point: that we have no, and cannot have any, cognitive access to the realm of abstract objects. However, also in this case, the epistemological point has de-

[38] S. Deheane, *The Number Sense*, 2nd edition, Oxford University Press, Oxford 2011, p. 225.

vastating consequences for mathematical Platonism: if we have no access to the abstract sphere, there are no grounds for postulating its existence. It is clearly visible that Deheane, similarly to Lakoff and Núñez, does not even consider that our grasping of abstract objects may be enabled by language.

It seems, moreover, that Lakoff and Núñez's, as well as Deheane's critiques, do not defeat any of the three arguments in favour of mathematical Platonism described above. In order to defeat the semantic argument one would have either to show that mathematical propositions cannot be ascribed truth or falsehood; or to reject the idea that mathematical propositions have no 'deep structure': that they are what they seem at their face; or to reject Quine's criterion of ontological commitment. Neither Lakoff and Núñez, nor Deheane, do so. Also, they fail to address the indispensability argument. To do so, they would need either to reject Quine's criterion; or the thesis that mathematical physics is our best theory of the world; or the realist stance towards physical theories. Finally, the intuition-based argument seems the easiest to attack from the point of view of the neuroscience of mathematics. As we have seen above, human 'intuitive' mathematical capacities are substantially limited. However, Gödel – the proponent of the intuition-based argument – does not claim that our intuition is a faculty that gives us access to the entire world of mathematical structures. His thesis is that intuition is the source of certainty in relation to relatively simple mathematical structures and relations; more complicated mathematical propositions are evaluated as true because they are justified by commonly accepted mathematical methods and have consequences controllable at the intuitive level. Of course, Lakoff, Núñez and Deheane may claim that the intuition Gödel speaks of is not an intuition of *abstract* objects; it is rather the capacity to use abstract mathematical concepts, which are ultimately shaped by the experiences of our bodies. But this criticism can be softened by a modification of Gödel's argument: instead of speaking of intuition, one can simply speak of mathematical experience, even conceived of in terms of Lakoff and Núñez's of of Deheane's theory. The crux of

Gödel's thesis, or so I argue, lies somewhere else: mathematical Platonism is true, because "there exist axioms so abundant in their verifiable consequences, shedding so much light upon a whole field, and yielding such powerful methods for solving problems that, no matter whether or not they are [a subject of direct experience], they would have to be accepted at least in the same sense as any well-established physical theory". Gödel points out to something important here: the full power of our abstract conceptions, which lie beyond any intuitive or 'direct' experience, is clearly visible in the consequences they produce within the sphere controllable by experience, as well as in the coherence they bring to entire areas of mathematics and the heuristic role they play in solving mathematical problems. It is reasonable, therefore, to assume that those highly abstract concepts describe some independently existing structures rather than claim that they are just 'metaphorizations' of more concrete concepts. The mathematics we can somehow experience directly is only the tip of the iceberg: and when Lakoff and Núñez, as well as Deheane believe that the rest of the iceberg is only an illusion, Gödel seems to claim that it is a rock-hard, even if abstract, reality.

All this is not to say that the three arguments supporting mathematical Platonism are irrefutable or incontestable: the heated debates in the philosophy of mathematics during the last century are the evidence to the contrary. However, Lakoff, Núñez and Deheane failed to provide a persuasive case against mathematical Platonism. Moreover, along the way, they have themselves accepted some philosophical assumptions, such as that the only access to the abstract objects is through some kind of rational intuition; or Quine's existence criterion; or realism in relation to biological theories.

The second problem I would like to address is that of the mathematicity of the universe. Michael Heller introduces the concept in the following words:

> In the investigation of the physical world one method has proved particularly efficient: the method of mathematical modeling coupled

> with experimentation (to simplify, in what follows I shall speak of the mathematical method). The advances in physics, since it has adopted the mathematical method, have been so enormous that they can hardly be compared to the progress in any other area of human cognitive activity. This incontestable fact helps to make my hypothesis more precise: the world should be ascribed a feature thanks to which it can be efficiently investigated with the use of the mathematical method. Thus the world has a rationality of a certain kind – a mathematical one. It is in this sense that I shall speak of the mathematicity of the universe.[39]

According to Heller, to say that the world is mathematical is equivalent to the claim that it possesses a feature which makes mathematical method efficient. In the quoted passage, Heller hints at one of the aspects in which the mathematicity of the world should be understood: the efficiency thesis. It says that the mathematicity of the universe is evident once one considers the enormous success of the mathematical method over the last 300 years. The success cannot be a matter of coincidence, as the efficiency of mathematics in uncovering the laws of nature seems 'unreasonable'.[40] The argument pertaining to the 'unreasonable effectiveness of mathematics' is not trivial. As Eugene Wigner observes:

> It is true, of course, that physics chooses certain mathematical concepts for the formulation of the laws of nature, and surely only a fraction of all mathematical concepts is used in physics. It is true also that the concepts which were chosen were not selected arbitrarily from a listing of mathematical terms but were developed, in many if not most cases, independently by the physicists and recognized then as having been conceived before by the mathematicians. It is not true, however, as is so often stated, that this had to happen because

[39] M. Heller, *Czy świat jest matematyczny?*, [in:] M. Heller, *Filozofia i wszechświat*, Universitas: Kraków 2006, p. 48.
[40] Cf. E. Winger, *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, "Communications on Pure and Applied Mathematics" 1960, no. 13(1), pp. 1–14.

> mathematics uses the simplest possible concepts and these were bound to occur in any formalism. [Moreover], it is important to point out that the mathematical formulation of the physicist's often crude experience leads in an uncanny number of cases to an amazingly accurate description of a large class of phenomena. This shows that the mathematical language has more to commend it than being the only language which we can speak; it shows that it is, in a very real sense, the correct language.[41]

There are some phenomena connected to the use of the mathematical method that leads to the conclusion that it is some feature of the world that must be responsible for the method's successes. It is often the case that mathematical equations describing some aspects of the universe 'know more' than their creators. The standard story in this context is that of Einstein's cosmological constant. When Einstein formulated his cosmological equations on the basis of the newly discovered general relativity theory, he realized that they imply a dynamic, expanding universe. In order to 'stop' the expansion, he introduced the cosmological constant. It quickly proved, however, that Einstein was 'wrong' and his equations were 'right': the expansion of the universe is a fact.

Another instructive example is given by Wigner. When Heisenberg formulated his quantum mechanics based on matrix calculus, the theory was applicable only to a few idealized problems. Applied to the first real problem, of the hydrogen atom, it also proved successful:

> This was (...) still understandable because Heisenberg's rules of calculation were abstracted from problems which included the old theory of the hydrogen atom. The miracle occurred only when matrix mechanics, or a mathematically equivalent theory, was applied to problems for which Heisenberg's calculating rules were meaningless. Heisenberg's rules presupposed that the classical equations of motion had so-

[41] *Ibid.*, p. 7.

> lutions with certain periodicity properties; and the equations of motion of the two electrons of the helium atom, or of the even greater number of electrons of heavier atoms, simply do not have these properties, so that Heisenberg's rules cannot be applied to these cases. Nevertheless, the calculation of the lowest energy level of helium (...) agrees with the experimental data within the accuracy of the observations, which is one part in ten million. Surely in this case we 'got something out' of the equations that we did not put in.[42]

The second aspect of the mathematicity of the universe may be called *the miracle thesis*. It is possible to imagine worlds which are *mathematical* in a certain sense, yet non-idealizable. Michael Heller considers a hierarchy of such worlds. 'The most non-mathematical' is a world in which no mathematical and logical principles are observed (including any stochastic or probabilistic laws). Next, he suggests to consider a simplified model of the world: let us assume that the world in question may be in one of only two states, represented by '0' and '1'. Now:

> The history of this world is thus a sequence of '0's and '1's. Assume further that the world had a beginning, what may be represented by a dot at the beginning of the sequence. In this way, we get, e.g., a sequence:
>
> .011000101011...
>
> The task of a physicist is to construct a theory which would enable to predict the future states of the world. Such a theory would amount to the 'encapsulation' of the sequence of '0's and '1's in a formula (which is shorter than the sequence it encapsulates). Such a formula may be found only if the sequence of '0's and '1's is algorithmically compressible. But this leads to a problem. Such a sequence may be interpreted as a decimal expansion of a number in [0,1] and – as well known – the set of algorithmically compressible numbers

[42] *Ibid.*, p. 10.

> belonging to [0,1] is of measure 0 (…). Thus (…) there is zero-measure chance that a sequence of '0's and '1's, representing our world, belongs to the set of algorithmically compressible sequences and so the physicist, who investigates such a world, may have no rational expectation to discover the theory she is looking for.[43]

This observation underscores 'the other side' of the mathematicity thesis: not only is the universe mathematical (and hence penetrable by *some* mathematical method), but it is also mathematical in a non--malicious way (and hence penetrable by *our* mathematical methods).

In connection to the problem of the mathematicity of the world, Lakoff and Núñez claim:

> No one observes laws of the universe as such; what are observed empirically are regularities in the universe (…); laws are mathematical statements made up by human beings to attempt to characterize those regularities experienced in the physical universe. (…) What [the physicists] do in formulating 'laws' is fit their human conceptualization of the physical regularities to their prior human conceptualization of some form of mathematics. All the 'fitting' between mathematics and physical regularities of the physical world is done within the minds of physicists who comprehend both. The mathematics is in the mind of the mathematically trained observer, not in the regularities of the physical universe.[44]

This, again, is an example of bad philosophy. Lakoff and Núñez fail to realize the far-reaching consequences of the efficiency thesis. What they leave unaccounted for are, at least, the fact that the mathematical method helped us to conquer the micro-scale phenomena; that equations often 'know more' than their creators; that mathematical models are often the basis for formulating qualitatively *new* predic-

[43] M. Heller, *Czy świat jest matematyczny?*, *op. cit.*, p. 51–52.
[44] G. Lakoff, R. Núñez, *Where…*, *op. cit.*, p. 344.

tions, and so serve as powerful heuristic tools. It seems that behind Lakoff and Núñez's observations there lies a very simplistic or naive view of science: that scientific progress comes from the *observations* of the regularities of real-world phenomena and their generalizations into the mathematically expressible laws of physics. What follows, within Lakoff and Núñez's framework one cannot even formulate the miracle thesis.

Stanislas Deheane offers a more sophisticated argument to explain – within a naturalistic framework provided by the recent findings of neuroscience – the effectiveness of the mathematical method. He notes:

> How can one explain the extraordinary adequacy of the purest products of the human mind to physical reality? In an evolutionary framework, perhaps pure mathematics should be compared to a rough diamond, raw material that has not yet been submitted to the test of selection. Mathematicians generate an enormous amount of pure mathematics. Only a small part of it will ever be useful in physics. There is thus an overproduction of mathematical solutions from which physicists select those that seem best adapted to their discipline — a process not unlike the Darwinian model of random mutations followed by selection. Perhaps this argument makes it seem somewhat less miraculous that, among the wide variety of available models, some wind up fitting the physical world tightly[45].

The problem with this argument is that it does not touch what is the crux of the efficiency thesis. In the above quoted passage Wigner admits that only a small fraction of our mathematical theories find application in physics; but his astonishment stems from a different source: given that mathematics provides us with "an amazingly accurate description of a large class of phenomena" and that physical equations often 'know more' than their creators, shows "that the mathematical language

[45] S. Deheane, *The Number…*, *op. cit.*, p. 232–233.

has more to commend it than being the only language which we can speak; it shows that it is, in a very real sense, the correct language".

Deheane continues by observing that:

> In the final analysis, the issue of the unreasonable effectiveness of mathematics loses much of its veil of mystery when one keeps in mind that mathematical models rarely agree *exactly* with physical reality. Kepler notwithstanding, planets do not draw ellipses. The earth would perhaps follow an exact elliptic trajectory if it were alone in the solar system, if it was a perfect sphere, if it did not exchange energy with the sun, and so on. In practice, however, all planets follow chaotic trajectories that merely resemble ellipses and are impossible to calculate precisely beyond a limit of several thousand years. All the "laws" of physics that we arrogantly impose on the universe seem condemned to remain partial models, approximate mental representations that we ceaselessly improve[46].

This is a gross misunderstanding. Surely, mathematical physics takes advantage of idealizations: in constructing mathematical models of reality, the strategy is to disregard a number of aspects of the phenomena under consideration. But this is exactly what has made science possible. Moreover, the fact that idealization is possible says something about the universe: that it has a feature thanks to which it can be efficiently investigated with the use of mathematical method, a method that takes advantage of idealization. This means, however, that idealization is possible *because* the universe is mathematical (in the Heller's sense of the word).

I believe that it is relatively easy to point to one of Deheane's assumptions which prevents him from appreciating that the possibility of idealization constitutes an argument in favour of the mathematicity of the universe. He says that "all the 'laws' of physics that we arrogantly impose on the universe seem condemned to remain par-

[46] *Ibid.*, p. 233.

tial models, approximate mental representations that we ceaselessly improve".[47] Notice the use of the verb 'impose': it suggests that, for Deheane, mathematics is only what our minds construct. With such a definition, the conclusion that the world is not mathematical trivially follows. However, this is not what philosophers have in mind while speaking of mathematical Platonism or the unreasonable effectiveness of mathematics in the natural sciences; they rather claim that our mathematics, i.e. the mathematical theories we have developed, somehow captures or 'resonates' with mathematical reality, be it the Platonic universe of pure abstract objects or some aspect of the physical universe. As Michael Heller puts it:

> It is obviously true that genetically our mathematics comes from the world: we abstract some of its features. However, one needs to carefully distinguish between *our mathematics* and *mathematics as such*. Our mathematics (which I also deem 'mathematics with a small m') has been developed by humans in a long evolutionary process: it is expressed in a symbolic language we invented; its results are collected in our scientific journals, books, or computer memory. But our mathematics is only a reflection of certain relations or structures, which governed the movement of atoms and stars long before biological evolution began. I deem those relations or structures mathematics as such (or 'Mathematics with a capital M'); it is what we think of when we ask, why nature is mathematical. The answer to this question, which posits that the nature is mathematical because mathematics has been abstracted from nature, turns out helpless, or even naïve, when one introduces the distinction between our mathematics and mathematics as such.[48]

Thus, it seems that neither Lakoff and Núñez nor Deheane are able to provide any tenable answer to the efficiency thesis; with regards the

---

[47] *Ibid.*, p. 239.

[48] M. Heller, *Co to znaczy, że przyroda jest matematyczna?*, [in:] *Matematyczność przyrody*, eds. M. Heller, J. Życiński, Petrus, Kraków 2010, p. 16.

miracle thesis, they do not even formulate it. However, the interesting fact is that the conception of mathematics, which draws on their theories, may shed some light on the efficiency thesis. The argument is quite general. Both our inborn mathematical capacities, as well as our conceptual apparatus have been shaped – in the evolutionary process – by our interactions with the environment. Now, given that our environment is mathematical (in Heller's sense of the word), it helps us to understand why our mathematical concepts are efficient in uncovering the laws of the universe. Of course, such an argument cannot explain fully the efficiency of mathematics in quantum physics, or the fact that physical equations sometimes 'know more' than their creators. However, it may serve to dismiss the idea that "all the 'fitting' between mathematics and physical regularities of the physical world is done within the minds of physicists who comprehend both. The mathematics is in the mind of the mathematically trained observer, not in the regularities of the physical universe". On the contrary: the mind is mathematical because it is a part of the mathematical universe.

The views of Lakoff, Núñez, and Deheane illustrate nicely that the approach of (some) neuroscientists towards philosophical problems is a kind of *scientific foundationalism*: they seem to believe that neuroscience provides us with knowledge freed from philosophical assumptions, when the opposite is true – scientific theories are often intertwined with philosophical contents. For instance, in the above described example, at least the following philosophical doctrines have been embraced by neuroscientists: that an alleged rational intuition is the only mode of grasping abstract objects (this assumption seems to be shared by Lakoff and Núñez, as well as Deheane); that science (including physics) proceeds by induction and abstraction from observations (Lakoff and Núñez); and that the only mathematics one can speak of is the mathematics we constructed (Deheane). I hope to have illustrated that these are not the only philosophical stances one may adopt. This is not to say that there is no controversy here, e.g. that there really exists Mathematics with the capital M or that the assumptions embraced by Lakoff, Núñez and Deheane are untenable;

to the contrary, they are perfectly acceptable *philosophical* claims. The illusion is, however, that these issues have already been settled by neuroscience.

3. Enrichment

I argue that in order to depict the relationship between philosophy and neuroscience one needs to acknowledge that neither discipline is *isolated* from the other, nor provide *foundations* for the other. Argumentation – both in philosophy and in neuroscience – is *non-foundational*.

In his essay "Against Foundationalism" Michael Heller observes that each philosophical argument has two components: the deductive and the hermeneutic:

> I believe that all arguments in philosophy, but also in the sciences, can be arranged in a sequence, such that at its – say – left end there are arguments without the hermeneutic component, while at the right – arguments without the deductive component. (…) Rationalistic arguments are relatively closer to the left-hand side of the sequence; visionary arguments are relatively close to the right-hand side. Crucially, any philosophical argument, which pertains to a non-trivial philosophical claim, is never devoid of the hermeneutic component.[49]

He also adds that:

> in a typical situation there exists a kind of feedback between the vision and the logical argumentation. Even if the chain of arguments is inspired by a vision, rational argumentation may influence it, giving rise to its revisions and, in a critical situation – even to its rejection.[50]

[49] M. Heller, *Przeciw fundacjonizmowi*, [in:] M. Heller, *Filozofia i wszechświat*, Universitas, Kraków 2006, p. 93.
[50] *Ibid.*, p. 94.

Such a view of philosophical argumentation leads firmly to the rejection of foundationalism: if argumentation is a constant interplay of the hermeneutic vision and deduction, there exist no indefeasible, 'clear and distinct' premises, or there exists no unshakable foundation of our philosophical constructions. Argumentation in philosophy takes on a different form:

> When one begins to solve a problem, (...), one accepts certain *hypotheses* (...). It is important to note that these are hypotheses, not certainties (...), and maybe even working hypotheses. By using them one arrives at a solution of a problem (…). The results of the analysis may either strengthen one's initial hypotheses, or lead to their modifications. Such a procedure may be repeated multiple times, resulting in the self-adjustment of the system.[51]

Heller's insightful remarks may be summarized – and given more precise form – in the following way. Any philosophical argumentation must meet four conditions:

(a) the revisability condition: at least some of the premises of any philosophical argumentation are hypotheses – they can be rejected or modified;
(b) the feedback condition: the modification or rejection of premises (hypotheses) must be based on the evaluation of their logical consequences;
(c) the background stability condition: the argumentation background (some previously accepted theories other than the evaluated hypotheses) is relatively stable in relation to the hypotheses; it should be easier to modify or reject the hypotheses than the background;

[51] M. Heller, *Nauki przyrodnicze a filozofia przyrody*, [in:] M. Heller, *Filozofia i wszechświat,* Universitas, Kraków 2006, p. 32.

(d) the disputability condition: any philosophical argumentation is open to formulating competing, even contradictory, hypotheses.

Heller rightly observes that arguments that meet the above stated conditions cannot be accounted for within classical logic. He urges us therefore to look for a 'non-linear logic', or such a logic that would encapsulate the structure of non-foundational thinking.[52] Although I cannot offer such a full-blooded logic here, I would like to suggest that non-foundational arguments can be explicated with the use of some non-classical but well-known formal tools and, in particular, the formal theory of belief revision and the formal theory of coherence.

The idea is simple: with a given language *L* and the background knowledge *K* one puts forward certain hypotheses *H1, H2, H3, ...,* each aiming at solving a problem at hand. We shall say – simplifying considerably – that a problem is defined by a pair of contradictory sentences {*p, ~p*}, and that to solve a problem means to determine which of the sentences, *p* or *~p*, is true. Thus, a hypothesis *H* solves a problem when it (together with some other previously accepted sentences) implies *p* or ~p. Importantly, any newly introduced hypothesis *H* together with the background knowledge *K* may yield contradiction. In such cases, one needs to revise or reject some parts of the background knowledge, and this procedure is well modeled in formal theories of belief revision[53]. In other words, the set *K*H1*, i.e., *K* revised by *H1*, may not include every sentence, which was originally in *K* (I simplify here, disregarding the fact that there usually are many ways of revising *K* by *H1*, and so the set *K*H1* is in fact chosen from among the possible ways of modifying *K* in order to accommodate

[52] The classical relation of logical consequence is a non-linear function. In addition, there exist formal systems called nonlinear logics. However, Heller speaks of something different – a logic of epistemological non-foundationalism – and hence I used the term 'non-linear logic' in quotation marks.

[53] Cf. P. Gardensors, H. Rott, *Belief Revision*, [in:] *Handbook of Logic in Artificial Intelligence and Logic Programming*, *vol. IV: Epistemic and Temporal Logic*, eds. D.M. Gabbay, Ch. Hogger, J.A. Robinson, Oxford University Press, Oxford 1995, pp. 35–132.

H1). To put it succinctly: revisions such as *K*H1, K*H2, K*H3* often result in the modifications to the background knowledge.

Whether such modifications are acceptable depends on whether an introduced hypothesis (*H1*, *H2, H3*) indeed solves a problem that has previously remained unsolved. However, this is not the only criterion for assessing the quality of a hypothesis. The other such criterion is coherence: we shall say that the better the hypothesis (solving some problem) is, the more coherence it generates in our system of beliefs. Coherence is determined by taking into account: (a) the number of nontrivial inferential connections in our belief set (so in *K*H1, K*H2, K*H3* respectively); and (b) the degree of its unification.[54] There exist nontrivial inferential connections between sentences belonging to a given set if they can serve together as premises in logically valid schemes of inference. In turn, a given set of sentences is unified if it cannot be divided into two subsets without a substantial loss of information.

Thus, the question is, which from among the considered hypotheses *H1, H2,* and *H3* (of which all solve the problem at hand), should be given priority? The answer lies in the interplay between two factors: the extent of modifications a hypothesis causes within our background knowledge (the less changes the better), and the degree of coherence it brings about in our belief set (the higher degree the better). There is no simple formula to settle this interplay, it is rather a matter of decision on case by case basis. However, it is reasonable to assume that if two hypotheses, *H1* and *H2*, bring about a similar level of coherence, and when *H1* causes substantial modifications in the background knowledge, while *H2* changes it only slightly, it is *H2* that should be preferred. Similarly, when both hypotheses produce similar modifications in the background knowledge, but one of them brings about more coherence, it should be preferred. It must also be added that there may be situations in which *all* of the considered hy-

[54] L. Bonjour, *The Structure of Empirical Knowledge*, Harvard University Press, Cambridge, Mass. 1985.

potheses cause so substantial changes to the background knowledge that they cannot be accepted, even if they solve the problem at hand and bring about much coherence.

The situation depicted above, i.e. one which takes into account only the background knowledge and the hypotheses, is a substantial simplification. However, it may easily be extended to give a more fine-grained description of non-foundational argumentation. For instance, one can utilize the concept of presuppositions, which enables to capture two important aspects of non-foundational thinking. Firstly, one can speak of the presuppositions *P* of the background knowledge *K*; in particular, the set *P* may contain the so-called existential and lexical presuppositions. Existential presuppositions posit the existence of a certain entity or a situation (e.g., when I say that "John has a new car" it presupposes that John exists); lexical presuppositions, on the other hand, are sentences which must be true in order for some concepts to be applicable (the lexical presuppositions of the sentence "John is not a bachelor" include "John is a male"). The introduction of the set of presuppositions *P* enables one to describe a situation in which a hypothesis leads to the modification not only of some fragment of our background knowledge, but also of our existential commitments and our conceptual scheme (when it causes the rejection of an existential or a lexical presupposition, respectively).

Secondly, the utilization of the concept of presupposition enables one to account for a situation in which one determines that a given problem is ill-stated. This requires a modification in the way we understand the process of solving problems. We shall say that a hypothesis *H* solves a given problem defined by the set $\{p, \sim p\}$ if *H* (possibly together with some other sentences belonging to the background knowledge) deductively implies *p* or $\sim p$, or it deductively implies $\sim s$, where *s* is a presupposition of *p*. In the latter case – where a presupposition of *p* turns out false – one can say that the solution to the problem defined by the pair $\{p, \sim p\}$ is that the problem is ill-stated, i.e. neither *p* nor $\sim p$ can be ascribed truth-values.

The introduction of presuppositions into our formal account of non-foundational argumentation requires two additional comments. The first is that while our background knowledge should be more stable (i.e., more immune to revisions) than our hypotheses, our presuppositions should be more stable than our background knowledge. Thus, when one chooses from among a number of hypotheses of which all solve the problem at hand and bring about much coherence into one's belief set, the hypothesis should be preferred which causes fewer modifications within one's system of presuppositions. Still, it must be stressed that taking advantage of the mechanism of presuppositions requires changes in the logic underlying non-foundational reasoning.[55]

The above described procedure meets all the conditions of non-foundational argumentation. Firstly, neither the hypotheses one considers, nor one's background knowledge, are immune to revisions, and so the revision condition is fulfilled. Secondly, the quality of hypotheses hangs together with the changes they bring about in our belief system, and they are modified or rejected if the changes are unacceptable (so, the feedback condition is met). Thirdly, the background stability condition is fulfilled since although background knowledge is not immune to revisions, from among the hypotheses that solve the problem and bring about a similar level of coherence the one should be preferred that saves most of the original background knowledge. Moreover, in cases when all the hypotheses cause substantial modifications of the background knowledge, they may all be rejected. Fourthly, as the above described formal framework enables one to work simultaneously with several hypotheses, the disputability condition is met (it must be stressed, however, that this requires a special underlying logic, e.g., the so-called defeasible logic).

In the passage quoted above, Michael Heller suggests that non-foundational thinking is typical not only in philosophy, but also in

---

[55] Cf. B. van Frassen, *Presupposition, Implication and Self-Reference*, "The Journal of Philosophy" 1968, no. 65(5), pp. 136–152.

science. This becomes clearly visible when considers the structure of argumentation in neuroscience. Let us begin with an idealization: although it is commonly accepted in the philosophy of science that there exist no theory-free observations and experiments, and that out theories play important heuristic and interpretation roles in our observational and experimental activities, let us assume that there are ‘pure’ neuroscientific facts (results of observations and outcomes of experiments). What does a neuroscientific explanation of such facts consist of? I posit that there are three different criteria at work here: empirical adequacy, convergence and coherence. An empirically adequate theory must connect facts in such a way that it may serve as a means of prediction (even if not an infallible one). For instance, a theory that posits the existence of an inborn Object Tracking System, which is capable of discriminating up to 4 objects, would be empirically inadequate if it turned out that infants are capable of tracking 10 or 15 object at once.

Still, there may exist various competing theories explaining the same set of facts. For instance, as I indicated above, there is a controversy regarding how children move from using numbers 1–4, which seems to be an innate skill, to mastering arithmetic. One proposal was put forward by Piazza.[56] She observes that the Approximate Number System (ANS) may be used to represent not only large numbers, but also small ones. Moreover, ANS quite quickly becomes very precise as regards small numerosities. Given the progression in the sensitivity of ANS, in order to distinguish between 2, 3, and larger numbers a ratio of 3:4 is needed. This happens at around three years of age, and coincides with the period when children become ‘three-knowers’. In other words, Piazza believes that no interplay between OTS and ANS is needed to ‘break the number four barrier’ – the increasing precision of the ANS system is sufficient to account for this ability.

Another hypothesis which addresses this problem is defended by Spelke. She observes that “children appear to overcome the limits

[56] M. Piazza, *Neurocognitive…*, *op. cit.*, p. 275–276.

of the core number system when they begin to use number words in natural language expressions and counting".[57] Children learn the first ten counting words by the age of 2, but initially use them without the intended meaning. At three they know that 'one' means one; at four they associate '2', '3' and '4' with the corresponding numerosities. Then, there is a kind of 'jump' – children learn quite quickly the next numbers. This, according to Spelke, requires two things: (a) to understand that every word in the counting list designates a set of individuals with a unique cardinal value; and (b) to grasp the idea that each cardinal value can be constructed through progressive addition of 1.[58] How this is possible? "For most children, the language of number words and verbal counting appears to provide the critical system of symbols for combining the two core systems (i.e., ANS and OTS), and some evidence suggests that language may be necessary for this construction".[59]

Thus, we have two competing explanations of the same set of facts: that human innate skills cannot account for simple arithmetic, and that something in individual development must facilitate – or even enable – 'breaking the number 4 barrier'. Piazza believes that the increasing sensitivity of ANS is sufficient to explain how it happens, while Spelke claims that it is the development of the language skills that plays the pivotal role here. How to decide which of those hypotheses is acceptable? One of the possibilities is to use the criterion of convergence. Let us recall some other facts, referred to above. First, both children and adults in remote cultures, whose languages have no words for numbers, when dealing with numbers larger than three only recognize their equivalence approximately. Second, deaf persons living in numerate cultures but not exposed to the deaf community use a gestural system called homesign; they use fingers to communicate numbers, but only with approximate accuracy. Third,

[57] E.S. Spelke, *op. cit.*, p. 304.
[58] *Ibid.*, p. 305.
[59] *Ibid.*, p. 305.

educated adults who suffer language impairments have problems with exact, but not approximate numerical reasoning. Fourth, when doing exact (but not approximate!) tasks adults spend more time with numbers that are difficult to pronounce, even if they are presented in Arabic notation. Fifth, bilingual adults who are taught some new mathematical facts in one of their languages have difficulties in the smooth production of exact number facts in the other language.[60] All these facts support Spelke's hypothesis – but not Piazza's – because it is Spelke's claim that language is essential to acquiring arithmetic skills that is empirically adequate for a larger set of facts. In other words, Spelke's hypothesis *converges* on more experimental and observational data than Piazza's.

Another criterion that may be used to pick from among competing – and empirically adequate – hypotheses is that of coherence. Spelke's claim that language is essential in the development of arithmetic skills seems highly coherent with Lakoff's theory of embodied mind (as presented above), while Piazza's hypothesis is not. This may be seen as an argument from coherence in favour of Spelke's hypothesis. At the same time, there may be other theories – e.g., some incarnations of the modular mind paradigm – which would favor Piazza's stance. The point is that the criterion of coherence constitutes an important justification standard in neuroscientific discourse.

Thus, even in the idealized picture of neuroscientific practice we assumed, one that posits the existence of pure, theory-free facts, there are *competing explanations* of the same set of facts, and the criteria for choosing from among them include convergence and coherence. Of course, the situation becomes even more complicated when we drop our idealizing assumption and admit that our theories – and, in particular, entire paradigms, such as embodiment or the modular paradigm – provide both heuristic and interpretation frameworks for neuroscientific practice. But the conclusion remains the same: neuroscientific thinking, *is* at its core, *non-foundational*. Neuroscientific

[60] *Ibid.*, p. 307.

hypotheses – as well as background knowledge – are revisable, and the revisions are caused not only by empirical inadequacy of our theories, but also by the consequences we draw from our new hypotheses. Background knowledge in neuroscience is usually quite stable (as illustrated by the persistence of entire paradigms, such as the embodied or modular one). Finally, neuroscientific argumentation fulfils the disputability condition: one usually formulates and chooses from among a number of hypotheses explaining the given phenomenon.

The fact that both philosophy and neuroscience take advantage of non-foundational reasoning has far-reaching consequences for any description of the interactions between them. Since neither discipline provides foundations for the other, the interactions in question are never uni-directional. Generally speaking, such mutual interactions are possible because some neuroscientific theories may constitute (a part of) the background knowledge or presuppositions of philosophical conceptions, and some philosophical theses may be elements of the background knowledge, or presuppositions of neuroscientific theories. It must be stressed, however, that it is only so if both philosophical conceptions under consideration and the given neuroscientific theories take advantage of the same language, or at least the languages they use are (partially) mutually translatable.

In a previously published paper I suggested that the interactions between philosophy and neuroscience take place at four different levels: the conceptual level, the presuppositional level, the problem level and the functional level.[61] At the conceptual level, the conceptual schemes of both philosophy and neuroscience may overlap, and concepts from one discipline may migrate to the other. This can be accommodated within the framework developed in this essay. Moreover, the framework enables also *conceptual change* brought about by the hypotheses formulated in one discipline leading to the rejection of the lexical presuppositions of the other. At the presuppositional level, the theories developed in one discipline may lead to the claim

[61] Cf. B. Brożek, *Philosophy in Neuroscience*, *op. cit.*

that some sentences within the other discipline transpire to be meaningless (e.g., when a neuroscientific theory revises some existential presuppositions of a philosophical conception). Given the above described framework, philosophy and neuroscience may pertain to the same *problems*, or one may generate problems taken up by the other. It is also possible that some existing problems may become ill-stated, when one of the disciplines (e.g., neuroscience) provides us with the grounds to reject a presupposition of some (e.g. philosophical) problem. Finally, at the functional level, both disciplines may provide heuristic tools as well as justification criteria for the other – ultimately, it is not *a priori* decidable what constitutes one's background knowledge.

Of course, it is not to say that there is no difference between neuroscience and philosophy: they both have their own methods, justification criteria and typical problems, which are built-in into neuroscientific or philosophical practice. The point is not that any belief is equally good (there may be better or worse arguments both in neuroscience and philosophy), but that there is no *critical* structural difference between the two types of argumentation. In both cases one has to do with non-foundational reasoning, and this is exactly what enables mutual enrichment between philosophy and neuroscience.