

Robert Piłat, *O istocie pojęć*, [*On the Nature of Concepts*], Warszawa: Wydawnictwo IFIS PAN 2007, ISBN 978-83-7388-137-2.

A Metaphysician in the Land of Cyborgs

In thinking we do not produce thoughts, we grasp (fassen) them. For what I have called thoughts stand in the closest connection with truth.
Gottlob Frege (1918, p. 59)

Hence the notion of the concept cannot be explicated without at the same time sketching the background against which it is set; and the "correctness" of a particular notion of concept cannot be evaluated without at the same time evaluating the world view in which it plays a role.
Ray Jackendoff (1992, p. 21)

1. The Sources

There is strong evidence that any time Robert Piłat produces a book, it is destined to become a publishing event that no passionate or meditative reader can miss. The 1993 book, *Czy istnieje świadomość?* [*Does Consciousness Exist?*], focused on phenomenological and psychological investigations, and it was only in his next work, *Umysł jako model świata* [*Mind as a World Model*], published in 1999, that Piłat reformulated his interests and redirected his writing passion to cognitive science, even though phenomenology remained an important point of reference and a rich source of inspiration. The book revealed in full Piłat's writing talent, in particular his skill of counter-argumentation; every thesis in the book meets its counter thesis. Reading *Mind as a World Model* is like sitting at a legal trial or watching a court drama, where after defining the controversy, the floor is given to the parties to engage in a heated dispute before the final judgement is pronounced. I must confess that it was a very important book for me personally, which

put in order many of the not-so-clear disputes in contemporary cognitive sciences.

The next book by Piłat, *Doświadczenie i pojęcie* [*Experiencing and Concepts*] (2006), was not at all accidentally subtitled *Studies in Phenomenology and Philosophy of Mind*, as it heralded a change of scientific outlook and anticipated the coming of an epic on concepts – the book now under discussion, *O istocie pojęć* [*On the Nature of Concepts*]. It was published in 2007 by IFiS PAN, with a title that is both daring and adequate. Already in the earlier book on the existence of consciousness there were traces to be found of the next book's core idea, namely that of mind as a world model. In the last chapters of this first book Piłat revealed his key concept – his personal world model – to be later meticulously developed in the next stage of his work. Similarly, *Experiencing and Concepts* gave a foretaste of the author's strong dissatisfaction with the theory of conceptual representation and, with it, the state of contemporary cognitive sciences. When read from a certain angle, *On the Nature of Concepts* is an indirect critique of cognitive reason. It is an expression of the author's discontent with the current grounding of cognitive theories in philosophy and, in general, with the weakness of contemporary cognitive sciences. To deepen the philosophical grounding and to better explain the specific problems of cognitive sciences, Piłat makes numerous references to Kant, Husserl, Hegel, Frege, Jaspers, and even to Heidegger. On the other hand, the book begins with a reconstruction of Quine's views on the irrelevance of concepts for scientific knowledge, and Putnam's on meaning as the contingent use of expressions – but not as a complete set of features necessary to specify the reference object. This multitude of citations makes the book exceptionally opulent, but it also makes the readers ask what connects such a varied and not necessarily coherent polyphony of inspirations.

2. The Intention

What I find disturbing and discouraging in the book is the scope of research interests and the daring way in which theses are formulated. This is both its advantage and its disadvantage. In his attempt to provide a comprehensive book on concepts, Piłat examines them at all possible levels. Hence, we have an extensive

section devoted to the empirical basis of concepts (the conceptual content of experience), but we also have parts on the normative nature of concepts, categorisation, categorical mistakes, the relation between concepts and language, learning concepts, the transformation and use of concepts, and finally, on meaning and cognition (in the context of truth, reasoning, thinking, and formulating theories). The scope of the work, resulting from the need to write a complete study (almost encyclopaedic) on the nature of concepts, explains why the chapters of the book are uneven. To make things even more complicated, Piłat employs multiple methods to justify the judgements he makes. After bold and absorbing chapters on the conceptual content of experience and the empirical nature of concepts (the conceptuality of existence), the final chapters sound somewhat unconvincing and biased. Piłat is well aware of this and toward the end he comments melancholically: “My argumentation is not a uniform line of reasoning, but rather a search for evidence, going into many directions and researching philosophy as well” (p. 303). However, it should be noted that this distracted detective-like approach is accompanied by clear and focused metaphysics.

The extensiveness of Robert Piłat’s book is matched by its philosophical intensity. The author is aiming higher than just providing encyclopaedic knowledge on concepts. The book’s concern, in fact, is to seek to formulate a certain radical philosophical (ontological) thesis, and to decide upon the existence or non-existence of philosophy itself. Even though, as we read in the introduction, Piłat declares that he is not interested in synthesising knowledge on concepts, but only in showing “why such a synthesis is not possible” (p. 9) and only in “working out a sufficiently precise notion of a concept”, it seems to me that the awareness of the impossibility of creating such a unitary theory of concepts came to Piłat only at the very end of the writing process, and is a statement of the failure, rather than the success, of the project.

Piłat is not content with achieving knowledge of how concepts function in the mind, in culture, in language, etc. Instead, he attempts to answer a challenging and everlasting philosophical question about the nature of things: “What are concepts as such?” It seems that finding an answer to this question is vital to Piłat not only for cognitive reasons. Piłat considers concepts to be the actual domain of philosophical work, and a philosophical project without

concepts, or deprived of conceptual truths, to be broken and diminished. He says: “To abandon the attempt to explain the principles of the conceptual universe – the general theory of concepts – is to abandon a crucial aspiration of the human mind. This aspiration has always animated philosophy, and this is why the history of the general theory of concepts parallels the history of philosophy itself” (p. 29). And this is indeed the aspiration of Robert Piłat’s book on the nature of concepts. He seeks to restore a particular philosophical hope, a hope that was gradually abandoned in the course of the 20th Century. We speak here of the philosophy inspired by Nietzsche, Heidegger and Derrida – a philosophy free from the need to justify claims through references to, or a construction of, universalistic dictionaries. And there is also the philosophy inspired by Darwin, Dawkins and Dennett - attempting to naturalise epistemological problems and to approach them in accordance with the project of reverse engineering.

Piłat makes this point crystal clear at the very beginning of his epic. He says: “I defend the thesis that concepts are ideal objects related to the modal properties of the world. I distinguish those concepts from the conceptual representations researched by cognitive psychology and semantics” (p. 10). This thesis reappears again and again, like a mantra. The trouble is that when Piłat undertakes analytical work, he refers to only two sources of inspiration: the semantic theory of Renate Bartsch, intending to provide the extensional conditions for creating and understanding concepts, and the theory of conceptual spaces of Peter Gärdenfors, aiming to understand how conceptual representations appear and how they relate to other representations and the operations performed on them. For Bartsch, representations are functions, abstract objects; and for Gärdenfors representations are processes in real nervous systems. *Prima facie* it is not clear what the relations between these three elements are: Piłat’s strong thesis of conceptual realism (the ontological thesis); Bartsch’s thesis that conceptual content is dependent not on the properties of mental representation, but on the qualities of formal construction modelling, the use of expressions in various situations (the semantic thesis); and Gärdenfors’ thesis, in which concepts refer to regions in conceptual spaces that possess no integrated coordinates, meaning that in actual experience they do not come along together (the cognitive thesis). Neither Bartsch nor Gärdenfors need to support their claims in

conceptual semantics (Bartsch) and conceptual geometry (Gärdenfors) with the strong thesis of conceptual realism. In other words, if we are to assess the value of Robert Pilat's proposal, it is absolutely necessary to learn the line of reasoning which supports his claim that approaching concepts as abstract objects can help explain something that the semantic or cognitive theory cannot do.

Pilat asserts that he is not after a synthesis, but an understanding of the differences between the paradigms, because "right there is the knowledge that can enrich philosophy of mind, not in potential syntheses" (p. 45). Even so, it is clear that this single articulation alone does not meet his own expectations, as we soon learn that apart from the cognitive and the semantic theory of concepts, this book has one more "protagonist", "namely concepts themselves, which cannot be reduced to the cognitive form of their representation in the mind, nor to the semantic structure of sets" (p. 50). Let me confess that, to me, this third "hero" is mostly a hollow figure, and the major problem that Pilat brings on himself is connected with his inability to clearly demonstrate the gains to the two other theories were such a postulate to be accepted.

If that was not enough, the acceptance of strong conceptual realism may bring more problems than gains. If we agreed to Pilat's thesis that concepts are abstract objects, then we would also have to assume that concepts are not (1) equivalent to intentional states of mind, or (2) parts of a language or any other symbolic system. If that were true, then in speaking of a relation of concepts to the world, we would in fact be speaking of a relation of objects to other objects, not of cognitive representations to the world. In place of reference, which belongs to the representational dictionary, we would rather choose a completely different ontological relation. Pilat examines how this relation is analysed in the second volume of Hegel's *Science of Logic*, but does not take it into account in further considerations.

3. Modality

The key concept for Pilat is possibility. He says: "Concepts are possible, abstract individuals, to which our mind and sensuality remain in a relation of partial exemplification" (p. 56), and "Necessity is not a feature analytically inferable from causality.

Causality is more of a cross-section within the scope of possibility” (p. 59). Piłat also says: “The conceptual universe is not simply a set of concepts ascribed to individuals or a group, but a field of dynamic conceptual transformations. Only in such way can the conceptual universe organise whatever is possible. Whatever is possible is incidental - and incidentality, by nature, excludes static representation” (p. 34). Finally, Piłat comes to the following conclusions: “the relevance of our conceptual systems follows from the fact that it brings order to the sphere of possibility” (p. 24) and we are not “putting reality in relation to possibility as its product but the other wayaround: possibility is emerging from the real. Reasoning is a transgression beyond the real to what is possible” (p. 308). Well, the question is: can we, from that set of claims, construe a coherent proposal for approaching modal terms? Can we make out what Piłat thinks about two difficult relations – possibility/necessity and necessity/causality? And what meaning of possibility is he referring to?

In asking about the category of possibility we could – for example – be thinking of at least three mutually non-reducible ideas: (a) logical and/or conceptual possibility (when we ask whether something can exist in a given world ruled by the principles of logic); (b) metaphysical possibility (when we ask whether something can exist in a given possible metaphysical world ruled by prerequisite propositions *a posteriori*); (c) nomological possibility (when we ask if something can exist in a given possible world ruled by the laws of physics/chemistry/biology of the existing world). These three ways of apprehending possibility/necessity are still not comprehensive enough to denote all the variants. We have heard of other possibilities (See: Marciszewski, 1972; Marciszewski, 1988; and more recently: Poczobut, 2000): logical possibility (related to a given set of laws of a given system of logic, perhaps not classical: what is logically impossible in classical logic may be possible thanks to the laws of paraconsistent logic authorising contradiction), metaphysical possibility (related to the set of synthetic statements *a priori* – as in Kripke’s sense: identicalness *a posteriori* e.g. “Water is H₂O”), ontological possibility (defined by the relations between objects), empirical possibility (related to the conditions of a given temporal system), physical (related to the set of physical laws valid in the present world), conceptual possibility (related to the conceptual abilities of a given subject), and linguistic possibility (related to the generative power of a given language). And still, these are not all the meanings of possibility that are in use. We have left out the category of Aristotelian potential (potentiality, power) and the Hegelian

category of becoming, which – as we shall soon see – is most attractive to Piłat.

Piłat provides no such distinctions and therefore – in various contexts – he uses “possibility” in different senses. I suspect that behind Piłat’s way of thinking about the category of possibility there lies the Kantian objection that we cannot determine the category of determination. According to Piłat, in the modal organisation of the events and processes (causal chains) there is a function which selects the chain of determination from all possible causal chains. The explanation of events is nothing else but the modelling of this function. The choice of the modelling structure from a pool of possibilities is the subject’s reflexive way of relating to reality and cannot be described in terms of the physical interactions between the organism and the environment. This is all fine, but then we should expect Piłat to elaborate some kind of a modal logic or at least to use an already existing model. What surprised me most was that despite the continuous references to modal concepts, there is only one – if I am not mistaken – reference to the already elaborated methods of modal logic. In the case of the category of possibility/necessity, Piłat relies solely on his own semantic intuition.

Let us take a closer look at that modal intuition. It seems to me that, once again, it is inspired by phenomenology. For Husserl, constitution is the act of separating the actual from the scope of the possible. But since the scope of possibility is theoretically infinite, and since it contains negative possibilities as well, it yields serious problems that can only hope to be solved with the help of the contextual content. If there exist a multiplicity of relations connecting a cognitive organism to a set of objects, then only the contextual content may have a referential function in relation to the scope of possibility. I must assume then that Piłat follows Husserl’s intuitions in that the intentional subject manifests itself in experience as invariant, focusing on itself even the simplest conscious perception. For the time being I will have to pass over two important problems – (1) how to interpret the contextual content which focuses the perception on the invariant and (2) does focusing on the invariant in experience have a purely intuitive and non-sensual component, or is it perhaps managed by a physical mechanism? Instead, I will focus solely on the work of modal concepts.

Jaakko Hintikka used to say, “without beating about the bush”, that a concept is intentional if and only if it takes into account many possible states of things or courses of events (in his interpretation – possible worlds). To Hintikka, this means that the semantics of possible worlds is the logic of intentionality, and the intentional is what requires a reference to the semantics of possible worlds (Hintikka, 1975, 34). A reference to Hintikka is justified here insofar as he considers speaking of possible worlds to be idle talk as long as the consequences for possible experience are not defined, which in turn implies that a possible world is defined by the related entirety of experience. The trouble is that after making this assertion Hintikka does get down to building a modal logic and an ontology of possible worlds, and Piłat does not. The references to Renate Bartsch and Peter Gärdenfors, and then to the effects of metaphorisation (Lakoff, Fauconnier, and Langacker) appear insufficient and only paraphrase the problem without outlining a solution. This arbitrariness and the unclear way of dealing with the dangerous category of possibility have far reaching implications.

4. What is a Concept, that a Person can't Grasp It?

The title of this chapter is of course a paraphrase of the title of a well-known essay by Ray Jackendoff on concepts and conceptual semantics (Jackendoff, 1992). To Jackendoff, there are reasons for speaking of “E-concepts” and “I-concepts”. An “E-concept” is something given in the external world, e.g. the concept of a natural number or a cardinal gauge, that conceptual realists approach as if it existed independently of whoever knew it or comprehended it. An “I-concept”, on the other hand, is a unit in one's head that can be passed to other persons only through language or some other channel of communication. To the degree that generative grammar allows us to better understand the human ability to use language, the choice of “I-language” as the subject of study (Noam Chomsky's choice) is entirely justified. Similarly, to the degree that conceptual semantics allows us to better understand the human ability to use concepts, the choice of “I-concepts” (Jackendoff's choice) as the subject of study is fully justified as well.

The problem is that even though Piłat implicitly refers to the above distinction, he does so without showing interest in a better understanding of the human ability to use concepts, and focuses

instead on the concepts themselves and their ambiguous ontological status. Hence, the emphasis in Piłat's book falls on "E-concepts". This forces Piłat to attempt to answer a challenging question: how are humans able to reach and to grasp concepts? If humans are not so much into generating concepts – since they can only generate and learn conceptual representations, and not concepts themselves – then how can we explain the phenomenon of the productivity of concepts? Here Piłat's answer is similar to the one once given by another giant of conceptual realism, Roger Penrose¹, whom Piłat appears to dislike. Reflexive (conceptual) intuition is the cognitive power which can explain how the unlimited number of varied syntactic structures corresponds to the unlimited number of varied concepts. All concepts share some of the features that Kurt Gödel ascribed to mathematical concepts: "as constructions of intentional objects, they come before any representation of these concepts" (p. 115). According to Piłat, we can formulate abstract intuitions, and in order to name them and then to determine if they are true, we reach for the arithmetic that allows such determination to be made. "The ability to formulate abstract intuitions is what distinguishes the human mind from machines" (p. 113). This however, may bring other controversies and doubts. Is the category of reflexive intuition to be trusted so much? Can the dark be explained with the darker? The reflexive intuition is defined through references to the category

¹ Penrose – who, like Piłat, was inspired by the findings of Kurt Gödel – claims that no non-trivial mathematical theorem can be proved solely in a formal way – that is, using axioms and tautological rules of reasoning. Definitions and mathematical operations are performative, that is to say they create new mathematical states of things, which, as a result, make the non-trivial mathematical theorems equal to synthetic statements – and not to purely analytical statements. Thus, these definitions and operations are authentic findings and not just simple consequences of the axioms. This fact makes mathematics similar to the empirical sciences, disciplines authentically experiencing the world. Mathematics is no more deductive than other theories in the empirical disciplines (Penrose, 1989). Robert Piłat once pronounced that the weakness of Penrose's proposal results from a lack of philosophical awareness, and even a lack of philosophy itself, in the work of the author. This may be true - a deepened philosophical consciousness could have warned Penrose not to make so many hasty statements, such as the one suggesting human consciousness is the result of quantum gravity effects in microtubules - but it makes us ask whether a deepened philosophical consciousness in Robert Piłat's book is not a reason for his excessive trust in the instructive power of philosophy, its ability to illuminate the "darkness of existence" and to explain the whole as a "pool of possibility of events".

of possibility, but the category of possibility can only be defined in the act of conceptual intuition.

Piłat's reasoning is as follows. The form of a given concept is at the same time a disposition to a certain type of mental activity and to certain actions. If so, the disposition must take the form of the acceptance of a certain possibility – namely, the possibility of acquiring some other concept. Piłat says: “If we speak of the possibility of a concept, it cannot belong to the same level of conceptual representation as a concept already acquired. Again, we see the necessity to distinguish between concepts effectively present in the cognitive system – that is conceptual representations – and concepts as such, which make up the layer of abstract objects built on top of conceptual representations. The possibility of acquiring a concept is simultaneously the actual acceptance of a given possibility, or more precisely, a disposition to accept a given possibility or a certain modal judgement” (p. 165).

It appears that in this line of reasoning the ontological order is continuously confused with the epistemic order. Cognitive representations, which belong to the epistemic order, are concepts effectively present in the cognitive system. But there are also concepts as such, which constitute the ontological order. Conceptual intuition is the acceptance of a certain possibility, but there is also the category of the “possibility of a concept” – right at the borderline between ontology and epistemology. If concepts are abstract objects, then the “possibility of a concept” is the ontological possibility of the existence of a certain object, but if concepts are conceptual representations, then the “possibility of a concept” results from a correction of our knowledge.

To better understand the complex relations between what is epistemic and what is ontological, let us turn once again to Saul Kripke. Kripke notices that the statement: “The evening star is the morning star” is *a posteriori* – its validity was established through astronomical observation. However – Kripke claims – this statement is also necessary. “The morning star” and “the evening star” are proper names, and as such, in all possible worlds, they refer to the same object of the real world – Venus. It means that the statement “the evening star is the morning star” is true in all possible worlds, and thus, it is necessary (Kripke, 1980).

However, is it not true that this statement, like all statements based on observation, is uncertain? Does it not imply that this

particular statement could possibly be false? The possibility we are referring to here is epistemological. If we consider all we know, it is possible that one day we might discover that our observations were inaccurate. The statement “the evening star is the morning star” is equipped with a particular quality that makes it valid in all possible worlds: if it is already true, then it is true in all possible worlds. In this sense, this statement is necessary. This is what Kripke describes as metaphysical necessity.

The statement “the evening star is the morning star” is thus a statement that is metaphysically necessary. Of course, other statements of the same kind are also metaphysically necessary, e.g. “monkeys are animals”, “heat is the movement of molecules”, “light is a stream of photons”, etc. All these statements refer to the nature of certain objects – a monkey, heat, light, and so on. Let us consider the statement “heat is the movement of molecules”. It is *a posteriori* – many years of thorough research and experimenting were needed to find out that heat is the movement of molecules. As any *a posteriori* statement, it is uncertain, that is, it is possible that one day it will come out that in spite of all our efforts we were mistaken and that heat is not at all the movement of molecules. This possibility is epistemological; it describes the cognitive situation we are currently in. It is an expression of the uncertainty related to all empirical knowledge. However, this does not undermine the metaphysical necessity of the statement.

Let us assume, for the time being, that we are not mistaken and that heat is indeed the movement of molecules. In which case this sentence is not only true but also necessary. “Heat” from some other world, “heat” found in a distant region of the universe, would be considered heat (as described in our language) only if it was the movement of molecules. A phenomenon similar to heat in all respects (that is, affecting our senses identically), but which is not the movement of molecules, cannot be (and would not be) described as heat. If heat is indeed the movement of molecules, it is necessarily the movement of molecules.

In this example we can clearly see that epistemological possibility is harmonised with metaphysical necessity. The contingency of our knowledge determines the possibility of correcting our concepts. The validity of our knowledge in all possible worlds decides in favour of its necessity. According to Pilat, concepts as such, unlike conceptual representations, express

the possibilities of the world. But what are these possibilities? Well, if the extensions of concepts are sets of possible things and states of things, then the possibilities Piłat is referring to consist in extending the concepts and increasing the complexity of their content as new referential objects appear. To me, this process is exclusively epistemological, as the concepts get affected by their extensions and their growing complexity. Piłat demonstrates how the contents of certain concepts evolve, as a matter of logical or conceptual – but not metaphysical – possibility. The problem is that according to his main thesis, Piłat should in fact be referring to the idea of metaphysical possibility and questioning whether something can exist in a metaphysical possible world ruled by necessary *a posteriori* statements.

Instead, Piłat claims that “contingent truths about the world expressed in scientific theories all have at their base certain conceptual truths *a priori*, that is, truths in the relations between concepts. The expression of these relations is the possibility of proving identical extensions of two concepts with different connotations” (p. 306-307). Excellent, but the truths contained in the relations between concepts still refer to logical or conceptual possibility. To put it bluntly: if concepts are objects and not representations, then Piłat should rather be searching for evidence for the *a priori* structure of the universe, which for some reason allows or excludes certain states of things, and not for the *a priori* form of knowledge. We should then begin some work on ontology similar to that done by Wittgenstein in the *Tractatus Logico-Philosophicus*, which aims at defining the space of possibility and impossibility for the state of all things. “Each thing is, as it were, in a space of possible states of affairs. This space I can imagine empty, but I cannot imagine the thing without the space” (Wittgenstein, 1961)². Piłat however, formulates epistemological arguments after putting forward ontological theses. This is not a weakness of his book, but a proof of how difficult it is to harmonise the successes of cognitive sciences with an interesting philosophical ontology.

Piłat defends his position in yet another way. To him, there exists a conceptual universe preordaining the existence or non-existence of certain objects. This means that conceptual relations have

² In Polish literature, a similar project was undertaken by Bogusław Wolniewicz (Wolniewicz, 1985).

metaphysical consequences. “While there are concepts referring *a priori* to non-existing objects, there seem to be no *a priori* conditions of existence of certain objects” (p. 288). If so, the power of concepts would be a negative one – excluding certain states of things. This means that concepts would more likely set the conditions of the impossibility of certain objects rather than the conditions of their possibility. Here, we discover Hegel again. Piłat quotes Hegel’s saying that concepts are the other side of things, which links the representation of existence and the existence of representation, and to describe the movement of substance. If we were to follow the flow of Piłat’s thought, then we would have to accept that his favourite category of possibility is not the conceptual, logical, metaphysical, nor even nomological, but rather the Hegelian category of possibility – possibility as potentiality in movement, which reveals the truth of the concept itself. Piłat, indulging himself with the notion of concept as the other side of existence, forgets about the other aspects of Hegel’s philosophy. Let me mention just one.

In the last part of the *Phenomenology of Spirit* Hegel says: “The self-knowing Spirit knows not only itself but also the negative of itself, or its limit: to know one’s limit is to know how to sacrifice oneself. The sacrifice is the externalisation in which Spirit displays the process of its becoming Spirit in the form of free contingent happening, intuiting its pure Self as Time outside of it, and equally its Being as Space” (Hegel, 1977, p. 462). This might explain why a concept is something that a person cannot grasp. What cannot be grasped is this constant becoming – the movement consisting in the externalisation and restitution of a concept’s existence, which dies away at its limit.

Would Robert Piłat want to attempt to describe this Hegelian odyssey of mutual references to the subsequent sequences of existence in search of the absolute concept? Would he ever wish to undertake this negative work himself, the work of concepts seeking their own content in constant and intensifying acts of overcoming? “The True” – Hegel says – “is the whole. But the whole is nothing other than the essence consummating itself through its development” (Hegel, 1977, 20). I dare to doubt Robert Piłat’s Hegelianism; but I would rather let it go here and conclude with a recognition that in this book the metaphysical nature of the analyst wandering about the land of Cyborgs felt an irresistible desire to

experience, if only for a little while, the true life of conceptual subjectivity. Someone else reading *On the Nature of Concepts* could, however, doubt the cognitive orientation of its author, and come to the conclusion that his metaphysical nature has prevailed over the dead language of the analyst (engineer), manifesting itself so badly that it may as well attempt to correct the ontological argument for the existence of God.

Szymon Wróbel
Polish Academy of Sciences

References

- Frege, G. (1918). *The Thought: A Logical Investigation*. (P. T. Geach & R. H. Stoothoff, Trans.) *Beiträge zur Philosophie des deutschen Idealismus, 1*, 58-77. (Original work published in 1918).
- Hegel, G. W. F. (1977). *Phenomenology of Spirit*. (A. V. Miller, Trans.). Oxford: Clarendon Press. (Original work published in 1807).
- Hintikka, J. (1975). *The Intentions of Intentionality and Other New Models for Modalities*. Dordrecht: Reidel.
- Jackendoff, R. (1992). *Languages of the Mind. Essays on Mental Representation*. London: The MIT Press.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge: Harvard University Press.
- Marciszewski, W. (1988) Logika modalna. In *Mała encyklopedia logiki*. Wrocław: Ossolineum.
- Marciszewski, W. (1972). *Podstawy logicznej teorii przekonań*. Warszawa PWN.
- Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and The Laws of Physics*. Oxford: Oxford University Press.
- Piłat, R. (1993). *Czy istnieje świadomość?* Warszawa: IFiS PAN.
- Piłat, R. (1999). *Umysł jako osobisty model świata*. Warszawa: IFiS PAN.
- Piłat, R. (2006). *Doświadczenie i pojęcie*. Warszawa: IFiS PAN.
- Piłat, R. (2007). *O istocie pojęć*. Warszawa: IFiS PAN.
- Poczobut, R. (2000). Superweniencja. Zarys problematyki. *Filozofia Nauki*, 8(30), 25-44.
- Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. (D. Pears & B. McGuinness, Trans.), London: Routledge, (Original work published in 1922).
- Wolniewicz, B. (1985). *Ontologia sytuacji*. Warszawa PWN.

POLISH JOURNAL OF PHILOSOPHY
Vol. II, No. 2 (Fall 2008), 145-160.

Book Reviews

Martha Nussbaum, *Frontiers of Justice: Disability, Nationality, Species Membership*, Cambridge, MA: The Belknap Press of Harvard University 2006, XV + 512 pp., ISBN 9780674019171.

Martha Nussbaum shares Alexis de Tocqueville's healthy suspicion against unbridled majoritarianism, the philosophical corollary of which may be called utilitarianism. Tocqueville's fear of the "tyranny of the majority" led him to believe that "if ever the free institutions of America are destroyed, that event may be attributed to the omnipotence of the majority." Consequently, he sought to examine those "Causes Which Mitigate the Tyranny of the Majority in the United States," which is contained in Chapter 16 of Volume I of his book *Democracy in America*. Although Tocqueville believed that "the authority which they [Americans] have entrusted to members of the legal profession, and the influence that these individuals exercise in the government, are the most powerful existing security against the excesses of democracy" (I, 272), he warned of potentially dangerous abuses:

In a state of society in which the members of the legal profession cannot hold that rank in the political world which they enjoy in private life, we may be rest assured that they will be the foremost agents of revolution. But it must then be asked whether the cause that then induces them to innovate and destroy results from a permanent disposition or from an accident. It is true that lawyers mainly contributed to the overthrow of the French monarchy in 1789; but it remains to be seen whether they acted thus because they had studied the laws or because they were prohibited from making them.

Unlike Tocqueville, however, according to Nussbaum, "theories of social justice should be abstract" and cosmopolitan, that is, "responsive to the world and its most urgent problems (p. 1)." Her book is both and is therefore required reading for anyone interested in political philosophy.

In *Frontiers of Justice* ("FJ"), Nussbaum launches a revolutionary manifesto that sets out to further develop and apply the "capabilities approach" to justice. Nussbaum's approach, originally introduced in *Women and Human Development* (2000), addresses three "unsolved problems of justice": our moral duties to the mentally and physically

disabled, to members of other nations and nationalities, and to non-human animal species. Her looking at injustice naturally leads her to deplore the ways that people with impairments and disabilities, different nationalities, and animal species have been omitted from participating in political choice in the classical social contract tradition. Building on John Rawls's work, which she deeply admires and to whom *FJ* is dedicated, Nussbaum embarks on an ambitious quest to extend democratic participation on a global scale: "to include more of the human beings on the Earth's surface" (p. 21). Her filial piety toward Rawls notwithstanding, *FJ* is a highly original major critical assessment of John Rawls's contractarian theory of justice.

Chapter One focuses on and examines classical and modern social contract theories in general and John Rawls in particular. Although Rawls's theory of justice surpasses Thomas Hobbes and John Locke (pp. 35-54), "Rawls is still a member of the social contract tradition"; and his theory of justice, like all social contract theories, suffers from fundamental assumptions and errors traceable to a "very deep adherence to the contract idea" (p. 57). In Rawls's theory, people in the Original Position negotiate as independent actors of relatively equal stature behind a "veil of ignorance." "They do not know their own race, or class, or birth, or sex, or conception of the good" (p. 57). Assuming, as Rawls does, they recognize that forming an organized government will be to everyone's mutual advantage, the framers behind the "veil of ignorance" would agree to "inequalities . . . acceptable only if they raise the level of the least well-off" in society (p. 64). This is what Rawls refers to as the "difference principle."

As Nussbaum notes, however, since those in the Original Position are primarily concerned with arriving at a fair division of the gains from social cooperation, his model of justice assumes that everyone in the Original Position is rational and productive. This leads Nussbaum to reject Rawls's standpoint because the severely disabled, inhabitants of poor nations, and animals are excluded from his conception of justice. As she points out, the utilitarian model of justice, which overemphasizes collective results at the expense of disadvantaged individuals, will not guarantee social justice for those overlooked in Rawls's social contract theory either. As an alternative to these theories, Nussbaum introduces her "capabilities approach" in order to "extend the Rawlsian" conception of justice to include and guarantee basic dignity and entitlements to the disabled, foreigners, and non-human animals "that Rawls was unsure his theory could cover" (p. 81).

She offers ten essential capabilities to which everyone is entitled as a matter of right —"life," "bodily health," "bodily integrity," "senses/imagination/thought," "emotions," "practical reason,"

“affiliation,” “other species,” “play,” and “control over one’s environment” (pp. 77-78). “The capabilities approach is fully universal: the capabilities in question are held to be important for each and every citizen, in each and every nation, and each person is to be treated as an end” (p. 78).

Chapters Two and Three deal with disabilities and the social contract tradition. Nussbaum sets the tone for Chapter Two with the following quotation from David Gauthier: “The primary problem is care for the handicapped. Speaking euphemistically of enabling them to lead productive lives, when the services [of care for the handicapped] exceed any possible products, conceals an issue which, understandably, no one wants to face” (p. 96).

Members of the social contract tradition, Nussbaum points out, all presuppose that “the contracting agents who design the basic structure of society as ‘free, equal, and independent,’ the citizens whose interest they represent as ‘fully cooperating members of society over a complete life’” (p. 98), have no natural mental or physical limitations. Consequently, the social contract tradition excludes from the outset the severely disabled and overlooks the needs of those who provide care for dependents (p. 100).

In Chapter Three, Nussbaum argues for the superiority of her capabilities approach against the social contract tradition. The “capabilities approach can do better” than the social contract tradition “because it begins from a conception of the person as a social animal, whose dignity does not derive from an idealized rationality, it offers a more adequate conception of the full and equal citizenship of people with physical and mental impairments, and of those who care for them” (p. 99). The strength of her capabilities approach, she argues, lies in her rejection of the Kantian and Rawlsian conception of the person as a “rational being.” “The capabilities approach, by contrast, sees rationality and animality as thoroughly unified. Taking its cue from Aristotle’s notion of the human being as a political animal, and from Marx’s idea that the human being is a creature ‘in need of a plurality of life-activities,’” the capabilities approach “sees the rational as simply one aspect of the animal, and, at that, not the only one that is pertinent to a notion of truly human functioning” (p. 159).

The capabilities approach to justice is, therefore, a synthesis of Aristotle and Marx (pp. 277-78, 346, 352). How Aristotle’s justification of slavery and rejection of democracy fits with her “capabilities approach” to justice that guarantees rights and entitlements to the severally disabled is unclear. How Marx’s idea of protecting human dignity, which requires the abolition of private property in the name of communism, fits in with the current globalization of capitalism where she wishes to extend her model of justice is also unclear.

Chapters Four and Five concentrate on the “morally alarming” inequalities and unacceptable physical and economic conditions of inhabitants living in other nations (p. 224). She ignores the abuses of human rights and unacceptable physical and economic conditions of inhabitants living in current or former Marxist governments. Although Nussbaum argues that justice must be globalized, she rejects a world state because such a state would require the use of military force or coercion (pp. 313–14).

How, then, in the absence of a world state, are human rights and entitlements to be enforced and the “brutal and oppressive discrimination on grounds of race” and religion to be eradicated? “The allocation of responsibility among these different parts of the global structure . . . is an ethical allocation, and political only in the sense that it is aspirational and . . . the requirements at the world level are moral requirements, not captured fully in any set of coercive political structures” (p. 315). Although Nussbaum asserts the need for such rights, she does not adequately give an account of the new state that would enforce these rights. Presumably, force is unnecessary if “we acknowledge right now that we are citizens of one interdependent world, held together by mutual fellowship as well as the pursuit of mutual advantage, by compassion as well as by self-interest, by a love of human dignity in all people, even when there is nothing we have to gain from cooperating with them. Or rather, even when what we have to gain is the biggest thing of all: participation in a just and morally decent world” (p. 324).

Chapter 6 stretches the frontiers of justice to include non human animals while Chapter 7 summarizes the strengths of the capabilities approach to justice. In Chapter 6, “Beyond ‘Compassion and Humanity,’” Nussbaum warns against the romanticizing of nature as some harmonious ecosystem. Following John Stuart Mill, she characterizes nature as a ferocious, disharmonious existence of kill or be killed (p. 367). Remarkably, Nussbaum calls to overturn nature completely. Nature should not “and cannot mean just leaving nature as it is, and must involve careful normative arguments about what plausible goals must be” (p. 369). Since Nussbaum intends the capabilities approach to justice and entitlements to include nonhuman animals, care and compassion toward animals becomes a moral imperative. “Whether among domestic animals or ‘in the wild,’ human beings are obligated to intervene to prevent” abuses and “egregious harms to weaker species members” (p. 399). Zoos have figured out a workable stratagem for protecting tender animals, such as the gazelle, from predatory animals.

The Bronx Zoo, for instance, has substituted a “large ball on a rope, whose resistance and weight symbolize the gazelle” that a tiger would otherwise crunch on in the wild. Although she reports that the “tiger seems

satisfied” (p. 371), Nussbaum provides no evidence to support her claim that this technique fulfils the tiger’s fundamental instinct to chase and kill prey. Instead, she argues that animals can and should be “treated as we currently treat children and many people with mental disabilities, which have a large menu of rights and are in that sense far from being ‘mere property,’ although those rights must be exercised through human guardianship (p. 377).” Setting aside Nussbaum’s standard for the treatment of animals, rights she argued for earlier in the manuscript are ignored (p. 14). It is not clear to me how carnivorous animals are to exercise their right to life by eating other animals, which is bound up with the preservation of their species.

She acknowledges that extending the capabilities approach to justice for non-human species “in an increasingly interdependent world” (p. 18) to Christian, Jewish, and Muslim parts of the globe will be difficult. Such religions have doctrinally ranked “human species metaphysically above the other species and [given] the human secure rights to the use of animals for many purposes” (p. 390). Apparently, “increasing people’s life chances” does not include certain religious duties to the contrary (p. 18). Nevertheless, she is convinced that a consensus is “fully available across lines of tradition and religion” (p. 163).

Although several important questions go unanswered, *FJ* draws attention to the importance of thinking about justice for the severally disabled and inhabitants of impoverished regions. Her analyses of Grotius’s view of property rights along with her critique of Rawls are outstanding. Although the section on justice for animals was ultimately not persuasive, *FJ* is written in a clear manner.

Khalil M. Habib
Salve Regina University
Newport, RI

Harold Kincaid, John Dupre, Alison Wylie (Eds.), *Value-Free Science? Ideals and Illusions*, Oxford: Oxford University Press 2007, XIV + 241 pp., ISBN 9780195308969

The question whether science ought to be “value-free” can be understood in a variety of ways. Which values may play a legitimate role in science, exactly what role values play, as well as the implications values may have for scientific objectivity, have been a matter of disagreement among scientists, philosophers, historians, and sociologists. This collection of essays provides rich and sophisticated analyses of the *legitimate* role that

social, political, and ethical values (or other non-epistemic values) play in science. Although the authors disagree about what values contribute to good science and how exactly values may or may not do so, their contributions challenge the dichotomy between the view that science is objective because it is free of all non-epistemic values and the view that science is radically subjective because it is inevitably saturated with such values.

The book includes three main sections: (1) Case Studies, (2) Evidence and Values, and (3) Values and General Philosophy of Science Perspectives. Since most of the articles employ a wide range of case studies and examples from fields such as evolutionary biology, economics, sociology, archaeology, developmental biology, astronomy, biomedical sciences, and environmental sciences and they offer some analysis of values and evidence, the division of the book is somewhat misleading.

The volume's essays might be better seen as responses to what I take to be the central argument for the view that non-epistemic values, i.e., moral, political, and social values, fail to play a legitimate role in scientific reasoning. The central argument of the "value-free" view is:

1. The only concern of scientific reasoning is that scientific theories are best supported by evidence.

2. Judgments about which theories are best supported by evidence should only be determined by (a) whether the empirical evidence supporting the theory is better than any of its competitors or (b) the extent of a theory's meeting other epistemic criteria, i.e., consistency, explanatory power, and simplicity, than any of its competitors.

3. Scientific theories are purely descriptive claims about the world, and value judgments about what *ought* to be the case are irrelevant to whether something *is* the case.

4. Epistemic values are solely a function of the aims of science (They are instrumentally valuable criteria in arriving at scientific knowledge, while non-epistemic values are not).

5. Therefore, non-epistemic values can never play a legitimate role in scientific reasoning.

The first two articles, "Fact and Value" by John Dupré and "How Should Sociologists Study Social Problems?" by Michael Root, take aim at the third premise. Premise three assumes a sharp distinction between scientific "facts" and "values." A strong motivation for this assumption is that values are subjective matters of preference (5-7). If values are subjective matters of preference, then they could not count as evidence supporting a scientific claim because they are not propositional. Even if values are propositional, it is unclear how ethical and social values support the truth or likelihood of scientific claims. Susan Haack has (in)famously

argued: “That this is false is manifest as soon as you express it plainly: that propositions about what states of affairs are desirable or deplorable could never be evidence that things are, or are not the case” (Haack, 2003, 13). So, it seems that values cannot support scientific claims.

Both Dupré and Root challenge the claim that scientific theories are *always* purely descriptive claims. They argue that scientific theories may contain useful or informative value-laden concepts. Dupré shows that evolutionary psychology and economics rely on concepts with evaluative content, such as “rape,” “violence,” “economic efficiency,” and “inflation.” Root demonstrates that we presuppose evaluative assumptions when social phenomena are said to be “social problems.” When theories employ value-laden concepts, they are no longer descriptive. As a result, value judgments will be relevant and important for testing and assessing scientific hypotheses. Neither Dupré nor Root believes this to be an alarming aspect of evolutionary psychology or economics. As Dupré notes, value-laden concepts can be useful for producing scientific knowledge about “matters that concern us” (35).

Other authors in the volume show that Haack’s argument is a red herring since no one believes that a claim about what *ought* to be the case can be deduced from a claim about what *is* the case. Lynn Hankinson Nelson and Alison Wylie, for example, argue that ethical and social value judgments, taken together with other descriptive and normative assumptions, may supply evidential support for scientific claims. They show how this occurred when scientists with feminist ethical and political commitments entered archaeology in larger numbers. Sherrilyn Roush shows that social values still play an important role *after* epistemic criteria have been applied in choosing between empirically adequate theories (171). Roush acknowledges that social values cannot provide reasons for thinking a scientific theory is more likely to be true but can give pragmatic reasons for the empirical adequacy of one theory over its alternatives.

Several articles challenge premise four of the central argument. These articles use Kuhn’s arguments (1977) that non-epistemic values are necessary for interpreting traditional epistemic values. Nelson and Wylie argue that political values play a role in interpreting and weighing epistemic values, such as generality, simplicity, and explanatory power. They use a detailed case study in developmental biology to show that feminist commitments provide good reasons for different epistemic tradeoffs. The result is better scientific theories and models. Gerald Doppelt argues that a consensus about certain social values must exist to endorse or to revise epistemic values (190-193; 197-8; 207-8). Doppelt’s view does not imply that the justification of a theory is “just politics” because scientific communities have good reasons for reaching a consensus (207).

Other articles in the collection argue against the first premise of the central argument. These selections improve upon arguments originally posed by Carl Hempel (1965) and Richard Rudner (1953). According to Hempel and Rudner, scientific reasoning involves more than making judgments about which theories are best supported by evidence. Elliot Sober and Heather Douglas argue scientists must make judgments about *how much evidence is needed* for a hypothesis, a model, an interpretation of data, and so on to be acceptable. This entails that scientists consider the possibility of error. When research has social, political, or ethical consequences, scientists must consult ethical and social values in weighing how *bad* the possible consequences could be and in deciding which risks are the most acceptable. Douglas demonstrates here and in her previous work (2000) how this occurs in dioxin cancer experiments. While Sober concludes that value judgments are important, he believes they do not provide any *new* evidence (116). Douglas, on the other hand, convincingly argues that value judgments play a necessary role in determining the availability of evidence for or against a hypothesis (124-5). She shows that these value judgments commonly occur when deciding whether to risk more false positives or false negatives in the classification of rat liver slides in dioxin studies.

Like Sober and Douglas, K. Brad Wray's article challenges the first premise of the central argument. He argues that scientists must make implicit and explicit value judgments in deciding whose results to trust, and in evaluating scientific competence and expertise (88-94). In particular, he shows that sexist attitudes and assumptions can negatively affect some judgments. Interestingly, Wray does not conclude that scientists eliminate the use of values altogether. He suggests that scientists are likely to be less biased when some personal interest is at stake.

The articles in the final section of the volume examine how different perspectives in the philosophy of science account for science *not* being value-free. The most surprising view is John T. Roberts'. He argues that even logical empiricism (the sort developed by Carnap, Hempel, and Reichenbach) can be viewed as inconsistent with the view that science ought to be "value-free." Roberts shows that logical empiricism seems committed to endorsing the non-epistemic value of egalitarianism to justify the conceptual tools that it endorses (158). Roush and Harold Kincaid defend a more robust role for values in science. Roush shows how Bas van Fraassen's Constructive Empiricism legitimizes an appeal to social values in choosing between competing theories. Kincaid develops a compelling account of how contextualism can make sense of the various ways that values can be both good and bad for science. His view implies alternative notions of "objectivity."

In general, the contributing authors provide clear arguments illuminating the relationship between values and science. Although Nelson and Wylie and Kincaid reveal the implications their accounts have for the objectivity and rationality of science, more attention should have been devoted to the implications of the authors' arguments. Several unanswered questions remain. For example, if we suppose that ethical, social, and political values contribute to good science, what does this imply about the distinction between epistemic values and non-epistemic values? Does the distinction collapse altogether? Given that values *do* appear to be a part of science, what implications does this have for objectivity? Wylie and Nelson, Douglas, and Kincaid seem to think that objectivity will be determined on a case by case basis. More work is necessary to determine what will guide our analyses. The collection extends the debate about the role of values in science and moves us well beyond the old debate of whether science is value-free.

Kristen Intemann
Montana State University, Bozeman, MT

References

- Douglas, H. (2000). Inductive Risk and Values in Science. *Philosophy of Science*, 67, 559-579.
- Haack, S. (1998). *Manifesto of a Passionate Moderate*. Chicago: University of Chicago Press.
- Hempel, Carl G. (1965). Science and Human Values. In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (81-96). New York: Free Press.
- Kuhn, Thomas. 1977. Objectivity, Values, and Theory Choice. In *The Essential Tension: Selected Studies in Scientific Tradition and Change* (320-39). Chicago: University of Chicago Press.
- Rudner, Richard (1953). The Scientist Qua Scientist Makes Value Judgments. *Philosophy of Science*, 20, 1-6.

Robert Hanna, *Rationality and Logic*, Cambridge, MA: MIT Press 2006, XXII+316 pp. ISBN 9780262083492.

Logical psychologism is the view that the prescriptive laws of logic are reducible to descriptive psychological laws. This is not the only way to define logical psychologism, but it is the one most often criticized. The most famous and most vocal critics, Gottlob Frege and Edmund Husserl, have suggested their antipsychologistic arguments are decisive because no rational person would support the view that logical principles depend on how a person, culture, or species conceive of them. Since no one would be willing to claim that the laws of logic are relative, Frege and Husserl believe it is hard to find a philosopher willing to defend logical psychologism. As a result, a normative science like logic ought not to be infected by the descriptive laws of empirical psychology.

Rationality and Logic by Robert Hanna is an eloquently written and well-argued book. In it, Hanna attempts to reunite logic with psychology after a century of the philosophical community's almost unanimous disapproval of such a thesis. His intention is to rethink and reformulate the philosophical upshot of the logical psychologism controversy. Hanna believes that antipsychologism has been grossly overstated, and, if logical psychologism is defensible at all, it must be more nuanced than the view criticized by Frege and Husserl. Hanna's view, therefore, is a defense of what might be better termed a broadly cognitivist conception of logic.

Logic on Hanna's view is an a priori normative discipline constitutive of rationality, and rational animals using an innate module, called "protologic," which belongs to the logic faculty, create it. He calls his psychologistic conception of logic "logical cognitivism." The view consists of two general claims: (a) "*logic is cognitively constructed by rational animals*" (xiii) and (b) "*rational human animals are essentially logical animals*" (xviii). Together, these two claims show that logic can be revealed by our cognitive psychology.

When Hanna supports the view that rational animals possess the ability to cognitively construct a logic, he means that rational animals have an innate faculty for the purpose of representing logic, and the construction, analysis, and evaluation of any classical or nonclassical logical system (202). He calls the innate module the logic faculty. The logic faculty is a part of the cognitive architecture of rational animals, and it is the synthesis of three very familiar philosophical ideas:

- (1) the traditional idea, drawn from Kant and Boole, that logic is the a priori science of the "laws of thought";
- (2) the mid-twentieth century idea, drawn from Quine, that logic has a universal, indispensable, and unrevisable basis...;
- and (3) the contemporary idea, drawn from Chomsky's psycholinguistics and Fodor's rational psychology, that the human animal carries out all its specifically rational

cognitive activities in a fully meaningful inner language or “language of thought,” which in turn is sufficient for our cognition of natural language. (25-26)

The innate faculty or module contains a special or supernatural single universal “protologic.” The protologic is a “set of schematic logical structures, in the form of a coherent repertoire of metalogical principles and logical concepts, that is used for the construction of all logical systems, and is not itself a classical or nonclassical logical system” (153). The protologic is modeled on Chomsky’s universal grammar. Whereas Chomsky’s universal grammar postulates that humans have an innate set of constraints allowing them to acquire language, Hanna’s protologic suggests that humans possess a set of innate, a priori, universal, unrevisable, and categorically obligatory logical concepts and principles. These principles revealed by the protologic are presupposed by any act of reasoning and any classical or nonclassical logical system.

The book consists of seven chapters defending logical cognitivism, and these chapters can be divided into three parts. Although I have neither the time nor the space to review all the details of Hanna’s arguments, I would like to give a brief summary of each chapter. Then, I would like to criticize Hanna’s main argument on two fronts. First, I will argue that the foundation of his logical faculty thesis, the “protologic,” may fail because it does not rule out a proto-protologic or any successive protologics above and beyond the initial protologic and Chomsky’s universal grammar applies to language without there having to be anything similar for logic.

The first three chapters attempt to justify Hanna’s logical cognitivism on metalogical grounds. Chapter 1 addresses the problem of logical psychologism: what is the relationship between the logical and the psychological? Traditionally, two options are possible: *thick* scientific naturalism and Platonism. Thick scientific naturalism is the view that “logic logically strongly supervenes on the natural facts” (20). Thick scientific naturalism about logic must be rejected because it is logically possible that logic does not exist if the natural contingent facts that are constitutive of logic could have been otherwise. Platonism, on the other hand, fails because it “metaphysically alienates” the human mind from logic (23-24). According to Hanna, logical cognitivism can effectively avoid the pitfalls of both logical scientific naturalism and logical Platonism by not relying on contingent natural facts or by not completely disconnecting the mind from logic.

In chapter 2, Hanna addresses the “*e pluribus unum*” problem: how can we reconcile the unity of logic with the plurality of logical systems distinct from classical or elementary logic? Hanna proposes a universal protologic distinct in structure from all classical or nonclassical systems. On Hanna’s view, “every linguistically competent being constructs an internalized logic... for the representation of the “natural logic” of her own natural

language” (51). The innate protologic does not presuppose any logical system, classical or nonclassical, thus we need not give up on the unity of logic even if there is a plurality of logical systems. The protologic accounts for both the unity of logic generally and the plurality of classical and nonclassical logical systems.

The metalogical focus of chapter 3 is the logocentric predicament, which arises from the disturbing fact that the explanation of any logical theory or the justification of any deduction must presuppose and must employ logic itself. According to Hanna’s cognitivist solution, the conclusion of a presuppositional argument “partially or completely *unpacks* truth-conditional information implicitly contained in the premise” (73). In this way, logic has a legitimate presuppositional explanation in the logical faculty thesis.

The next two chapters offer an empirical psychological justification for Hanna’s conception of logic. Chapter 4 shows how human thinking conforms to the standard cognitivist model of the mind. The cognitivist model has its origins in Kant’s transcendental psychology and in Chomsky’s psycholinguistics. The cognitivist model of mind includes among its components the language of thought, and Hanna argues that the language of thought presupposes a mental logic we possess but which we sometimes do not exercise. In chapter 5, Hanna defends his protologic competence theory, and he applies this doctrine to the heated debate about human rationality. The psychology of reasoning has shown that humans are really bad at performing well on reasoning tests. Hanna argues that the possession of an innate protologic explains the nature of our reasoning competence. On the one hand, human reasoning is following and deploying the normative metalogical principles and logical concepts of the protologic. On the other hand, the highly variable performance of cognizers on reasoning tests can be explained by appealing to three facts: (1) the plurality of the language of thought suggests a plurality of mental logics or logics of thought, (2) each cognizer is limited by his own finite embodiment, and (3) the cognitive function of the logic faculty must mediate between the salient logical information of the innate logical module and the central processes of belief, desire, emotion, and decision. Given the protologic competence model, according to Hanna, we can easily explain why it is people do poorly on reasoning tasks.

The final two chapters use the cognitive conception of logic to develop an account of logical knowledge (chapter 6) and the normativity of logic (chapter 7). In chapter 6, Hanna bases his theory of logical knowledge on logical intuition in response to natural extensions of Wittgenstein’s famous worry about rule-following, and Paul Benacerraf’s worry about reconciling our face-value or standard sensitivities of mathematical truth with our best epistemology of intuitive knowledge. Intuition is a mental

act, involves semantic content, is clear and distinct, authoritative, cognitively indispensable, and fallible (171-173). That intuition is a mental act does not mean it is analogous to sense perception. In fact, on Hanna's account intuitions are not related to sense perception but to "memory, imagination, or conceptual understanding" (191-200). Hanna's antiperceptualist account denies ostensive contact between the object cognized and the cognizer. In Chapter 7, Hanna defends the normativity of logic by comparing the principles of the protologic with Kantian categorical imperatives. For Hanna, logic is a moral or "prescriptive" science because the principles and concepts of the single universal protologic must be intrinsically categorically normative for human reasoning. Consequently, we commit errors in logical reasoning because we are free to obey or not obey the logical principles.

Despite the rigor and lucidity of Robert Hanna's argument, two problems deserve to be explored. Both problems have to do with the so-called "protologic." First, Hanna's argument fails to rule out an infinite regress of protologics. Second, the analogy of the protologic with Chomsky's universal logic seems ill-conceived given the abstract nature of logical systems.

First, the protologic is a foundational concept for the logical faculty thesis. But the protologic may fail to rule out the possibility of a further protologic justifying the general structure of the protologic itself. If the protologic is not the foundation upon which we may build the logic faculty thesis, then the logic faculty thesis fails and with it Hanna's logical cognitivism fails too. Similarly, the logic faculty thesis does not rule out some sort of coherentist view of the protologic. A coherentist version of the logic faculty thesis may sacrifice the unified vision of the single universal protologic Hanna propounds. Without a foundation, logical cognitivism would fail.

Second, the analogy between Chomsky's universal grammar and Hanna's logical faculty seems ill-conceived. On page 48 Hanna emphasizes the analogy between his proposal and Chomsky's universal grammar. He writes, "My proposal is that *the protologic stands to the many and classical or elementary, extended, and deviant logics, precisely as Chomsky's [universal grammar] stands to the many native, idiolectic or dialectic, and foreign natural languages.*" But natural languages and logical languages seem to behave very differently. Natural languages, for instance, arise naturally, while logical languages seem to be artificially constructed. Also, if we have learned anything from Wittgenstein, it is that we cannot say whether natural languages are constrained by normative rules or procedures; however, logic is constrained by rules that we must apply correctly. Although the analogy is convenient and interesting, some

limitations could support the view that the logical faculty thesis fails to be anything like Chomsky's universal grammar.

Logic and Rationality is an important contribution to a debate that has been ignored for too long. Students and researchers who have thought that Frege's and Husserl's dismissal of logical psychologism was decisive should not ignore Hanna's compelling argument for logical cognitivism. Hanna's very impressive work deserves to be studied intently by those with an interest in the often criticized but largely overlooked relationship between psychology and logic.

Joseph Ulatowski
University of Wyoming, Laramie, WY

Józef Bremer, *Jak to jest być świadomym. Analityczne teorie umysłu a problem świadomości*. [What it is like to be conscious. Analytic theories of mind and the problem of consciousness] Warszawa: Wydawnictwo IFiS PAN 2005, pp. 372, ISBN 83-7388-069-0, Series: Umysł. Prace z Filozofii i Kognitywistyki.

Józef Bremer's book *What it is like to be conscious. Analytic theories of mind and the problem of consciousness* (written in Polish under the title *Jak to jest być świadomym. Analityczne teorie umysłu a problem świadomości*), published as the 5th volume in the well-known Polish series "Mind. Works from Philosophy and Cognitive Science," is an interesting overview of selected naturalistic theories of mind considered with regard to the problem of consciousness. The author also presents his own research, which aims to solve the problem of subjectivity.

He starts his discourse with an introduction to the contemporary analytic philosophy of consciousness (Chapter 1), where he defines the basic vocabulary, aim, and method of his research, and also makes some methodological and historical remarks on the problem of subjectivity. He then focuses his attention on the question of a philosophical-psychological description of the subjectivity of consciousness (Chapter 2). As a philosopher, Bremer distinguishes between the ontological and the epistemological aspects of consciousness, asking both about its nature (the ontological question) and our access to it (the epistemological question). It is obvious that this issue requires adopting some theory of consciousness. However, in his search for an appropriate theory Bremer is actually very cautious. At the beginning, he gives a model of phenomenal consciousness, but this model is metaphorical, attempting to capture some of the intuitions needed to understand phenomenal consciousness instead

of being strictly theoretical, which is more common in cognitive science. The core question focuses on the relation between subjective conscious states and brain processes, and hence on a certain version of the classical mind-body problem considered in the context of neuroscience. In Bremer's view, the solution to the mind-body problem cannot be given without also providing a solution to the problem of subjectivity. In investigating this relation, Bremer gives an overview of the different standpoints and contemporary theories (Chapter 3. The place of consciousness in chosen theories of mind), making a critical analysis of the materialistic approaches as regards their reductionism or eliminativism (especially in Chapter 4. Daniel Dennett—consciousness as activity of the brain). His critical remarks also touch upon some points of Thomas Nagel's well-known conception, in which there is a place for subjectivity, but only within a certain objective theory („objective phenomenology”, as it is called by Nagel) (Chapter 5. Thomas Nagel – subjective properties of consciousness). Only after giving these critical analyses does Bremer begin to look for a new theory of consciousness (Chapter 6). His search leads him to examine neuroscience and to attempt to join the approach of this field with that of philosophy. In his view, a satisfactory theory of consciousness needs to be developed in the context of neuroscience and needs to be multidimensional. Despite the earlier criticism, Bremer does follow Thomas Nagel's methodological approach in his own research program and determines proto-theory of consciousness in which phenomenal consciousness has its neural correlates (Chapter 7. Neural correlates of consciousness). Bremer's program consists in (a) giving definitions of the methodology of neural correlates of consciousness and (b) explaining the connections between neurobiological systems or states and conscious states.

His theory is based on two main assumptions concerning consciousness: (i) holism and (ii) emergentism. In adopting holism Bremer is making a reference to Kant's idea of the transcendental unity of apperception (i.e. self-knowledge, *Selbstbewusstsein*); with emergentism he is referring to those philosophers of mind who argue for non-reductive materialism, which has been an attractive proposal for many philosophers, especially concerning the question of mental properties, although some of them (as J. Kim) find it methodologically controversial. For Bremer, who is looking for a good explanatory theory of consciousness and is aware of its ontological assumptions, non-reductionism and antirealism are attractive. But at the same time he is well aware of the difficulties of emergentism, for he critically analyzes the different theses that are put forward in various emergentist theories and shows the dilemmas they face. Following R. Sperry, he shows that if a theory admits the causal role of mental states (with mental properties) in the „top-down” direction, then in

such a theory the causal closure of the physical domain has to be rejected. The solution to this problem is epiphenomenalism (according to which mental properties are not causally efficacious). If in turn epiphenomenalism and the „top-down” direction were rejected, then there would be no emergent properties. In the light of these and other difficulties with different versions of emergentism, Bremer chooses synchronic emergentism, which assumes nonreductive properties (in his analysis they are qualia) which are systematic (i.e. according to his definition they cannot be analyzable either micro- or macro- behaviorally) and unpredictable (that is why they are really novel). In his analysis, holism and emergentism combine the philosophical and neurological dimensions in a way that provides food for serious thought to both philosophers and cognitive scientists. Bremer's book raises a number of serious questions for discussion. By appealing to the hypothesis of correlation, the author refers to some interesting aspects of the classical mind-body problem (here mainly the question of the correlation between the subjective conscious and physical neuronal states). The unsolved question is: what kind of relation is such a correlation? How do physical brain processes (if any) generate subjective conscious states? Is our consciousness subjective by necessity? If so, what determines it: what kind of physiological process or what structures of the brain?

Bremer shows the limitations of the known scientific theories of consciousness given in neuroscience, but he also points out their philosophical assumptions and this inspires him to seek a satisfactory theory of consciousness from a wider (philosophical) perspective. Without an account of the problem of subjectivity, it is impossible to construct any adequate and satisfactory theory of consciousness. In his own work he determines emergent laws (known as the so-called bridge laws in the classical model of explanation). In Bremer's project they are theoretical constructs which require empirical confirmation. Thus an important question remains open: can this project be realized empirically? The reply to this question is the aim of cognitive scientists.

To conclude: Bremer's proposal is noteworthy in today's complex and interdisciplinary cognitive science. Written from the position of analytic philosophy of mind, it is also a very good contribution to the field, providing not only a rich conceptual network, correct definitions of many technical terms, and detailed analyses, but also a certain coherent (proto-) theory of consciousness which seems to fit very well with the results of current scientific research and evokes many interesting and serious questions both for philosophers and cognitive scientists.

Urszula M. Żegleń
Nicolaus Copernicus University in Toruń