DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 4:2 (2015)

Contents

DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 4:2 (2015), pp. 3—12
DOI: 10.1515/sh-2015-0011

Salva Phaenomenis. Phenomenological Dimension of Subjectivity in the Frame of the Reductionist Paradigm of the Cognitive Sciences

Filip Kobiela

The University School of Physical Education in Krakow, Krakow, Poland

e-mail: filipkobiela@interia.pl

Abstract:
The paper addresses the family of questions that arose from the field of interactions between phenomenology and the cognitive sciences. On the one hand, apparently partial coextensivity of research domain of phenomenology and the cognitive sciences sets the goal of their cooperation and mutual inspiration. On the other hand, there are some obstacles on the path to achieve this goal: phenomenology and the cognitive sciences have different traditions, they speak different languages, they have adopted different methodological approaches, and last but not least, their prominent exponents exhibits different styles of thinking. In order to clarify this complicated area of tensions, the paper presents the results of philosophical reflections of such topics as: 1) philosophical presuppositions and postulates of the cognitive sciences 2) abstraction of some phenomena during idealisation and the dialectical model of science's development 3) argumentation based on prediction of future development of the cognitive sciences. This finally leads to the formulation of a phenomenology-based postulate for adequate model of mind and the discussion of humanistic dimension of cognitive sciences.

Keywords: phenomenology, the cognitive sciences, philosophical presuppositions, reductionism, naturalism, idealization, methodology, Husserl, *salva phaenomenis*

1. Is cognitive science based on philosophical presuppositions?[i]

In this paper I would primarily like to refer to the first segment of this issue's guiding question: philosophical presuppositions of the cognitive science that results in formulating a certain programme assumption (postulate) which was inspired by phenomenology. As a starting point, I would like to ask a question, which is typical for the phenomenological tradition, concerning possible philosophical presuppositions of particular sciences, referring them to the cognitive sciences. I would now like to clarify a way of understanding this question in two steps: by briefly describing the peculiarity of cognitive sciences and their possible philosophical presuppositions. Thus, I will be penetrating the field of meta-cognitive-science considerations.

The concept of mind, which is a key concept in the cognitive sciences, is regarded either as referring to the existing, physically realized minds (human minds, also minds of higher animals), or physically possible minds (Artificial Intelligence), or minds possible only in terms of nomology. Modelling of cognitive processes is usually understood as the main task of these sciences. This task can be accomplished by *constructing a theory of mind* in its cognitive dimension (constructing theory that models the human mind)—in this aspect, the cognitive sciences research are continuation of cognitive psychology. This task can also be accomplished as *constructing a mind*, but what is the case is an artificial mind, commonly referred to as artificial intelligence—in this

aspect, the cognitive sciences are derived from computer science. While the very task of the cognitive sciences can be presented in a relatively homogeneous manner: constructing the models of the cognising mind by creating empirical theories or functional prototypes, within the sub-disciplines of cognitive science which are responsible for accomplishing this task, things are significantly diversified. There is no doubt that the cognitive science's project is multi- and interdisciplinary, and therefore also open for using different research methodologies. I think, however, that the belief that the unity of the cognitive sciences is emerging from such heterogeneous sources is based on the concept of supervenience [25] or, to use Ingarden's terminology–a homogenous *figural quality* emerges from different component disciplines. Irrespectively of whether the philosophical disciplines are present or not as part of (especially the philosophy of mind) the cognitive sciences project, I consider it to be a relatively independent, specific research program. The starting question is therefore as follows: does the mind modelling oriented theoretical and/or practical activity of cognitive sciences depend on philosophical presuppositions?

In the history of science, examples were recorded of important inspirations arising from the philosophy that contributed to scientific discoveries. The most famous case of this heuristic role of philosophy in science is the context of the discovery of the special theory of relativity. But still, is the already formulated special theory of relativity, as some kind of a finished outcome of scientific cognition, based on some of philosophical presuppositions? The context of „philosophical" justification for some kind of scientific theory and the possible philosophical consequences of this theory belong to the same group of issues–in both cases it is a question of the dimension of logical consequence (of course with different sets of premises and conclusions). In the first case, philosophical theses would be premises in inferences, where theses of a given scientific theory would be conclusions; in the second case, it would be the opposite way. It seems that according to a fairly widespread opinion, asymmetry occurs here: the existence of philosophical presuppositions is denied, assuming, however, the existence of philosophical consequences. In the context of the above-mentioned special theory of relativity, it is quite commonly recognised that it radically changed certain philosophical beliefs concerning, among others, the nature of time and space (therefore it has important philosophical consequences), and although in fact some philosophical premises were important to develop it, they are not included in the set of its presuppositions. Regardless of whether these beliefs are right, which will not be considered here, it becomes explicit that one should distinguish between: 1) the philosophical context of discovery 2) the context of its philosophical justification 3) philosophical consequences of discovery. In further considerations I will be dealing only with philosophical presuppositions related to the context of discovery, which contains also programme presuppositions.

Is the task set by the cognitive sciences (the cognitive task, but it could also be an engineering task) in its methodological dimension somehow dependent on philosophical postulates? I would like to regard this problem as a question of a *normative* nature, so what I am interested in is mainly the specifically understood methodological aspect of the cognitive sciences.

2. Can we quine Husserl?[ii]

When considering the family of concepts (as understood by Wittgenstein) that defines the area of the cognitive sciences research, one may notice that some of them are also part of the field of interest of phenomenology. Since phenomenological research were initiated more than half a century before cognitive science emerged, phenomenologist is, in a sense, a host in the area, to which an interdisciplinary project of the cognitive sciences is entering. Phenomenology provides description of human subjectivity, which is the basis of knowledge about what it means to be a human subject. And yet, despite this historical dependence, the popular image depicts the role of a phenomenologist as an auxiliary one, at best, for the cognitive sciences.

One would not find a separate course in phenomenology in cognitive studies syllabus. I believe that one of the reasons for making phenomenology peripheral was the opinion of some representatives of philosophy of mind, who, by mistake, identify the phenomenological method as an untrained psychological introspection [4][iii]. Given that, phenomenology would only be a kind of folk psychology, perhaps the aspiring one, but still *folk* psychology.

Another reason could be the issue of the conceptual apparatus: compatibility of the specific language of phenomenology with the scientific discourse. The basic contemporary discourse concerning reality uses the language (sometimes the trivialised one) of natural sciences; therefore the basic concepts are *gradient*, *tensor* or *a state of matter* rather than *essence*, *accident* or *a state of affairs*. Given that, as noted by A. Klawiter, the philosophers must accept (willingly or not) the fact that *lingua franca* of the contemporary knowledge about the mind is the language of science. If so, what can be the function of the phenomenology-oriented philosopher? According to Klawiter's analysis, the function of the philosopher it is to show that: „An attempt to ignore the philosopher's proposals will result in creating mind-like or even para-mental structure model rather than the model of the mind" [10][iv]. I believe this remark grasps one of the central relation between phenomenology and cognitive sciences. In the following discussion, I will formulate a postulate which allows to determine the methodological function of phenomenology in cognitive science as well as its wider context that goes beyond methodology. Now, apart from its basic methodological dimension, phenomenology also has an important anthropological dimension. For phenomenology, the human version of the mind is a model one, and its task is description and eidetic analysis of mental phenomena: awareness, intentionality and axiological dimension of the mind. Especially the last item allows to clearly articulate the typically Husserlian questions such as: *Are the cognitive sciences humanities?* [9].

3. In search for the natural analogon of mental events: phlogiston or air resistance/friction? Towards the dialectic model of science's development

So far, the attempts to indicate the place of mind in the physical world did not bring widely accepted solution that indicates the natural process which would be identified with mental phenomena. Instead of such identification, some analogies are indicated[v]. Besides, these analogies are of heuristic rather than explanatory nature. A group of eliminativistic theories also includes an analogy that compares the mind to phlogiston—the concept used in the 17th and 18th-century science to explain phenomenon of combustion. This analogy is meant to show that the mentalistic discourse is (or will become) obsolete. I'll try to offer another analogy designed to show that mentalistic discourse is indispensable not only for the daily life, but also in the theoretical context.

A field of science, which is relatively young and rapidly growing, is in a natural way future-oriented, as evidenced by the far-reaching saturation of scientific jargon with terms such as *project* or research *program*. Despite numerous objections regarding predicting the future that futurology raises in the world of science (it, after all, raises serious methodological difficulties bringing thus the risk of abusing the epistemic authority of science), it is remarkable how many arguments in the philosophy of mind and the cognitive sciences *refer* to the future, trying to anticipate development of science, although they are also often based on analogies found in the past (history) of science. The so-called *reductivism*, position which states that due to previous success of the micro-reduction, at the moment, science should prefer neuroscientific theories in relation to theories that are not based on neuroscience, is a clear example of this argument. Therefore, the postulate to abandon the phenomenology-oriented research concerning the mind would be a consequence of the radical version of this position. The structure of such arguments is usually as follows: „future science's development will explain X, and therefore one can reduce ... etc. etc." However, the presupposition included here is *de facto* questionable, therefore the structure of this reasoning should be explicated as the following: „*if* the future development of science explains X, then... etc. etc." However, a brief reaction is possible: „if", that indicates who bears the burden of proof of the presupposition. This

prediction is based on a certain extrapolation, which means it is based on an inductive inference. Due to the risk of failure of such inference with regard to development of science, which is largely unpredictable, what we are actually dealing with is a speculative position. Thus, some forms of naturalism appealing to future development of science seem to have serious hallmarks of ideology, or even of wishful thinking.

In opposition to the hypothesis of linear structure of the development of science, which is linked with reductionism, I would like to offer the hypothesis of dialectical development of science, which is also based on an analogy taken from the history of science. This analogy allows, I believe, to show that some forms of reductionism might be seen as a certain stage of the development the natural sciences. Now, in physics, one can refer to Newton's equations without taking account of the friction force or air resistance. While we know that these phenomena occur in a real system, solving the model without taking account of them might provide insights into how behaves the real system, where these forces might be neglected. Resigning from taking account of air resistance enabled formulation of the laws of motion, but the laws alone, which were obtained through this idealisation, are not sufficient for complete description of reality. A description like this becomes possible once one take account of what has been previously omitted. As noted by Weizsäcker, philosophy asks questions, whose if not asked, were a prerequisite for the scientific method to succeed. This includes therefore the statement that science owes its success to, among other things, resigning from asking certain questions [26, p. 167]. A similar idea might also be found in Husserl who claims that positive sciences in their exclusive orientation toward the acquisition of more and more results fail to reflect on their own epistemological and metaphysical presuppositions, but 'these deficiencies are part of what enables science to progress as fast as it does' [32, pp. 44, 151].

Thus, what emerges is some kind of dialectical model of the development of science: the success of a particular research program is determined by *omitting* a certain aspect of the field of phenomena examined, which has to be taken into account, however, in the subsequent studies, if it aspires to a comprehensive explanation of particular class of phenomena. To use the names of the two great scientists as figures that are symbols of this evolution: a „Galilean" model (leaving aside the air resistance), associated with the rapid development of modern science, must be complemented by „Aristotelian" taking account of the full complexity of phenomena[vi]. I maintain that there is an analogy between air resistance and mind-typical phenomena the described by phenomenology. Both of them are reality, and they are not fictions or theoretical entities (like phlogiston). Although at a certain stage of development of individual sciences they can be skipped, eventually it is necessary to take them into account. The latter stage of research can now be linked to the basic slogan of phenomenology—*Zurück zu den Sachen selbst!—Back to the Things Themselves!* If within certain limits the following analogy is valid: testing the subject without fully taking account of the specific nature of consciousness is like studying motion without taking account of friction (or of air resistance), then the phenomenology-inspired slogan of the cognitive sciences could be as follows: B*ack* (after the research that leaves aside the phenomenological dimension of consciousness) *to the Consciousness Itself*. How, however, should this return be understood?

As stated by H. Spiegelberg—the basis for the unity of phenomenology is the persistent strive to view phenomena and to be *faithful* (emphasis added, F.K.) towards them [24, 700]. Thus, phenomenology is set to oppose the various forms of phenomena deformation that originate both from the scholastic (in the pejorative sense) verbalism and to oppose a specific alienation, which is a by-product of the methodology of empirical research. Natural sciences express being in a certain aspect, using at the same time idealisation methods, therefore they provide (at least at a given stage of development) an abstract scheme rather than the complete knowledge of reality. The postulate of phenomenologically inspired research can be thus expressed as—*salva phaenomenis*—saving the phenomena.

4. *Salva phaenomenis:* phenomenological standard for project of constructing models of subjectivity

The expression ‘saving the phenomena’ goes back to the ancient Greek astronomy and had different interpretations throughout the history of science and philosophy. The phrase is connected with the work on the methodology of Aristotle, and in particular on his astronomical studies, [19], [12], but because of its methodological importance, it has gone beyond this original historical context. Contemporary usage was strongly influenced by Pierre Duhem’s discussion [6]. In Van Fraasen’s constructive empiricism ‘a theory is empirically adequate exactly if what it says about the observable things and events in the world is true—exactly if it <<saves the phenomena>>’. [28, p. 12], [29]. ‘Saving the phenomena’ consists in isomorphism between an empirical sub-model of a theory and phenomena [13, pp. 129-130].

Although the discussion concerning ‘saving the phenomena’ takes place mainly within the instrumentalist (anti-realistic) tradition of philosophy of science [3], it is necessary to confront it with the phenomenological tradition. One of Husserl's tasks was to replace the narrow empiricist concept of experience with an enlarged one, which enables a researcher to get access to wider range of phenomena. The term *phenomenon* is often interpreted as opposed to reality, as a mere *appearance*—the way in which an object appears to us, but it is not the object as it is in itself. The study of such understood phenomena is classified as something at least partly superficial; the phenomena should be transcended in order to reveal actual reality. This understanding lies behind the distinction between so called *phenomenological* thermodynamics—the study of actual phenomena with avoidance of complete microscopic explanation, in contrast with statistical thermodynamics which based its explanation on investigations of microscopic level. Now, in phenomenology *phenomenon* is understood in a different way: as the manifestation of the thing itself, the way in which objects show themselves [32, p. 55], [8, p. 84-85]. Regarding the subject matter of the cognitive science, one might say that phenomenological approach takes mental phenomena seriously: how the mind appears is an integral part of what it really is. Thus, any serious model of the mind can not dismiss such mind-specific phenomena, as intentionality [21] or the subjective, phenomenal aspect of consciousness, to use the most prominent examples.

After this introduction, we are in the position to formulate the version of ‘saving the phenomena’ principle for the cognitive science. Owen himself used the Greek formula *tithenai ta phaenomena*, but due to the analogy with the *salva veritate*[vii] formula, it will be more handily to use its Latin version. Additionally, the Latin version is meant to emphasise that the formula differs from the slogan used in anti-realistic tradition in the philosophy of science.
Here is the formulation of the principle of *salva phaenomenis* that I suggest as referred to the cognitive sciences:

> The model M* of the mind M is an adequate model if and only if its design saves all the M-specific phenomena.

The criterion of adequacy of the model of the mind must be based, first, on the appropriate selection of phenomena that are to be modelled. An arbitrary selection or omitting certain phenomena is excluded in adequate model. Secondly, an adequate model must contain the design of equivalents of all the phenomena explained—if the model suggested is to be a reliable explanation of *explanandum*, it has to be complete (it has to *explain*, not *explain away*). Of course, from the point of view of the finished construction, it is probably of secondary importance whether some occurrences were disregarded as early as when selecting *explanans* or whether they have been omitted later, during the process of explanation. Any form of ignoring (failure to save or leaving out) a class of phenomena will make the constructed model inadequate. The phenomenologist could serve here as a kind of a supervisor, being an ‘advocate’ of these aspects of subjectivity which are

best revealed by phenomenological investigation[viii]. He refers to the obligation to protect certain classes of phenomena against being ignored when selecting the phenomena that are to be explained and modelled.

Although the principle discussed is of fundamental nature, however, in the context of the history of cognitive science research it is evident that it was not always observed. The eliminativistic theory and the behaviourist theory are the most typical examples of how it was violated. As noted by Auyang: 'Behaviorists dismiss first-person experiences in the name of <<science>>, thus manifesting dogmatism opposing to science'. In contrast to such a dogmatism, 'science uses whatever concepts necessary to save the phenomena. If to characterize some human phenomena scientists require the concept of subjectivity or first-person perspective, let it be' [2, p. 134]. One of the causes of this state of affairs is that the locus of cognitive science—infrastructural level of mental faculties—is located 'below' the consciousness. The following comments by Searle accurately represent this phenomenon, at least at a certain stage of cognitive science development: 'Neither the study of the brain as such nor the study of consciousness as such is of much interest and importance to cognitive science. (...) The cognitive mechanisms we study are indeed implemented in the brain, and some of them find a surface expression in the consciousness, but our interest is in the intermediate level where the actual cognitive processes are inaccessible to consciousness. (…) The processes which explain cognition are unconscious not only in fact, but in principle. For example, Chomsky's rules of universal grammar (1986), or Marr's rules of vision (1982), or Fodor's language of thought (1975) are not the sort of phenomena that could become conscious'. [23, pp. 197-198][ix]. However, the fact that the essence of cognitive science research tends to explain the mind by showing certain unconscious structures of the mind is not, of course, a sufficient reason to ignore certain phenomena associated with the mind. Thus, the principles *salva phaenomenis* must then be formulated and what must be presented in its light is both certain regularities of how the cognitive sciences develop, and certain peculiarities of them that result from how certain examined properties of the mind are co-determined by phenomenal data.

Let us also note that *salva phaenomenis* is a more fundamental principle than *the Occam's razor principle*. The latter, although it prefers certain types of explanation, at the starting point does not define (and therefore it does not limit) the range of phenomena that are subject to the rigor of explanation. A the same time, the *salva phaenomenis* blade is aimed, so to speak, in the opposite direction than the Occam's razor blade: against the excessive reduction (or elimination) of entities that are subject to explanation.

The set of issues in cognitive sciences, where researchers can use phenomenological inspiration, is very extensive. Let us now consider some issues that may be interpreted as relevant for the principle discussed.

A. Klawiter, as an embodiment of his abovementioned postulate, suggests a conceptualisation rooted in Heidegger's distinction between a thing and a tool. The concept of a tool, initially created by Husserl and developed by Heidegger, provides a starting point to correct the standard model of perception. Well, according to Heidegger, the features of a tool can not be derived from the features of a thing. The impulse to explain the mechanism of creating the useful („handy") visual representations of qualities of objects, which deviates from the standard (Marr's and Biederman's) models of perception, was inspired by philosophy, or more specifically, phenomenology. The 'Perceiving Mind' Model, which does not take into account the above insight provided by phenomenology may not be an adequate model.

A. Miah delivers an analysis concerning the adequacy of the game of chess as a test in the context of AI. Miah argues (based on the philosophical theory of game playing delivered by B. Suits) that chess is a bad model for testing intelligence because it does not provide the opportunity for certain phenomena, which are typical for human intelligence, to appear, including „opportunity for nonlinear decision-making and deviance form preconceived strategic pattern", [17, p. 22]. I believe that Miah is right when he argues that chess programs are not adequate models of 'the mind of a chess player', and treating them as a paradigm of artificial intelligence turned out to be missing

the target as they failed to take account of many important phenomena. Better opportunities for testing humanlike intelligence could rather be provided by other kinds of games, like sport games, due to their temporal element and essential function of the player's bodies (compare e.g. Merleau-Ponty's phenomenological description of football player experience [18, pp. 168-169] as well as investigations concerning embodiment [7, pp. 129-150].

In his Cartesian-type 'empirical dualism', U. Meixner argues that materialistic account of human persons can not save certain empirical phenomena. For example: Meixner claims, that I am to be found at the location from which I am looking at the world (at the origin of my perspective on things, my eye-point), and no material object at that location is, plausibly, me because this location does not correspond to my body or any part of it [15, pp. 419-425]. An attempt to answer the question 'where am I?' along with other examples are intended to undermine materialism not by using a priori reasoning, which is traditional in philosophy, but rather using the phenomenological (in a certain sense) analysis of some experiences.

An important attempt to link phenomenological studies with neurosciences program is neurophenomenology [7, pp. 33-38]. Achievements obtained under this program can be interpreted as a successful application of the *salva phenomenis* principle. It turns out that inspiring to some empirical research may be a heuristic function of phenomenological studies [30].

The meaning of thought experiments such as J. Searle's *Chinese room* or F. Jackson's *Mary's room* can now, in the light of the *salva phaenomenis* principle, be interpreted as indication of inadequacy of a certain type of the mind models, the inadequacy that results from disregarding certain phenomenal qualities. On the other hand, a radical position called *misterianism* [16, p. 349-366] may be understood as a thesis on the impossibility to construct an adequate model. The *salva phaenomenis* principle is neutral towards the question of whether construction of adequate model of the mind is possible, since it is an implication: if an M model does not meet the *salva phaenomenis* standard, then this is not an adequate model of the mind. It is sufficient, however, to indicate the limitations of certain existing models of the mind, as well as in indicating the direction of research of more adequate models.

5. Closing remarks

While examining the problem of the relation of the cognitive sciences and phenomenology, I assumed that cognitive sciences' homogeneity is based on supervenience on many sub-disciplines, and that their primary task is to construct models of the mind. I tried to show that phenomenology understood as a study of mind-related phenomena can be an important factor in the cognitive science project. Now, any construction of scientific models, including the models of the mind, in its higher stage of development should go back to phenomena that were omitted in the first, rapid period of their development. The same mechanism of development of science through abstraction and idealization, which contributes to the initial advancement of research, in the longer term, is responsible for simplifying the initial model. To avoid such simplification I postulate to apply the *salva phaenomenis* principle that defines the overall condition of the adequacy of models of the mind. I believe that this principle does not reflect so much the idea of *naturalisation of phenomenology* [22], [31], as, on the contrary, *phenomenologisation of naturalism,* or, more precisely, phenomenologisation of the cognitive sciences.

Therefore, I believe that phenomenology has an important role to fulfil as a partner of the cognitive sciences, indicating the primacy of phenomena, whose theoretical importance—together with the increasing perfection of the models of the mind—will probably grow. In this indirect way, due to phenomenology, an important humanistic sense—the human dimension of experience—may become an inherent part of the cognitive sciences project, which is essentially naturalistic. This would mean a kind of Husserl's triumph in how farsighted he was. The *salva phaenomenis* principle belongs to the set of 'humanistic requirements' (as described by H. Arendt; the others are simplicity, beauty and harmony) of the scientific theory. 'A theory was still supposed to be

<<satisfactory>>, namely, satisfactory to human reason in that it served to <<save the phenomena>>, to explain all observed facts' [1, p. 47]. If the cognitive sciences in their teleological dimension are oriented towards constructing as faithful models of the mind as possible, they must rely on the principle *salva phenomenis*. The extent of how much they will be able to meet this challenge will also indicate the extent of how sensible it is to maintain not just that the models of the mind they offer are adequate, but also whether they might be considered as a part of the humanities.

Acknowledgement:

This article was written with the support of the Polish National Science Centre Fund (2012/07/D/HS1/01099).

References:

1. Arendt, H. The Conquest of Space and the Stature of Man. *The New Atlantis*. Number 18, Fall 2007, 43-55.
2. Auyang, S.Y. *Mind in Everyday Life and Cognitive Science*. Cambridge, MA, MIT Press, 2001.
3. Bogen, J. 'Saving the phenomena' and saving the phenomena. *Synthese* 182 (1), 2011, 7-22.
4. Dennett, D. *Consciousness explained*. Penguin, London, 1991.
5. Drummond, J.J. Phenomenology: Neither auto- nor hetero- be, *Phenomenology and the Cognitive Science* 6 (1-2), 2007, 57-74.
6. Duhem, P. *To save the phenomena*. Dolan, E. and Maschler C. (transl.), Chicago and London, University of Chicago Press, 1969.
7. Gallagher, S. and Zahavi, D. *The Phenomenological Mind*. London, Routledge, 2012.
8. Husserl, E. *Ideas Pertaining to a Pure Phenomenology and to Phenomenological Research: First Book: General Introduction to a Pure Phenomenology*. Kersten, F. (tranls.), Dordrecht, Springer, 1982.
9. Husserl, E. *Philosophy and the Crisis of European Man*. [in:] Husserl E., *Phenomenology and the Crisis of Philosophy*. Laurer, Q. (transl.), Harper & Row/Torchbook, New York 1965, 49-192.
10. Klawiter, A. Thinking like a philosopher, arguing like a cognitive scientist. *Diametros* 3, 2005, 176-181. (in polish)
11. Klawiter, A. Why did Husserl not become the Galileo of the science of consciousness?, *Poznań Studies in the Philosophy of the Sciences and the Humanities* 82, 2004, 253-271.
12. Kosman, A. *Saving the Phenomena. Realism and Instrumentalism in Aristotle's Theory of Science*. [in:] Kosman, A. *Virtues of Thought. Essays on Plato and Aristotle*. Harvard University Press, 2014, 138-156.
13. Liu, C. *A Study of Model and Representation Based on Duhemian Thesis*. [in:] Magnani, L. and Li, P. (eds.) *Philosophy and Cognitive Science: Western & Eastern Studies*. Berlin Heidelberg, Springer-Verlag, 2012, 115-142.
14. Marbach, E. No heterophenomenology without autophenomenology: Variations on a theme of mine. *Phenomenology and Cognitive Science* 6, 2007, 75-87.
15. Meixner, U. *Materialism Does not Save the Phenomena – and the Alternative Which Does*. [in:] Koons, R.C. and Bealer, G. (eds.), *The Waning of Materialism*. Oxford/New York, Oxford University Press, 2010, 417-437.
16. McGinn, C. Can we Solve the Mind Body Problem? *Mind* 98 (1989), nr 891, 349-366.
17. Miah, A. *The Deep Blue Grasshopper: Playing Games with Artificial Intelligence*. [in:] Hale, B. (ed.), *Philosophy Looks at Chess*. Chicago, Open Court, 2008, 13-23.
18. Merleau-Ponty, M. *The Structure of Experience*. Fischer, A.L. (transl.), Beacon Press, Boston, 1963.

19. Owen, G.E.L. *Tithenai ta Phenomena*. [in:] Mansion, S. (ed.), *Aristote et les problemes de la methode*. Publications Universitaries, Lovain, 1961, 113-126.
20. Rinofner-Kriedl, S. *What is Wrong with Naturalizing Epistemology? A Phenomenologist's Reply*. [in:] Feist, R. (ed.) *Husserl and the Sciences. Selected Perspectives*. University of Ottawa Press, Ottawa, 2010, 41-68.
21. Roy, J.M. *Saving the intentional phenomena: Intentionality, Representation, and Symbol*, [in:] Petitot, J., Varela, F.J., Pachoud, B. and Roy, J.M. (eds.) *Naturalizing Phenomenology*. Stanford University Press, 1999, 111-147.
22. Roy, J.M., Petitot, J., Pachoud, B. and Varela, F.J. *Beyond the Gap: An Introduction to Naturalizing Phenomenology*, [in:] Petitot, J., Varela, F.J., Pachoud, B. J. and Roy, M. (eds.) *Naturalizing Phenomenology*. Stanford University Press, 1999, 1-83.
23. Searle, J. *The Rediscovery of the Mind*. MIT Press, Cambridge, MA, 1992.
24. Spiegelberg, H. *The Phenomenological Movement. A Historical Introduction.* The Hague, 1965, vol. II.
25. Stalnaker, R. *Varieties of Supervenience*. [in:] Stalnaker, R. *Ways a World Might Be*, Oxford University Press, Oxford, 2003, 86-108.
26. Weizsäcker, C.F.v. *Deutlichkeit*. Hanser, München, 1978.
27. Wilson, R.A. and Keil, F.C. (ed.), *The MIT Encyclopedia of the cognitive sciences*, MIT Press, Cambridge, MA 1999.
28. van Fraassen, B. *The Scientific Image*. Oxford University Press, Oxford 1980.
29. van Fraassen, B. To save the Phenomena. *Journal of Philosophy*, 73 (18), 1976, 623-632.
30. Varela, F. Neurophenomenology: A methodological remedy for the hard problem. *Journal of Consciousness Studies* 3/4, 1996, 330-49.
31. Zahavi, D. Phenomenology and the Project of Naturalization. *Phenomenology and the cognitive sciences* 3/4, 2004, 331-347.
32. Zahavi, D. *Husserl's Phenomenology*. Stanford University Press, Stanford 2003.
33. Zahavi, D. Killing the straw man: Dennett and Phenomenology. *Phenomenology and the Cognitive Science* 6 (1-2), 2007, 21-43.

Notes:

---

[i] The author wishes to thank the anonymous reviewers for their valuable comments that contributed to improving the text, as well as Mr. Szymon Szymczak, M.A., for stimulating discussions regarding the relationship between phenomenology and cognitive sciences.

[ii] I'm using this verb in the sense of denying existence of entities that cannot be individuated or identified (following ontological methodology of W. V. Quine). As regards comparison of the naturalised epistemology by Quine with anti-naturalistic epistemology by Husserl cf. [20].

[iii] As regards discussion of Dennett's views on phenomenology, including the issue of relation between auto and hetero-phenomenology, cf. [5], [14], and [33].

[iv] Of course, apart from these fundamental functions, there still are secondary ones–the traditional philosopher's task: being familiar with the philosophical tradition, using formal techniques of argumentation etc.

[v] Although analogies, metaphors etc., can not replace arguments, their heuristic function is crucial. E.g. Dennett sees his contribution to the study of consciousness as a replacing one family of metaphors and images with another: 'It's just a war of metaphors, you say—but metaphors are not <<just>> metaphors; metaphors are the tools of thought. No one can think about consciousness without them, so it is important to equip yourself with the best set of tools available', [4, p. 455].

[vi] As regards presenting differences between Galileo's method of idealization in his studies of nature and the method of examining consciousness by Husserl cf. [11].

[vii] The *salva veritate* (or intersubstitutivity) formula originated by Leibniz states that two expressions are said to be synonymous if the substitution of one for the other does not change the truth value or meaning of any context in which either expression appears.

[viii] Although the interpretation of phenomenology I suggest here, is primarily based on the ideas of Husserl, it also admits (see examples given below that illustrate how the *salva phaenomenis* principle operates) subsequent modifications of the phenomenological method.

[ix] The above quotation is a Searle's statement, in which (in order to raise controversies) he speaks as the representative of the cognitive sciences.

DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 4:2 (2015), pp. 13—15
DOI: 10.1515/sh-2015-0012

On Perceiving Tropes

Przemyslaw Spryszak

Jagiellonian University, Krakow, Poland

e-mail: p.spryszak@iphils.uj.edu.pl

Abstract:

In this short paper I consider Professor Bence Nanay's suggestion that representationalism can be supported by the theory of tropes. I argue that from a philosophical point of view such a support is nevertheless not very strong.

Keywords: Tropes, Representationalism, Perception.

Any scholar interested in the philosophy of perception will enjoy reading Professor Bence Nanay's fairly recent article on a supportive link between the theory of tropes and representationalism [3]. As far as I am concerned, I like the idea of such a connection very much and find Nanay's constructive approach valuable. I must admit, however, that I am not convinced that the alleged link is in fact as supportive as it is thought to be. I hope, therefore, that the reader will allow me to briefly comment on this issue and will not find my observations completely off the mark.

On page 10, under the heading "The Particularity of Perception", the author says:

> Contrast the following two scenarios. In the first one, I am looking at a pillow, x, and unbeknownst to me, it gets replaced with another, indistinguishable pillow, y. I do not notice that the two pillows are different: I have no idea that the first pillow, x, was replaced by the second pillow, y. In the second scenario, I am looking at a pillow x all along – it does not get replaced with a different pillow. The relationalist points out that the representationalist needs to have a story about how to distinguish between these two experiences.

In what follows Professor Nanay argues that if we suppose that our experiences represent, *inter alia*, abstract particulars (tropes), the problem for representationalism regarding the answer to that question will virtually cease to exist. Although I will present his point of view below, let me first slightly alter the initial thought experiment, that is, place a similar one under the above heading. For the sake of clarity, I will present my idea step by step.

1. Suppose that I perceive a property P of an object X and that this property is a trope, and that object X itself is also a particular. Let me assume in addition, that object X is unknown to me and that my experience is not enriched by previous experiences, thoughts, etc., and that the only property of X my sense-experience represents is P.

2. I think that is fair to say that in such a case my experience tells me merely that "something is P".

3. Let us suppose subsequently, that object X was replaced by a particular object Y by destroying and immediately duplicating every part of X, with the exception of that part of X which is causally responsible for that object having the property P. One can also assume that I have no knowledge about that change and that the new object Y is similar to object X in every respect relevant to my previous perception of X, and that my attitude to perception is as before, that is, as in 1.

4. I think it is fair to say that the object Y has the property P (even if we agreed that property P is a trope), and that when I later perceive object Y, my experience tells me merely that "something is P".

5. Now the question arises: when I perceive object Y at present, does my experience tell me wrongly that X is P or rightly that X is P? I do not think the question can be justifiably answered either way, even if we maintain that X, Y and P are all particulars and that my sense-experience can represent them. My perception is in a way undecided, it does not have one definite object, *the* object, much less a particular one.

Naturally, I am well aware that Professor Nanay might respond immediately by citing the rest of part 4.1 of his initial response regarding the peculiarity of perception and repeating that what is at issue here is that

> If we accept that the properties we perceptually represent are tropes, then we have an easy answer to this question. I attribute the very same tropes to the pillow in the two perceptual experiences. Importantly, I attribute the trope of being the very same particular token object as the one I saw a moment ago (as object permanence is supposed to be perceptually represented, see Spelke 1990, 1994), (p. 10),

and that I have missed that important factor, which obviously allows us to conclude that the peculiar experience mentioned in 4. would simply tell me (wrongly) that X is P, particularly if there were no time gap between the two experiences, as in the case that Professor Nanay considers. Alas, I would respond that even if it is an empirical fact that there is a sort of object permanence representation of a token object, as Professor Elisabeth Spelke argues, that fact seems to be of little philosophical significance [4, 5]. What is still missing, in my opinion, is a proof that such a representation is necessarily contained in every experiential representation of any particular. Unless we are provided with such a proof, we are entitled to say only (at best) that representationalism is somewhat strengthened by a theory of tropes in the case where the *standard* conditions of our sense-perception obtain.

If the reader does not find my remarks very compelling, I ask him to consider the following example, which is a short story about a possible world. Suppose that I am overwhelmingly (and rightly) convinced that I live in an unstable or dynamic world where there is probability of 0.5 that the particular object I perceive was already perceived by me a moment ago, and of 0.5 that it is numerically different, that is, that it was not perceived by me a moment ago and was not replaced by the object perceived. Let us assume that in this world I always perceive objects as similar to each other, like the pillows in the above example. As a result, when I perceive an object X, my experience tells me that that the object is F (where the property F is a trope); one can arguably suppose that the trope of "being possible to be the same object as the one perceived before" is also represented. If the trope of "being possible to be the same object as the object perceived before" is preserved in the unusual (though possible) way outlined in 3-4, then there is no way of distinguishing between my experience of one persisting object and of two similar (and successive) objects. Accordingly, even if we say that objects (properties included) are indeed particulars and accept the general idea of representationalism, nevertheless as philosophers and perhaps on a less empirical level, we still face the problem of the peculiarity of perception.

References

1. Campbell, K., *Abstract Particulars (Philosophical Theory)*, Basil Blackwell, Oxford 1990.

2. Nanay, B. *Between Perception and Action*. Oxford University Press, Oxford 2013.

3. Nanay, B., Perceiving tropes. *Erkenntnis*, vol. 77, no. 1, 2012, p. 1-14.

4. Spelke, E., Principles of object perception, *Cognitive Science*, vol. 14, 1990, pp. 29-56,

5. Spelke, E., Initial knowledge: Six suggestions, *Cognition*, vol. 50, 1994, pp. 431-445.

DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 4:2 (2015), pp. 16—25
DOI: 10.1515/sh-2015-0013

Language as a Tool. An Insight From Cognitive Science

Bartosz Brożek

Jagiellonian University; Copernicus Center for Interdisciplinary Studies, Krakow, Poland

e-mail: bbroze@yahoo.com

Mateusz Hohol

Pontifical University of John Paul II; Copernicus Center for Interdisciplinary Studies, Krakow, Poland

e-mail: mateuszhohol@gmail.com

Abstract:
In this paper it has been argued that the theory of conceptual maps developed recently by Paul M. Churchland provides support for Wittgenstein's claim that language is a tool for acting in the world. The role of language is to coordinate and shape the conceptual maps of the members of the given language community, reducing the cross-individual cognitive idiosyncrasies and paving the way for joint cognitive enterprises. Moreover, Churchland's theory also explains our tendency to speak of language as consisting of concepts which correspond to things we encounter in the world. The puzzle of common sense reference is no longer a puzzle: while at the fundamental level language remains a tool for orchestrating conceptual maps, the fact that the maps encode some communally shared categorization of experience fuels our talk of concepts capturing the essences of things, natural kinds, prototypes, etc.

Keywords: cognitive science, conceptual maps, language processing, meaning, philosophy of language.

1. Prelude

In *Philosophical Investigations* Ludwig Wittgenstein famously argues that language is not a "mirror of reality"; rather, it is a tool for acting in the world. What this claim amounts to, is a revolutionary change in the understanding of the nature of linguistic phenomena. Wittgenstein underlines this point by contrasting his novel conception with the Augustinian picture of language, in which

> the individual words in language name objects—sentences are combinations of such names.—In this picture of language we find the roots of the following idea: Every word has a meaning. This meaning is correlated with the word. It is the object for which the word stands [18, §1].

What does it mean to reject this view? It is commonly believed that in the *Investigations* Wittgenstein posits that the meaning of an expression is determined by the ways in which it can be used (let us mark this fact by referring to the Wittgenstein's understanding of meaning with an asterisk '*' and use no asterisk with the "received" understanding of the term). Arguably, it is whole sentences, rather than individual words, that have meaning*, since it is sentences that are tools for acting in the world (even if we sometimes use single words – for example, when shouting "Brick!" – what we intend to say by this is something like "Watch out, because there is a brick falling down on you from the roof of the building by which you are standing!"). If we say that the word "brick" has a meaning, then it is only derivative in character: the meaning of a word can be abstracted from the meanings* of the utterances in which the word occurs. Let us consider another example. The sentence "Jane likes mangoes" means* something, but *not* because the words "Jane," "like," and "mangoes" have particular meanings. It is sentences – or their functional equivalents – that serve as the means for acting in the world and it is sentences that have meaning*, with the meanings of the component words being just *ex post* theoretical reconstructions. We can "construe" the meanings of "Jane," "like," and "mangoes" because we know in what sentences these words can occur and in what contexts those sentences can be used. This constitutes a break in the long philosophical tradition which posited the primacy of a concept over a sentence: in Plato, Aristotle, Descartes, Kant, and even in Frege, Russell and the early Wittgenstein (though with certain reservations in the latter cases), it is the meaning of individual concepts that is primary, while the meaning of a sentence is composed of the meanings of the words contained in it. Meaning* is not meaning: the latter is primarily predicated of concepts, while the former – of sentences. The latter is connected to the idea that words in language somehow reflect the nature of the things in the world they name, why the former sees linguistic expressions as ways to act in that world.

However, this dramatic shift in the understanding of the nature of language leads to some serious theoretical difficulties. Let us consider two of them. Firstly, it seems plainly clear that at least some of the words in our languages express concepts which correspond to things in the world. The word "table" expresses a concept which refers to tables, the word "apple" expresses a concept which refers to apples, and the word "water" expresses a concept that refers to water. Of course, the exact nature of such concepts is a focus of fierce philosophical debates. For example, the classical theory of concepts holds that "a lexical concept C has definitional structure in that it is composed of simpler concepts that express necessary and sufficient conditions for falling under C" [12]. According to the prototype theory, on the other hand, "a lexical concept C doesn't have definitional structure but has probabilistic structure in that something falls under C just in case it satisfies a sufficient number of properties encoded by C's constituents" [12]. The point is, however, that on both competing accounts linguistic concepts correspond to some things, which seems incompatible with the view suggested by Wittgenstein that words are mere tools for acting in the world.

Secondly, it is also undeniable that the theories we construct and express in language often capture the objective features of reality. This is particularly visible in physics, where mathematical equations not only enable to precisely predict a course of physical events, but often "know more" than their creators. An 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 solutions 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 [17, p. 10].

This, according to Wigner, 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" [17, p. 7]. But if so, it is difficult to conclude that the language of mathematics is just a tool for acting in the world (*pace* mathematical instrumentalism). Mathematical equations of the contemporary physics seem to correspond – in one way or another – to the structural aspects of reality.

Therefore, the question is, how to reconcile the persuasive picture of the nature of language as a tool provided by Wittgenstein with two different facts: that we are forced to claim that at least some of our linguistic concepts have referents in the world we encounter in our everyday experience; and that the theories developed in physics capture such aspects of the structure of the universe which lie beyond that experience. Let us deem those problems, respectively, the puzzle of the common sense reference and the puzzle of the scientific reference. In what follows, we would like to show how the first of those puzzles may be solved – or, at least, illuminated – by adopting the view of human cognition proposed by Paul M. Churchland. Further, we shall argue that although the second puzzle remains a mystery, its significance transcends the framework of the Wittgensteinian view of language as a tool.

2. The received views of cognition

In order to grasp the significance and the novelty of Paul M. Churchland's theory of cognition, it is reasonable to begin with an outline of the two traditional views of human cognitive abilities – of thinking as seeing and of thinking as conversation. The model of thinking as seeing is best encapsulated in the conceptions championed by Locke, Hume, Berkeley, and other early modern philosophers. It is a view that all we know comes from experience: when one perceives something, a mental picture (referred to as an idea or an impression) is produced, which is stored in one's memory, from which they can be retrieved even if the object which caused the experience is no longer present. Those pictures may also be processed in a number of ways. For instance, Locke believed that simple ideas may be combined into compounds, or brought together so as to establish a relation between them, or abstracted from to produce a general idea [11]. On this view, understanding something boils down to "seeing clearly" the ideas representing it, their compounds, relations between them, or some features they share. It further presupposes a special relationship between ideas and things in the external world: the former *resemble* the latter, just as pictures resemble the pieces of reality they portray. Still, Locke, Hume and Berkeley play down the role of language, which – on their account – becomes a system of purely conventional symbols used to designate ideas. That a table is called 'a table' in English, 'der Tisch' in German, or 'stół' in Polish, is a result of a convention; but the important thing is that all three words correspond to the same idea. Our understanding of reality is extra-linguistic: we know what tables are, what they look like, and what their functions are since we have the relevant ideas, and not because we use certain expressions.

This view of thinking faces two kinds of objections. On the one hand, and contrary to much of the current psychological and neurobiological theories, it encapsulates a particularly passive view

of cognition, in which judgment and action are logically depend on prior perception. For example, Descartes believes that "there are only two modes of thinking in us, viz., the perception of the understanding and the action of the will. (…) To perceive by the senses (*sentire*), to imagine, and to conceive things purely intelligible, are only different modes of perceiving (*percipiendi*); but to desire, to be averse from, to affirm, to deny, to doubt, are different modes of willing. (…) The will as well as the understanding is required for judging. I admit that the understanding is necessary for judging, there being no room to suppose that we can judge that which we in no way apprehend; but the will also is required in order for us to assent to what we have in any degree perceived" [5, XXXII, XXXIV].

On the other hand, the model of thinking as seeing totally disregards the role of language in cognition. The observational data gathered, and the theories developed since the times of Benjamin Lee Whorf, strongly suggest that the language one speaks influences one's perception and thinking; and even if the strong reading of linguistic relativity may not be supported by facts, there is no escape from the conclusion that the language and culture into which we are born decisively shape the way we perceive and the way we act. The failure to *see* this is one of the grave errors of those who believe that thinking resembles seeing. This problem is best illustrated in relation to the question of how abstract concepts are generated. Let us recall that according to Locke one produces a general idea by abstracting from concrete ideas; moreover, Locke claimed that the mechanism of abstraction is an inborn ability, something we have at our disposal from the very beginning of our lives. But what does "seeing" a general idea consist in? For instance, what would seeing a general idea of a triangle be like? Such a general idea of a triangle can neither represent an acute, a right, nor an obtuse triangle. Immanuel Kant was quick to realize that fact – he claimed that all we can 'see' in intuition are particulars: if one 'sees' a triangle, it is a particular acute- or right- or obtuse-angled triangle; if one 'sees' a tree, it is a particular tree, with a particular height, shape, number of branches, type of leafs, etc. Of course, says Kant, knowledge does not consist of 'pictures' of particular instances of triangles or trees; knowledge is universal and conceptual. Hence, Kant claimed that the concepts we have are general and refer to classes of things. The concept of a triangle embraces all triangles, be they acute, right or obtuse; the concept of a tree refers to all trees; etc. Naturally, a concept (of a triangle or of a tree) is not something one sees; rather, every concept is associated with the so-called transcendental scheme, i.e. a set of instructions one has to follow in order to construct in the intuition a particular instantiation of a given concept. Thus, the concept of a triangle is associated with a transcendental scheme which tells us how to 'draw' an imaginary triangle in the intuition. The outcome of the procedure is always a particular triangle; but the procedure itself, the set of instructions which constitute the transcendental scheme, encapsulate the general properties of *any* triangle. Kant might have failed to realize the extent of the influence language has on perception and action, but he clearly identified the shortcomings of the "mental eye" view of cognition [10], [3].

Unfortunately, the alternative to the model of thinking as seeing – the view that thinking is like a conversation of the soul with itself – is equally lacking. This is a theory which, with some reservations, and in different forms, may be ascribed to numerous philosophers and psychologists such as the late 19th century German psychologist Max Müller, who emphasized that true thought cannot be carried on without words [13]; Lev Vygotsky, who claimed that the structure of language determines the deep structure of thought [16]; or Jerry Fodor who believes in the existence of a universal language of thought or *Mentalese* [8].

The claim that thinking cannot take place outside language – be it talking to oneself in the language our mothers teach us or the use of the mysterious *Mentalese* – is easily dismissed when we consider that according to such a view thinking would be an exclusively human affair, and hence constitute an evolutionary enigma. Moreover, when we limit thinking to the use of words, much of our conscious and meaningful experience would need not only a different label, but also a different explanation. Imagine yourself 'playing in your mind' Bach's *Kunst der Fuge* or remembering someone's face; it surely is non-verbal, but it is deserving of the label "thinking". It may be argued

that the problem is purely definitional and is based on a misunderstanding. The proponents of thinking as conversation are simply not interested in non-verbal mental experiences – they do not deny their existence, but claim that they have different foundations than thinking in words and as such do not deserve the name. But in such a case, one is entitled to ask for an explanation what is the relationship between thinking (which requires language) and other forms of mental life, as well as what are the evolutionary reasons for such a dualism of thought.

3. Conceptual maps

Therefore, it seems that we need an alternative to the two received models of thinking. Such an alternative has recently been sketched by a number of scholars, most notably Peter Gädenfors and Paul Churchland. In a nutshell, their views – even if different in motivations, scope, empirical basis, and other details – boil down to the thesis that our mental representations and, in consequence, the process of thinking which takes advantage of them, are neither (structured) collections of 'mental pictures', nor (structured) sets of propositions. For Gärdenfors, the best description of human individual knowledge is provided by what he calls *conceptual spaces* [6], [7], while Churchland prefers to speak of *conceptual maps* [4]. Let us consider in some detail the latter proposal.

When one's brain receives some input from the environment, it is through the sensory neurons (e.g., those located in rods and cones in the retina, mechanoreceptors within the skin, etc.). However, those populations of sensory neurons do not produce the final representation of the perceived phenomena; rather, they only serve as the first rung in a complex, hierarchical neural structure, where information retained at one level is transformed – through synaptic connections – to a higher level. "Each rung of each of these ladders constitutes a unique cognitive canvas or representational space, a canvas or space with its own structured family of categories, its own set of similarity and difference relations, and its own peculiar take on some enduring aspects of the external world. What happens, as sensory information ascends such a ladder, is its progressive transformation into a succession of distinct representational formats, formats that embody that brain's background 'expectations' concerning the possible ways in which the world can be" [4, p. 35]. In this way, face recognition, the categorization of colours, and any other cognitive activity is enabled by the existence of an appropriate neuronal 'ladder'; it follows, *inter alia*, that the representation of some concept – say, of 'red' – is not an activation pattern at the level of sensory neurons, but rather 'redness' is a part of the conceptual system encapsulated in a hierarchy of neural networks. In such a hierarchy, the information from the lower-level network is transformed through the synaptic connections to the higher-level, and hence the ability to recognize colour red, and to distinguish it from other colours, is determined by the configuration of the connection weights across the transforming population of the synapses. Such a cognitive architecture cannot be properly captured by any system of logic; it is rather a vector space, and thus our representation of the world is geometrical. It is why Paul Churchland is justified in speaking of conceptual maps in our brains, and Peter Gärdenfors – of conceptual spaces.

Crucially, such vector spaces provide a framework for representing both perception and action. "The brain's representations of the world's enduring categorical and causal structure (its 'factual' knowledge), and the brain's representations of its various acquired motor skills and abilities (its 'practical' knowledge), are both embodied in the carefully sculpted metrics of similarities and differences that provide lasting structure to each one of the brain's many activation spaces" [4, p. 49]. Thus – contrary to many philosophical theories and the "received wisdom" of folk psychology – no unbridgeable gap exists between the representations of perception and action.

Another important aspect of the geometrical view of the mind is that the neural activation patterns are a result of the long process of individual learning. We are not born with the ability to discern between red and orange, or distinguish female from male faces. But our brains are so constructed as to enable the gradual formation of neural activation patterns, and this process leads firmly to the establishment of hierarchies of neural networks which make it possible for us to

categorize and conceptualize our experience and act in the world. If so, our mental representations are clearly embodied and enacted. They emerge in the interactions of our bodies with the environment, and they are shaped by what we do, and not – at least not only – by what we passively experience.

There is one further aspect of the above described cognitive mechanism that should be stressed here. The mechanism "displays the capacity to reach beyond the shifting vagaries of one's sensory inputs so as to get a grip on the objective and enduring features of one's perceptual environment" [4, p. 67]. The training of a given neural network, the painstaking orchestration of the numerous synaptic weights, takes place over a long period of time, fine-tuning the entire system to the environment, and hence capturing its unchanging, invariant features. Naturally, not all such features are represented in the human brain – only those which happen to be relevant for our actions. Moreover, the cognitive apparatus of other animals, from quite simple organisms to non-human primates, also capture objective aspects of reality. For example, bats must have representations of the environment which are quite distinct from human conceptual maps, even in relation to the same aspect of the world. "A blind bat, (…) may know what flamingos are, and be perfectly able to discriminate flamingos from other flying creatures, but it will do so by focusing (actively) on the distinctive acoustic signature of the sonar echo returned from such a bird, or (passively) on the equally distinctive acoustic signature ('swoosh, swoosh') given off when any flamingo takes flight" [4, p. 89]. But conceptual maps differ not only across but also within species. No two people have exactly the same neural representations of the given phenomenon: the relevant synaptic connections and their weights may differ to a greater or lesser degree (even though the fact that humans have similar genetic underpinning, as well as that they deal with the same type of environment, significantly reduces such differences).

Churchland claims further that in addition to the slow process of learning, which consists of the formation of the hierarchies of neural networks as described above, there exists another, faster, dynamic process which may also be called "learning". It is conceptual redeployment and it takes place when – without altering any synaptic connections or adjusting the network's weights – the sensory input receives a significantly novel conceptual interpretation. An example of such an occurrence may be the realization that dolphins – with all their fish-like features – are best categorized as mammals; or Darwin's observation that the concepts of natural variation and selective reproduction are applicable to all species throughout the history of life on Earth and that they are capable of explaining the diversity of those species; or Newton's insight that the Moon's elliptical trajectory is only an instance of a flung stone's trajectory [4, p. 188-192].

3. The nature of language

But the two kinds of learning Churchland identifies – the long process of the formation of conceptual maps, and a relatively quick one of conceptual redeployment – do not exhaust the list of mechanisms operating at the human cognitive *niveau*. What is typically human, and what constitutes one of the greatest discoveries in the evolutionary history of our species, is the ability to use language. When a child learns to speak, the effect on her conceptual maps is that they are formed in a way which is typical for the society in which she was born. Of course, some level of idiosyncrasy remains, but it is by far smaller than it would be were there no language, since "most obviously (…), a shared language and a shared theoretical vocabulary allows us to coordinate our individual conceptual maps of a given domain, both in our background understanding of that domain's abstract structure and in our local sensory and instrumental indexings of that map in the business of interrogating the objective world" [4, p. 269]. Apart from amplifying, regulating, and coordinating conceptual maps, there is one more effect that language has on human cognition – it becomes collective:

> It then involves a number of distinct brains – at least a handful, and perhaps many millions – engaged in a common endeavor. Their consensual understanding of the world's general structure, of its local and present configuration, of its immediate past and expected future, is then shaped, not by the activities of and the sensory inputs to a single brain, but by the activities of and inputs to a large number of different brains, similarly but not identically situated [4, 252].

This fact resonates with a number of theories which stress that human cognition is informed by culture, and culture is our common enterprise. In this context, Michael Tomasello speaks of "the cultural ratchet" – a genetically rooted mechanisms for accumulating patterns of conduct and passing them on from generation to generation [15]. Culture – at its core – consists of such patterns, which are not limited to purely linguistic behaviour. Thus, the claim that language amplifies, regulates, and coordinates conceptual maps among individuals, is not precise; what plays those roles are any socially shared patterns of conduct, not only linguistic.

All this reinforces the Wittgensteinian claim that language is not a mirror of reality, but rather a tool humans use to coordinate their perception and action. The fundamental role of linguistic expressions is not to mimic the structure of the world, but to coordinate conceptual maps among the members of the given society. To come back to the already considered example, the sentences such as "Jane likes mangoes" help to form the abstract structure (i.e., the relevant set of conceptual maps) of the domain of experience pertaining to persons, their attitudes and feelings, etc., and a particular utterance of the sentence "Jane likes mangoes" serves to index some concrete part of the interlocutor's conceptual map. At the same time, it is now easier to understand why we tend to think of words as associated with concepts: we believe that the word 'mango' denotes a concept of a certain fruit since our perceptual apparatus is so shaped – through the long process of the formation of conceptual maps – that our brains have a vector activation space, which enables us to distinguish mangoes from other fruits. Therefore, we can repeat that it is sentences that have meaning, since they are the functional units for acting in the world, while the meaning of their component words is something derivative, a theoretical construction – or better a reconstruction – of the "living language". Meanwhile, this reconstruction is made possible by the fact that our conceptual apparatus is organized in such a way that it encodes – in its vector space – the relevant aspects of prototypical mangoes, dogs, tables, and other entities, as well as actions such as grasping, speaking, etc. We believe that in this way the puzzle of the common sense reference disappears: within the framework provided by Churchland's theory of conceptual maps, it is possible to reconcile the Wittgensteinian picture of language as a tool with our strong conviction that at least some linguistic concepts refer to things in the world.

But is coordinating conceptual maps everything that language does? It may be argued that the answer to this question is negative, and that the role of language far exceeds the mere cross-individual orchestration of perception and action: it opens up a completely new cognitive niche. Let us compare the following *ex post* reconstructions of two scientific discoveries. Richard Hamming tries to imagine how Galileo arrived at the discovery of the law of falling bodies. He says:

> Well, Galileo was a well-educated man and a master of scholastic arguments. He well knew how to argue the number of angels on the head of a pin, how to argue both sides of any question. He was trained in these arts far better than any of us these days. I picture him sitting one day with a light and a heavy ball, one in each hand, and tossing them gently. He says, hefting them, "It is obvious to anyone that heavy objects fall faster than light ones – and, anyway, Aristotle says so." "But suppose," he says to himself, having that kind of a mind, "that in falling the body broke into two pieces. Of course the two pieces would immediately slow down to their appropriate speeds. But suppose further that one piece happened to touch the other one. Would they now be one piece and both speed up? Suppose I tied the two pieces

> together. How tightly must I do it to make them one piece? A light string? A rope? Glue? When are two pieces one?" The more he thought about it – and the more you think about it – the more unreasonable becomes the question of when two bodies are one. There is simply no reasonable answer to the question of how a body knows how heavy it is – if it is one piece, or two, or many. Since falling bodies do something, the only possible thing is that they all fall at the same speed – unless interfered with by other forces. There's nothing else they can do. He may have later made some experiments, but I strongly suspect that something like what I imagined actually happened [9, p. 87].

It is clear that Galileo's reasoning – as portrayed by Hamming – may be understood in terms of a sophisticated conceptual redeployment. Language might have played a heuristic role here, but a negligible one. Crucial was the insight that the conceptual map which associates the speed of a falling body with its weight is in fact a blend of two conceptual maps, and that these can be separated.

Now, let us consider Abraham Pais' reconstruction of Heisenberg's discovery of quantum mechanics:

> What, in one dimension, is a classical orbit? It is described by one coordinate *x* that varies continuously as a function of the time *t*, an orbit is given symbolically as *x(t)*. Now Heisenberg seeks inspiration from his previous work with Kramers. There the issue had been to find amplitudes *A(ν)* for the scattering of light with frequency *ν* by an atom. *A(ν)* should depend on the transitions from atomic states *n* to states *m*, as indicated by the symbol (*not* used by Kramers-Heisenberg) $A_{mn}(\nu)$. Now Heisenberg reasoned (I think), let us try to do something similar for *x(t)*, represent it by the 'quantum symbol' $x_{mn}(t)$, where, to fix ideas, *m* and *n* refer to quantum states of a harmonic oscillator, the simplest example he discussed in his paper. There are two possibilities. Either *m* equals *n:* $x_{nn}(t)$, which shall represent the coordinate at time *t* insofar as the system is in state *n*. Or *m* does not equal *n*, when $x_{mn}(t)$ shall represent what one might call a coordinate in transition. Likewise the classical velocity *v(t)* in the orbit shall be represented by $v_{mn}(t)$. All these quantities satisfy Heisenberg's criterion of being 'in principle observable'. Classically the continuous orbit *x(t)* satisfies an equation of motion which tells us how the particle moves from one position and velocity to another. Heisenberg assumes that each of the quantities $x_{mn}(t)$ satisfies that same equation. Next he asks: what is the energy of the particle in the state *n*? Again he takes over the classical expression for the energy, a function of *x(t)* and *v(t)*. Classically one proceeds by finding solutions of the equations of motion, substituting those in the energy expression and so obtaining the corresponding energy values. Heisenberg proceeded likewise in trying to find the quantized energies E_n. But now he had to make a crucial decision. The energy of the oscillator (our example) depends on the squares of *x(t)* and *v(t)*. If one represents *x(t)* by $x_{mn}(t)$, then how should one represent $x^2(t)$? It seems that the simplest and most natural assumption would be:
>
> $$x^2_{mn}(t) = \sum_k \left(x_{mk}(t) \times x_{kn}(t) \right),$$
>
> where $\sum_k$ means: sum over all possible values of *k* [14].

There is no need to continue with Pais' reconstruction. It should be clear by now that the discovery of quantum mechanics would be impossible without the essential, non-heuristic use of language – in this case, the language of mathematics as applied to the fundamental problems of physics. Of

course, Heisenberg's reasoning required quite complex conceptual redeployments – the connections between his conceptual maps must have changed to attain new insights pertaining to the behaviour of the hydrogen atom. However, it is quite unlikely that if someone else independently discovered quantum mechanics, her conceptual redeployments would resemble what happened in Heisenberg's brain. The reason is that mathematics and mathematical physics are quite abstract, and abstract language does not necessarily lead to cross-individual coordination of conceptual maps. For example, many studies suggest that the differences between brain representations of relatively simple mathematical concepts (such as natural numbers) are quite substantial [2].

This point may be generalized: arguably all abstract linguistic concepts in any domain, not only in mathematics, have no well-defined counterparts at the level of the conceptual maps in the brain. Let us consider, for example, the concept of justice. It is unlikely that one would be able to pinpoint any (aspect of) a particular conceptual map or a set of such maps that would correspond to 'justice' (while there may be representations of some classes of acts which usually count as just or unjust). This is not surprising: given the function of language there is no – and there cannot be any – one-to-one correspondence between linguistic concepts and the vector spaces which represent reality at the brain level. Abstract concepts feature in linguistic expressions which, when uttered by parents, teachers, and other people we talk to, do in fact shape our conceptual maps, making them more similar to the maps of the members of our society. But while one can identify corresponding aspects of the conceptual maps for linguistic concepts such as 'mango', 'table' or 'grasping', the same cannot be said of 'justice', 'punitive damages', 'noncommutative geometry' or 'muon'. Higher mathematics, mathematical physics, philosophies of law and of literature, etc., do not exist in our brains – they are publicly shared theories expressed in a suitable language. Of course, through the coordination of individual conceptual maps they do influence the way we perceive and act in the world. But they do more than that. As the example of Heisenberg's discovery and other theories of contemporary physics show, language can give us access to the world which by far exceeds the grasp of reality enabled by the mechanism of the formation of conceptual maps. This fact only reinforces the mystery behind the puzzle of scientific reference: language is not only a tool for acting in the world. While it serves us well in coordinating our conceptual maps, it does more than that – it enables an insight into the structure of reality which lies beyond any individual cognitive apparatus.

4. Coda

In this paper we have argued that the theory of conceptual maps developed recently by Paul M. Churchland provides additional support for Wittgenstein's claim that language is a tool for acting in the world. The role of language is to coordinate and shape the conceptual maps of the members of the given language community, reducing the cross-individual cognitive idiosyncrasies and paving the way for joint cognitive enterprises. Moreover, Churchland's theory also explains our tendency to speak of language as consisting of concepts which correspond to things we encounter in the world. The puzzle of common sense reference is no longer a puzzle: while at the fundamental level language remains a tool for orchestrating conceptual maps, the fact that the maps encode some communally shared categorization of experience fuels our talk of concepts capturing the essences of things, natural kinds, prototypes, etc. At the same time, Churchland's theory fails short of resolving the puzzle of scientific reference: that the mathematical equations of physics help us capture some enduring aspects of reality, which lie far beyond the limits of our sensual experience, remains a mystery [2, p. 203-236]. We believe, however, that this is not an argument against the Wittgensteinian view of language as a tool or Churchland's theory of conceptual maps. Rather, it is a phenomenon which requires metaphysical explanations [1, p. 151-187].

References

1. Brożek, B. *Rule-Following. From Imitation to the Normative Mind*, Kraków: Copernicus Center Press, 2013.
2. Brożek, B., Hohol, M. *Umysł matematyczny*, Kraków: Copernicus Center Press, 2014
3. Brożek, B., Olszewski, A. "The Mathematics of the Transcendental Ego", *Copernicus Center Reports* 2, 2011, 75-124.
4. Churchland, P.M. *Plato's Camera How the Physical Brain Captures a Landscape of Abstract Universals*, Cambridge, MA: The MIT Press 2012.
5. Descartes, R. *The Principles of Philosophy*, translated by J. Veitch, http://www.gutenberg.org/, 1901.
6. Gärdenfors, P. *Conceptual Spaces: The Geometry of Thought*, Cambridge, MA: The MIT Press, 2000.
7. Gärdenfors, P. *The Geometry of Meaning: Semantics Based on Conceptual Spaces*, Cambridge, MA: The MIT Press, 2014.
8. Fodor, J. *The Language of Thought*, Crowell Press: New York, 1975.
9. Hamming, R. "The Unreasonable Effectiveness of Mathematics", *The American Mathematical Monthly*, 87(2), 1980, p. 81-90.
10. Kant, I. *Critique of Pure Reason*, translated by P. Guyer and A.W. Wood, Cambridge University Press: Cambridge 1998.
11. Locke, L. *An Essay Concerning Human Understanding*, Hazleton: Pennsylvania University Press, 2009
12. Margolis, E., Laurence, S. "Concepts", *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), E.N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/concepts/>.
13. Müller, M. "On Thought and Language", *Monist* 1(4), 1891, p. 572-589.
14. Pais, A. *Niels Bohr's Times in Physics, Philosophy, and Polity*, Oxford – New York: Oxford University Press, 1991.
15. Tomasello, M. *The Cultural Origins of Human Cognition*, Cambridge, MA: Harvard University Press, 2001.
16. Vygotsky, L. *Thought and Language*, Cambridge, MA: Harvard Unviersity Press, 1986.
17. Winger, W. "The Unreasonable Effectiveness of Mathematics in the Natural Sciences", *Communications on Pure and Applied Mathematics* 13(1), 1960, p. 1-14.
18. Wittgenstein, L. *Philosophical Investigations*, translated by G.E.M. Anscombe, Oxford: Basil Blackwell, 1986.

DE GRUYTER OPEN

studia humana
QUARTERLY JOURNAL

Studia Humana
Volume 4:2 (2015), pp. 26—38
DOI: 10.1515/sh-2015-0014

Two Types of Visual Objects

Błażej Skrzypulec

Jagiellonian University, Krakow, Poland

e-mail: blazej.skrzypulec@gmail.com

Abstract:
While it is widely accepted that human vision represents objects, it is less clear which of the various philosophical notions of 'object' adequately characterizes visual objects. In this paper, I show that within contemporary cognitive psychology visual objects are characterized in two distinct, incompatible ways. On the one hand, models of visual organization describe visual objects in terms of combinations of features, in accordance with the philosophical bundle theories of objects. However, models of visual persistence apply a notion of visual objects that is more similar to that endorsed in philosophical substratum theories. Here I discuss arguments that might show either that only one of the above notions of visual objects is adequate in the context of human vision, or that the category of visual objects is not uniform and contains entities properly characterized by different philosophical conceptions.

Keywords: perception, metaphysics, visual objects, content, individuation

Our usual visual phenomenology suggests that the human visual system represents the environment in terms of objects. In various circumstances these "visual objects" can differ in complexity, ranging from shapeless bundles of features that appear at the periphery of the visual field to objects with complex parts and structures that are recognized as exemplars of general categories. Nevertheless, they always seem to be individuals, located in space and possessing certain qualitative properties. The above intuition about the importance of objects in vision is preserved within the majority of scientific models of human perception, which explain, *inter alia*, how figures are discerned from ground [e.g., 26], how objects are represented as being the same despite changes [e.g., 13], or how perceived objects are categorized as exemplars of general types [e.g., 41]. Because of this it seems to be generally accepted that the human vision represents objects.

However, from the philosophical perceptive, it seems that that statements claiming that vision represents objects are quite vague. This is because there exist various, mutually inconsistent philosophical notions of objects. In this context it may be asked which conception of objects is correct in the case of visual objects. One of the main demarcation lines separates bundle and substratum theories of objects. According to the bundle theory, objects are identical to relational combinations of features [5], while in substratum theories each object is constituted by an additional element that differs from usual features and is often characterized as an individuator, or as a subject of features [19; 27].

In the paper, I show that in psychological models visual objects are characterized in two distinct ways, one related to the bundle and another related to the substratum conception of objects. More specifically, I argue that the bundle notion of visual objects is connected with models of perceptual organization, while models describing mechanisms that allow us to represent persistence through change characterize visual objects along the lines of substratum theories. Because these two approaches to visual objects are connected with different classes of psychological models, discussion is limited when it comes to whether we should postulate only "bundle" visual objects, only "substratum" visual objects, or if both types are needed to properly characterize the representational abilities of the human visual system.

The main goal of the paper is to sketch a "conceptual map" of the problem, i.e. to investigate the main reasons for stating either that there is only one type of visual object or that we should postulate both "bundle" and "substratum" visual objects. My goal is not to provide a final answer, which in fact cannot be given based on the current state of empirical investigations, but rather to identify conditions whose satisfaction would provide a strong argument for one of the considered options. I start by explicating my usage of the term "visual object", and then explain the difference between bundle and substratum conceptions of objects (section 1). Subsequently, I show how the distinction between the bundle and substratum conceptions of objects is connected with the treatment of visual objects in models of perceptual organization and models of visual persistence (section 2). The central focus of the paper is discussion of arguments in favour of accepting one or both types of visual objects (sections 3–7).

1. Bundles, substrata, and visual objects

As was stated above, the human visual system seems to represent objects, which may vary in complexity according to their place within the perceptual process. For example, early visual objects may be characterized as combinations of co-located features [40], and later ones as persisting individuals possessing complex parts [33]; while at the highest level of visual information processing, objects can be represented as exemplars of general categories [12]. It is important to note that the characteristics of visual objects do not have to be exactly the same as characteristics of physical objects that causally interact with the visual system. In particular, the same physical object can be represented differently at subsequent stages of the perceptual process. What is more, in cases of misperception, represented characteristics will be different from those possessed by the object that stands in a causal relation with perceptual mechanisms.

Because of this, statements concerning visual objects should be interpreted as concerning the content of visual representation, or, in other words, as specifying the necessary adequacy conditions of visual representations [35]. In the case of human vision, where the category of objects seems to be crucial, adequacy conditions will usually be connected with the presence of certain objects within the visual field. From this perspective, stating that a red square is the visual object of a representation, or, in other words, that a red square is visually represented, means that a necessary adequacy condition of the concerned representation is the presence of a red square in the visual field.

The specification of visual objects as "objects of content", i.e. objects whose presence is a necessary condition for a visual representation's adequacy, makes it possible to interpret the claims of the philosophical bundle and substratum theories in a way that allow to treat them as characteristics of representational content. According to the main claims of bundle theories, every object is identical to a relational combination of features [1: 90; 34: 78; 39). In the context of the functioning visual system, this statement can be formulated as a thesis that representing an object is equivalent to representing a proper combination of features (Which combination are the "proper" ones may vary between different bundle theories):

(B) An object is visually represented if and only if a proper combination of features is visually represented.

In other words, thesis (B) states that the necessary adequacy conditions for visual representations representing objects are the same as necessary adequacy conditions for visual representations representing combinations of features. If the thesis (B) is correct, then to characterize the content of visual representations it is sufficient, at least as far as representations of objects are concerned, to describe features and the ways in which they are related, without the need to postulate any additional elements.

Proponents of substratum theories deny the main claim of bundle theories by arguing that objects cannot be identified with combinations of features [9; 18: 140–143). Instead, it is postulated that the structure of every object contains a special element, different from features, often called a "substratum". This additional element is thought to serve at least one of three roles. First, it may be an individuator, if having the same substrata is an identity criterion of objects [27]. Second, it may be a subject, if it "instantiates" other elements of an object, i.e. features, where "instantiation" is an asymmetric relation [7]. Third, and most common, it may be a unificator, if without the presence of the substratum the other elements of a given structure would not constitute an object [22].

Relying on the main claims of substratum theories—irreducibility of objects to feature-bundles and the presence of a substratum interpreted as an individuator, subject, or unificator—the following thesis can be postulated:

(S) *An object is visually represented if and only if a "substratum" element (serving the role of an individuator, subject, or unificator), usually combined with features, is visually represented.*

According to (S), and in contrast to (B), the content of visual representations representing objects cannot be sufficiently characterized in terms of combined features. This is because representing an object is always connected with representing a substratum, which serves one of the three roles described above.

In the next section, I show that theses (B) and (S) are assumed within the characteristics of representational content postulated in certain psychological models of vision.

2. Bundles, substrata, and psychological models

One of the most important categories of psychological models of vision are those of perceptual organization [25: 255–257]. Such models describe how the visual system creates a representation of objects in the visual field by representing simpler elements as composing more complex wholes. According to a popular view, connected with classical physiological research by Hubel and Wiesel [11] and Marr's fundamental work in the cognitive science of vision [21], the human perceptual system starts by representing local discontinuities between different surface features, like brightness or hue. These simple discontinuities, if they stand in proper spatial relations, are combined to form more complex edges, which together compose an early sketch of the visual field.

This sketch, based on the spatial configuration of edges, serves as a starting point for more complex operations. Most importantly, closed edges designate regions filled with uniform surface features, which are regarded as basic units from which visual objects are constructed [26]. First, some "uniform regions" can be distinguished from the ground and thus gain the status of figures, i.e. simple visual objects. This figure-status can be obtained if a region stands in appropriate qualitative relations to its neighbouring regions, where relevant relations are connected with, *inter alia*, being smaller, more convex, or more symmetrical [28; 32]. Second, simple figures, if they are spatially connected, may be combined into complex objects then serving as parts of a higher-level whole [26; 32]. Third, nearby objects possessing similar features can be connected into a perceptual

group, according to some Gestalt-like laws, forming a whole constituted by spatially-dispersed elements [8; 16]. Finally, objects can gain internal structure from the pattern of edges by which they are designated. In particular, points of concavity, at which different borders of an object are close to each other, allow for distinguishing parts within an object's structure [10]. An important subclass of perceptual organization models can be found in conceptions of visual binding, of which the most influential example is the Feature Integration Theory [40]. According to this theory, vision does not start from the representation of local discontinuities, but from representing simple features, without representing any relations between them. Subsequently, due to the work of attentional mechanisms, co-located features forms feature-bundles, which may be regarded as simple visual objects.

Despite the huge variety of perceptual organization models, they all seem to agree on a general view of the nature of visual objects. According to these models, at the beginning of the perceptual process uncombined features, or simple combinations of them (like local discontinuities) are represented, which, by standing in appropriate spatial relations, compose visual objects. Within this approach it seems that representing an object is always equal to representing a relational combination of features. Because of this, models of perceptual organization accept thesis (B), which is a perceptual counterpart of the main postulate of the philosophical bundle theory of objects.

It should be noted that the above conclusion are independent from considerations about the exact point in the perceptual process at which the representational content starts to be present. It may be the case that elements in which the visual system detects local discontinuities are not representations at all, but merely detectors of some causal influences, and the first representations are in fact those of uniform regions or figures distinguished from ground. However, no matter what the earliest elements of content are, representing them, according to models of perceptual organization, will be always equal to representing some combinations of features.

Models of perceptual organization usually focus on stationary phenomena, and lay out the steps of constructing a visual representation of the environment. In this context it is interesting to observe that models whose main focus is to explain how the visual system represents persistence through time assume a different view on visual objects.

Such models of visual persistence postulate perceptual devices, often called "visual indices" or "object files", that are engaged in representing objects but whose role is not to represent common visual features like hue, shape, or localization [13; 17; 31: 37–39**)**. A single index or file is activated when a causal interaction occurs that is, in ordinary circumstances, connected with the presence of an object within the visual field. According to models of visual persistence, indices and files allow us to recognize the numerosity of objects, to track them despite movement and qualitative changes, and to refer to them in order to gain further information regarding their features [30; 36].

The representational content connected with such perceptual devices is limited to representing several numerically distinct individuals and representing whether a presently represented individual is the same as the one represented at an earlier moment. In usual situations, when other perceptual mechanism are also activated, it is not only represented that there are numerically distinct individuals, but also that these individuals possess some features. Because of this, according to models of visual persistence, the structure of visual objects is composed of two types of element: a simple individual and features associated with it.

The above notion of visual objects is inconsistent with thesis (B). Representing an object is not equal to representing a combination of features, but rather to representing a simple individual that may possess some features. Every such individual is numerically distinct from other individuals and so may serve as an individuator of visual objects. It also possesses features and so may be interpreted as a subject. What is more, such individuals seem to be unificators, since other elements of the object's structure, i.e. its features, would not compose an object without being related to an individual. Because of this, models of visual persistence, in contrast to models of perceptual organization, seem to incorporate thesis (S), formulated by relying on the main postulate of substratum theories of objects.

The two incompatible conceptions of visual objects described above—"combinations of features" and "individual plus features"—were developed independently in connection with different classes of psychological models. Because of this, it may be asked whether we really need to postulate both types of visual object in the context of the human vision. In what follows, I investigate different answers to this question and the reasons that underpin them.

3. Simplicity and explanatory power

The bundle notion of visual objects has an initial advantage over the substratum notion because it relies on weaker assumptions. It offers a simpler account of visual content according to which it can be sufficiently described in terms of combinations of features. What is more, it is uncontroversial that features such as colours, shapes, and localizations are represented by the human visual system, and so it seems intuitively plausible that the characterization of visual content consists in specifying arrangements of features. On the contrary, proponents of the substratum approach expand our visual "ontology" by postulating an additional type of element, i.e. substrata, that play some formal roles but cannot be identified with any entities that are usual visually represented. Because of this, the burden of proof seems to be on proponents of the substratum notion of visual objects. If there is no justification for postulating an additional element of visual content in connection with the representations of objects, then the proper account of visual objects will be some version of the bundle approach.

The substratum view of visual objects can be justified by pointing to a perceptual phenomenon, connected with the representational abilities of human vision, that cannot be explained by treating visual objects as relational combinations of features. However, presenting such a phenomenon is no trivial task, since the bundle notion of visual objects, connected to models of perceptual organization, possesses significant explanatory power.

Starting from low-level elements of content, representing local discontinuities may be interpreted as the representation of mutually-exclusive surface features (like different levels of brightness) combined with spatially connected localizations. Further, representing edges may be equal to representing local discontinuities standing in appropriate spatial relations. In similar fashion, a uniform region designated by edges can be characterized as spatially coherent localization that is connected with a surface feature and not a part of a bigger localization connected with the same surface features.

The same approach to characterizing representational content naturally extends to representations connected with further perceptual processes, which allow for figure/ground discrimination and grouping. Representing a figure can be identified with representing a uniform region whose features stand in certain relations to features of neighbouring regions (like "being more convex" or "being more symmetrical"). Processes of grouping are connected with representing figures standing in relations of similarity and spatial proximity; while representing spatially connected figures is often sufficient for representing complex objects composed of them.

It seems that a huge variety of processes responsible for modelling the spatial arrangements of elements in the visual field produce representations whose content can be described in terms of features standing in relations of spatial arrangement, similarity, or comparison (e.g. "being smaller"). In the next section I consider phenomena that may force us to modify this coherent picture and thus make a room for visual objects as characterized by the substratum notion.

4. Synchronic and diachronic individuation

Because one of the main roles of substrata is that of being individuators, it seems plausible that perceptual which consist in representing objects as being the same or different, may justify the application of the substratum notion of visual objects. Amongst phenomena of perceptual

individuation we may distinguish those connected with synchronic individuation, when the visual system differentiates between objects composing a single stationary scene, and those connected with diachronic individuation, when it is recognized whether changing objects are the same or different.

4.1 Synchronic context

In synchronic individuation, the status of features describing the location of objects is controversial. According to some it is impossible to visually represent that two objects simultaneously occupy the same location [4]. However, many authors claim that such situations actually occur in human perception when, for example, overlapping semi-transparent patterns are observed or reflections on glass surfaces are seen [6; 23; 31: 40–42]. If the former are correct and locations play a privileged role in synchronic individuation, then it may be that a visual object *x* is synchronically identical to a visual object *y* if and only if *x* has the same location as *y*. In the second case, where spatially overlapping objects may be represented as being distinct, a different rule will apply: a visual object *x* is synchronically identical to a visual object *y* if and only if *x* has exactly the same features as *y*.

Despite this controversy concerning the role of locations, both solutions use only the bundle notion of visual objects. In these cases, representing that there are two objects consists in representing them as having different features concerning localization or representing them as having any other difference in features, and so no new element, like a substratum, is needed. In order to justify applying the substratum notion of visual objects in order to explain synchronic individuation, it would be necessary to show that the visual system can represent completely overlapping objects, which are different while at the same time have share the same localization as well as other features.

I believe that there are two types of perceptual phenomena that can be regarded as those of "complete overlap". However, in both cases one may raise serious doubts as to whether they really occur in the context of the human vision. First, it can be represented that moving objects overlap at some point in their spatial trajectories. If two such objects are represented as having exactly the same features, such as size, shape, and colour, then during spatial overlap they would be represented as sharing all of their features. If such objects, during overlap, are represented as being distinct, while being represented as having the same features, then they are individuated in virtue of something other than a difference in possessed features. This additional aspect of visual objects serves as a synchronic individuator and can be identified with a substrate. Nevertheless, one may doubt whether such situations of complete overlap during movement are really represented by the human visual system. For example, it may be the case that in every situation of overlap one object is represented as being below the other one due to some depth cues, or that only one object, resulting from merging the two earlier ones, is represented [see 42].

A second type phenomenon that may be regarded as an instance of "complete overlap" is connected with occlusion. It is well-known that an object that hides behind an obstacle and reappears after a short time is visually represented as being the same [38]. What is more, investigations in the field of developmental psychology show that infants can distinguish whether one or two objects have been hidden behind an occluder and represent that these objects still exist during the occlusion [3]. If the visual system is able to represent the presence of two simultaneously occluded objects, then it may seem that representing the difference between them is not founded on representing that they have different features. These objects are represented as having the same approximate localization, connected with the localization of an obstacle, and during the occlusion are not represented as having any features like colour or shape. Again, in this case representing the distinctiveness of objects requires representing something more than an arrangement of features, and this additional "something" can be seen as a substratum. Nevertheless, one may doubt whether situations in which two or more objects are simultaneously represented as being occluded are still

genuine visual representations. Maybe such representations are a product of higher-level reasoning that is not strongly connected with the visual processes.

The above considerations show that the bundle notion of visual objects would be inadequate in fully accounting for perceptual phenomena of synchronic individuation—that is, if there are cases of "complete overlap" in which objects are represented as being distinct while at the same time being represented as having exactly the same features. Such situations may happen when objects overlap during movement or are hidden behind an occluder, but the current state of empirical investigations probably does not allow for deciding whether they occur in the context of human vision. However, if cases of complete overlap really occur, then the representational content they involve should be characterized using the substratum notion of visual objects, in which substrata play the role of individuators.

4.2 Diachronic context

Within the field of cognitive psychology it is widely accepted that the human vision is able to represent objects as being the same despite movement and changes in qualities like colour or shape [see 37]. Because of this, in diachronic individuation it cannot be simply stated, as was plausible in the synchronic cases, that a visual object *x* is identical to a visual object *y* if and only if *x* has exactly the same features as *y*. An object can be represented as having different features at different times while still being represented as the same individual.

It may seem that representing sameness in the diachronic context automatically leads to the conclusion that the bundle notion of visual objects is inadequate. According to the bundle notion, representing a particular object is equivalent to representing a particular combination of features. If a change occurs, then a different combination of features is represented. So it seems that, according to the bundle notion of visual objects, a different object is also represented. In this case it would be impossible to explain how the visual system represents the diachronic identity of changing objects by treating them as simply relational combinations of features. However, the bundle notion of visual objects does not in fact entail that representing different combinations of features existing at different times cannot be equivalent to representing a single persisting object. For example, it may be claimed that a visual object *x* is diachronically identical to a visual object *y* if and only if (I) *x* stands to *y* in a similarity-like relation, which is reflexive, symmetric, but not transitive; or (II) *x* is connected to *y* by a chain of visual objects standing in such a relation. In this case representing diachronic identity would be equivalent to representing certain relations between combinations of features, so this phenomenon seems to be adequately grasped in terms of the bundle notion of visual objects.

In fact, in psychological works it is frequently claimed that the role of this similarity-like relation is served by spatiotemporal continuity. Results of experiments involving tracking and reidentifying changing objects suggest that continuous movement is usually connected with representing objects as being the same, even if other features, like shape or colour, change [2; 31: 37]. On the other hand, disturbances of spatiotemporal continuity, except cases of occlusion when the object is briefly hidden behind an obstacle, makes object tracking significantly harder and can break the identity of visual objects [38; 20]. Because of this, it may be proposed that representing diachronic identity between objects is equal to representing combinations of features standing in certain patterns of continuity relations.

Nevertheless, the above picture becomes more complicated if we consider ambiguous cases in which the occurrence of spatiotemporal continuity does not determine the pattern of identity relations. Such cases might include splitting-like situations, when a visual object A, at some moment T_1, is spatiotemporally continuous with two objects B and C at a subsequent moment T_2. In such a case, if all considered objects have the same features in terms of size, color, etc., three alternative patterns of identity relations are possible. First, object A is not identical to either B or C. Second, object A is identical to both B and C. This would lead to the conclusion that visual

diachronic identity is not the classical identity relation, because characterizing visual identity as transitive would entail a contradiction by identifying objects B and C. The third option is that object A is identical to only one of the objects B and C.

If the third option adequately captures the behaviour of the human visual system in splitting-like cases, then the occurrence of spatiotemporal continuity appears to be insufficient for the identity of visual objects. What is more, such ambiguous situations do not break diachronic identity, because object A is still identified with one of the objects in the subsequent moment. Because of this, representing diachronic identity is not equal to representing continuous bundles of features, and some additional element of content should be postulated that determines whether object A is identified with B or with C. Of course, this additional element may be interpreted as a substratum, serving the role of individuator. However, patterns of visual identity in splitting-like situations have rarely been studied, and there is no decisive account of how the human perceptual system behaves in such situations [see 15; 24].

In summary, the representational content connected with usual cases of synchronic and diachronic visual individuation can be adequately characterized using the bundle notion of visual objects. However, if cases of "complete occlusion" are represented by human vision, or if in splitting-like cases only one "later" object is identified with the "earlier" object, there are serious reasons for applying the substratum notion of visual objects in order to characterize the content connected with these situations.

5. Numerical difference

The perceptual phenomena of individuation in synchronic and diachronic contexts are not the only ones that may justify use of the substratum notion of visual object. In particular, the bundle notion of visual object is clearly inadequate if situations exist in which human vision represents only numerical sameness or difference of objects without representing any of their features. Within such cases, representing an object would be the same as representing a simple individual that is not qualitatively characterized to any extent. Visual objects of this kind can be regarded as the simplest objects satisfying the substratum notion, which are identical with substrata and are not constituted by any features. Of course, such substrata would be individuators as well as unificators of the visual objects to which they are identical.

According to some models of vision, of which Pylyshyn's FINST model is probably the most influential example [31: 39–40], experiments involving simultaneous tracking of several objects reveal that objects can be represented as being the same even when we don't represent any of their features. Similarly, the numerosity of a small set of objects can be immediately grasped without serial counting (a phenomenon known as "subitizing", [29]), which may suggest that it is possible to represent several numerically different individuals prior to representing their features.

However, it is far from being universally accepted whether the human visual system is actually able to represent such featureless but numerically distinct objects. It is commonly observed that in tracking experiments, changes in targets' features, such as colour or shape, not only compromise tracking, but participants are also often unable to report seeing these changes [31: 37]. This might show that during tracking objects can be represented as not having any features. Nevertheless, it is harder to argue that objects can be represented without representing their localization. In fact, the proximity of locations is the most important factor determining identity between objects represented at subsequent moments [15].

Because of this, it may seem more plausible that visual representations of objects also always involve representations of their localizations. While accepting this position substantially weakens claims about representing featureless but numerically distinct objects, it might still be sufficient to justify the application of the substratum notion of visual objects. For example, Ronald Rensink, in his works presenting the coherence theory of attention, postulates an early form of visual representation called "layout" [33]. Layout represents some locations as containing objects

without representing any other features of those objects. Layout is thought to serve as a guide for the serial attentional mechanism, which in virtue of layout can quickly gain access to the most important elements of the visual field and then allows us to form more detailed representations of qualitatively rich objects.

A distinguished location represented by layout is a visual object composed of two elements: a feature describing localization and an additional element that differentiates this location from others that do not contain interesting objects. This additional element cannot be identified with any usual visual features, because no such features are represented by layout, and so it can be interpreted as a type of substratum. In this case, the substratum would not serve the role of individuator, due to the fact that different distinguished locations are individualized by the component describing spatial position. However, the substratum within distinguished locations fulfils another usual role, as postulated in the substratum theories. It will be an unificator, because the structure of a distinguished location would not exist without containing a substratum.

The above considerations shows that the perceptual phenomena of individuation are not the only ones that might justify use of the substratum notion of visual object. If the human visual system is able to represent featureless but numerically distinct objects, or to distinguish locations without ascribing any other features to objects, then there will be cases of representational content that cannot be characterized using only the bundle notion of visual objects.

6. No representations without objects

In the two previous sections, I described phenomena whose presence would justify describing visual objects using the substratum notion, but not the bundle notion of visual objects. In addition, it may be asked whether there are reasons to postulate a stronger thesis, namely that the visual objects properly described by the substratum notion are the only ones which are represented by the human visual system. I believe that two such arguments can be found in the literature concerning models of visual persistence.

First, it has been claimed that it is impossible to visually represent combinations of features without representing that they characterize a single object [14: 318–321; 31: 87–89]. To adequately represent the visual field, the perceptual system needs information not only about the presence of certain features but also about their arrangement [4]. However, from the mere fact of representing some features it cannot be inferred which of them coincide. In particular, a colour-feature can, in principle, be combined with many location-features, but combining them randomly would lead to severe misperceptions. To avoid such a result, a colour-feature should not be combined with just any location-feature, but with a feature of proper size, shape, and position within the spatial framework. Since it seems that, from representing a colour-feature, the location-feature with which should it be combined cannot be inferred, one might think that the visual system does not start from the representation of features but from the representation of objects. Further perceptual processes, for example those connected with attentional mechanisms, may allow us to represent the features that this early objects possess. Such objects would be different from usual visual features and thus may be interpreted as substrata. In such a case, there would be no visual object simply identical to combinations of features, since each of them would also be constituted by an additional substratum element.

Second, one may doubt, as mentioned in the section 3, that the early visual processes described in models of perceptual organization play any representational role at all [31: 74–76, 81]. For example, detecting edges or uniform regions may not be connected with representing such patterns, but may consist in merely registering certain causal influences. Only later processes, using this registered information, produce actual representations that possess adequacy conditions and thus also content. If it were the case that the very first representations were connected with mechanism that recognized the numerosity of objects or their diachronic sameness, prior to representing any of their features, then it would be the case that representing objects would always

involve representing numerically distinct individuals that could, but would not have to, possess features.

While the above arguments may have some intuitive appeal, they both assume some rather strong views about the nature of the human vision, which are not sufficiently justified by current empirical results. First of all, even if representing the presence of a certain combination of features has to be preceded by the detection of an object, it is not obvious that such detection has to be accompanied by our representing that there is an additional element, different to these features. It may be the case that processes connected with detecting an object, as well as processes that allow us to decide which of the features should be represented as combined, consist merely in the registration of a certain causal influence and are not connected with the modification of the representational content.

According to the second argument, it should be observed that a crucial feature of representations is the possibility of misrepresentation. If something cannot be inadequate, such as the effect of a causal influence that is not "right" or "wrong", but is just determined by the properties of the elements engaged in an interaction, it is not a representation. While it may be the case that simple operations described in models of perceptual organization, like those responsible for detecting edges, are not representational but just register causal influences, this is less likely in the case of grouping and distinguishing figures from a ground. To represent some elements as forming a single perceptual group it has to be judged that they are significantly similar. Analogously, to represent a certain region as a figure distinguished from a ground its features have to seem more important than those of neighbouring regions. The effects of both these operations may intuitively seem prone to error and not automatically determined by the properties of stimuli. In addition, the second argument assumes that human vision may represent numerically different objects without representing their features, which is also far from being obvious (see section (5)).

7. Bundles, substrata, or both?

Based on the above investigations, I shall now try to characterize the conditions in which it would be justified to postulate the existence of one or both types of visual object in the context of human vision. More specifically, three hypotheses may be considered:

(I) All visual objects satisfy the bundle notion, and so, according to thesis (B), representing an object is equivalent to representing a relational combination of features.

(II) All visual objects satisfy the substratum notion, and so, according to thesis (S), representing an object is equivalent to representing a substratum that may be combined with some features.

(III) Some visual objects satisfy the bundle notion and some satisfy the substratum notion, i.e., in the context of human vision there are two types of visual object with different ontological structures.

The considerations presented in sections 3 to 6 reveal five questions that are relevant in judging hypotheses (I), (II), and (III):

(1) Are there cases in which human vision represents distinct objects that share all their features (that is, situations of "complete overlap")?

(2) Is it the case that in splitting-like situations diachronic identity is maintained between the earlier object and one of the later objects?

(3) Are there cases in which human vision represents numerically distinct objects without representing any of their features?

(4) Are there cases in which human vision represents locations as containing objects without representing any other features of these objects?

(5) Is it the case that a combination of features cannot be represented without representing an element that is different from usual features and serves one of the usual roles of substrata?

To justify hypothesis (I), we have to answer 'no' to all the above questions. The phenomena referred to in questions (1)–(4) involve visual objects containing elements that are different to features; as such, these visual objects cannot be adequately characterized using the bundle notion. What is more, a positive answer to question (5) would entail that there are no visual objects that are simply combinations of features.

The last question is also crucial for hypothesis (II). Only if we answer 'yes' do all visual objects satisfy the substratum rather than the bundle notion. The other questions are not relevant for hypothesis (II), since answering them positively leads to the applicability of the substratum notion in some, but not all cases in which objects are visually represented. If hypothesis (II) is true, then all visual objects are constituted by a substratum playing the role of unificator; but this substratum is not necessarily an individuator, since objects may be individuated in virtue of possessing different arrangements of features.

Finally, hypothesis (III) would be justified if the answer to at least one of questions (1)–(4) was positive, but the answer to question (5) negative. In such a case, there would be some phenomena involving visual objects that could not be reduced to combinations of features, but there would be no reason to suppose that there were no visual objects satisfying the bundle notion. In the case of phenomena related to questions (1)–(3), the substrata constituting visual objects would be both unificators and individuators. However, in the case of distinguished locations referred to in question (4), substrata would only serve as unificators, because such visual objects are individuated by features describing locations. It is worth noting that none of the discussed phenomena, whose occurrence would justify the presence of visual objects satisfying the substratum notion, suggest a need to postulate substrata serving the role of subjects of features. This is because there is no salient reason to assume that the relation between substrata and features in the structure of visual objects has to be asymmetric.

In earlier sections, I claimed that the current state of research concerning vision prevents us from determining which of hypothesis (I)–(III) is true for human perception. Nevertheless, the investigations conducted within this paper suggest that hypothesis (II) may be least probable, since it relies on strong and controversial assumptions about the nature of visual processing; and, by contrast, the bundle notion of visual objects possesses significant explanatory power. However, to reject hypothesis (I) it would be sufficient to prove the existence of a single phenomenon, related to individuation or the representation of numerical difference, like those described in sections 4 and 5, such that the representational content connected with it could not be fully characterized in terms of a combination of features. Because of this, one may rationally suppose that the category of visual objects is not ontologically uniform, but contains two types of objects adequately described by the bundle or substratum theories respectively.

8. Conclusion

I argued that scientific models of human vision assume two, mutually inconsistent notions of visual objects. Models of perceptual organization adopt the main thesis of the philosophical bundle theory of objects and so characterize visual objects as relational combinations of features. Models of visual persistence, however, characterize visual objects in accordance with the philosophical substratum theory of objects and because of this characterize visual objects as constituted by a substratum that cannot be identified with any usual features.

After investigating reasons for claiming the existence of only one or both types of visual object in the context of human vision, I argued that it would be difficult to defend the thesis that all visual objects satisfy the substratum notion. However, it is not necessarily the case that every visual object can be adequately described as a combination of features. If within the context of human vision certain phenomena occur, connected, *inter alia*, with the individuation of objects or with representing distinct objects in the absence of features, then the substratum notion is required for

their characterization. In this case, the category of visual object would contain two types of ontologically different structures.

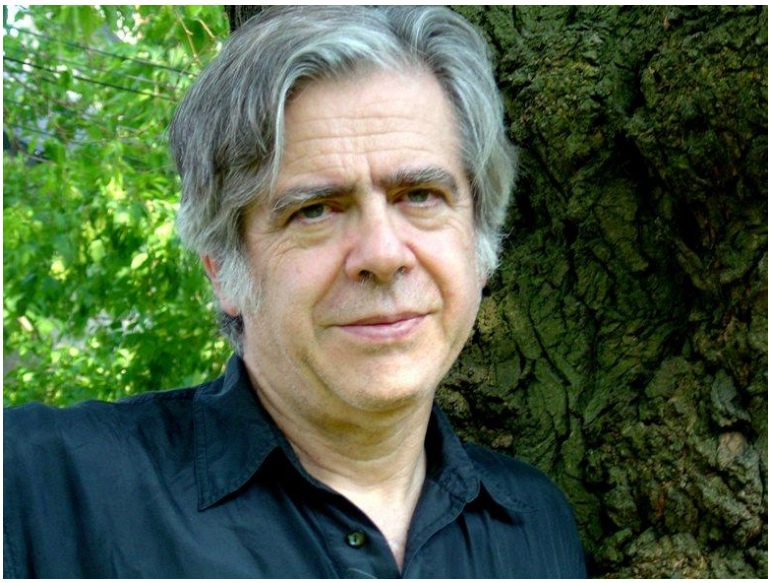
References

1. Armstrong, D. M. 1978, *Universals and Scientific Realism. Vol. I: Nominalism and Realism*. Cambridge: Cambridge University Press.
2. Bahrami, B. 2003, "Object Property Encoding and Change Blindness in Multiple Object Tracking", *Visual Cognition* 10, pp. 949–963.
3. Carey, S., Xu, F. 2001, "Infants' Knowledge of Objects: Beyond Object Files and Object Tracking", *Cognition* 80(1–2), pp. 179–213.
4. Clark, A. 2004, "Sensing, Objects, and Awareness: Reply to Commentators", *Philosophical Psychology* 17(4), pp. 563–589.
5. van Cleve J. (1985), *Three Versions of the Bundle Theory*, Philosophical Studies, 47(1), 95–107.
6. Cohen, J. 2004, "Objects, Places, and Perception", *Philosophical Psychology* 17(4), pp. 471–495.
7. Denkel, A. 2000, "The Refutation of Substrata", *Philosophy and Phenomenological Research* 61(2), pp. 431–439.
8. Elder, J. H., Goldberg, R. M. 2002, "Ecological Statistics of Gestalt Laws for the Perceptual Organization of Contours", *Journal of Vision* 2(4), pp. 324–353.
9. Hochberg, H. 1965, "Universals, Particulars, and Predication", *The Review of Metaphysics* 19(1), pp. 87–102.
10. Hoffman, D. D., Richards, W. A. 1984, "Parts of Recognition", *Cognition* 18(1–3), pp. 65–96.
11. Hubel, D., Wiesel, T. N. 1962, "Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex", *The Journal of Physiology* 160, pp. 106–154.
12. Hummel, J. E. (2013), *Object Recognition*, in D. Reisburg (ed.), *Oxford Handbook of Cognitive Psychology*, Oxford University Press, Oxford, pp. 32-46.
13. Kahneman, D., Treisman, A. M., Gibbs, B. J. 1992, "The Reviewing of Object Files: Object-Specific Integration of Information", *Cognitive Psychology* 24(2), pp. 175–219.
14. Keane, B. P. 2009, *Visual Objects as the Referents of Early Vision: A Response to A Theory Of Sentience*, in D. Dedrick, L. Trick (eds.), *Computation, Cognition, and Pylyshyn*, The MIT Press, Cambridge, MA, pp. 303–333.
15. Keane, B. P, Pylyshy, Z. W. 2006, "Is Motion Extrapolation Employed in Multiple Object Tracking? Tracking as a Low-Level, Non-Predictive Function", *Cognitive Psychology* 52(4), pp. 346–368.
16. Kubovy, M, Holcombe, A. O., Wagemans, J. 1998, "On the Lawfulness of Grouping by Proximity", *Cognitive Psychology* 35(1), pp. 71–98.
17. Leslie, A. M., Xu, F., Tremoulet, P. D., Scholl, B. J. 1998, "Indexing and the Object Concept: Developing 'What' and 'Where' Systems", *Trends in Cognitive Science* 2(1), pp. 10–18.
18. Loux, M. J. 1978, *Substance and Attribute. A Study in Ontology*. Dordrecht: D. Reidel.
19. Lowe, J. E. 2000, "Lock, Martin, and Substance", *The Philosophical Quarterly* 50(201), pp. 499–514.
20. van Marle, K., Scholl, B. J. 2003, "Attentive Tracking of Objects vs. Substances", *Psychological Science* 14, pp. 498–504.
21. Marr, D. 2010, *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, MA: MIT Press.
22. Martin C. B. (1980), "Substance Substantiated", *Australasian Journal of Philosophy* 58(1), pp. 3-10.
23. Matthen, M. P. 2004, "Features, Places, and Things: Reflections on Austen Clarks's Theory of Sentience", *Philosophical Psychology* 17(4), pp. 497–518.
24. Mitroff, S.R., Scholl, B.J. Wynn, K. 2004, "Divide and Conquer: How Object Files Adapt When a Persisting Object Splits Into Two", *Psychological Science* 15, pp. 420–425.

25. Palmer, S. E. 1999, *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
26. Palmer, S., Rock, I. 1994, "Rethinking Perceptual Organization: The Role of Uniform Connectedness", *Psychonomic Bulletin and Review* 1(1), pp. 29–55.
27. Park, W. 1990, "Haecceitas and the Bare Particular", *The Review of Metaphysics* 44(2), pp. 375–397.
28. Pomerantz, J. R., Kubovy, M. 1986, *Theoretical Approaches to Perceptual Organization. Simplicity and Likelihood Principles,* in K. R. Boff, L. Kaufman, J. P. Thomas (eds.), *Handbook of Perception and Human Performance,* Wiley, New York, pp. 1–46.
29. Pylyshyn, Z. W. 1994, "Some Primitive Mechanisms of Spatial Attention", *Cognition* 50, pp. 363–384.
30. Pylyshyn, Z. W. 2001, "Visual Indexes, Preconceptual Objects, and Situated Vision", *Cognition* 80(1), pp. 127–158.
31. Pylyshyn, Z. W. 2007, *Things and Places. How the Mind Connects with the World*. Cambridge, MA: MIT Press.
32. Qiu, F. T., von der Heydt, R. 2005, "Figure and Ground in the Visual Cortex: V2 Combines Stereoscopic Cues with Gestalt Rules", *Neuron* 47(1), pp. 155–166.
33. Rensink, R. A. 2000, "The Dynamic Representation of Scenes", *Visual Cognition* 7(1/2/3), pp. 17–42.
34. Russell, B. 2009, *Human Knowledge: Its Scope and Limits*. New York: Routledge.
35. Schellenberg, S. 2011, "Perceptual Content Defended", *Nous* 45(4), pp. 714–750.
36. Scholl, B. J. 2001, "Objects and Attention: The State of Art", *Cognition* 80(1), pp. 1–46.
37. Scholl, B. J. 2007, "Object Persistence in Philosophy and Psychology", *Mind and Language* 22(5), pp. 563–591.
38. Scholl, B. J., Pylyshyn, Z. W. 1999, "Tracking Multiple Items through Occlusion: Clues to Visual Objecthood", *Cognitive Psychology* 38, pp. 259–290.
39. Simons, P. 1994, "Particulars in Particular Clothing: Three Trope Theories of Substance", *Philosophy and Phenomenological Research* 54(3), pp. 553–575.
40. Treisman, A. M. 1998, "Feature Binding, Attention and Object Perception", *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences* 353(1373), pp. 1295–1306.
41. Ullman, S. 1996, *High-Level Vision. Object Recognition and High-Level Vision*. Cambridge, MA: MIT Press.
42. Viswanathan, L., Mingolla, E. 2002, "Dynamics of Attention in Depth: Evidence From Multi-Element Tracking", *Perception* 31(12), pp. 1415–1437.

Studia Humana
Volume 4:2 (2015), pp. 39—44
DOI: 10.1515/sh-2015-0015

Are Our Emotions True Cognitions?



Ronald Bon de Sousa Pernes (born 1940 in Switzerland) is an Emeritus Professor at the Department of Philosophy of the University of Toronto which he joined in 1966. He is best known for his work in philosophy of emotions, and has also made contributions to philosophy of mind and philosophy of biology. He was elected a fellow of the Royal Society of Canada in 2005. De Sousa possesses both UK and Canadian citizenship. Educated in Switzerland and England, he took his B.A. at New College, Oxford University in 1962, and his Ph.D. at Princeton University in 1966. He has contributed to and is frequently cited in the Stanford Encyclopedia of Philosophy.

Andrzej Dąbrowski: Philosophy is the systematic study of world, human being, knowledge, mind, thinking, rationality. The affective life is not the most important subject of philosophy. First existentialism and philosophical anthropology have established philosophical interests in human emotions. You are working in philosophy of emotions. How did you first become interested in this subject? Why emotions? And what are the differences between the modern emotion philosophy and Max Scheler and Jean-Paul Sartre's first ideas?

Ronald de Sousa: I'm not sure that existentialism and philosophical anthropology can be credited with "first establishing" the philosophical interest in emotions. Historical figures such Plato, Aristotle, Spinoza, Descartes, Hume and William James, to name just some of the major names in the philosophical canon, were all very interested in emotions. But it is certainly true that in analytic philosophy as well as in psychology emotion theory suffered a hiatus for most of the first half of the 20th century. In psychology, that can probably be attributed to the dominant influence of behaviorism; in philosophy, to the primacy of linguistic philosophy. Emotions seemed just too complicated to tackle from the perspectives of those research programs. The study of emotions requires multiple points of view and emotions themselves touch on many areas of philosophy. My own preferred view of philosophy is as a license to be a dilettante; so it is perhaps precisely because I was reluctant to specialize in any narrow branch of philosophy that I became interested in emotions. For they seemed to constitute a domain on which many of my interests could converge: my original specialization of the philosophy of language (in which I wrote my doctoral thesis), and my interests in epistemology, in ethics, in aesthetics, in literature and, of course, in the philosophy of mind. Emotions raise problems in all these areas. For the philosopher of language, the different

senses in which emotions have "objects" demand that we understand what Donald Davidson called the "variable polyadicity" of events in general and mental representational states in particular. Whether emotions bring us knowledge of the external world, of ourselves, or of both, is a question that concerns both epistemology and the philosophy of mind. Emotions play a crucial role – though certainly not an uncontroversial one – in ethics and aesthetics. And since literature is the richest source of available discourse about emotions, that provides a philosopher of the emotions with a good excuse to spend time reading fiction and poetry. Max Scheler's chief contribution to the theory of emotion was to emphasize their role in our apprehension of values. This highlights their relevance to both epistemology and the theory of value – both ethical and aesthetic. It is probably fair to say that contemporary philosophy of emotion, my own in particular, is more hesitant to posit an objective realm of value to which emotions provide epistemic access; I myself am more inclined to think of values as projections of emotions; in practice, however, I'm not sure that this makes a great deal of difference. As for Sartre, I associate him with two ideas, both of which are intriguing but ultimately unconvincing. The first, which was much emphasized by Robert Solomon in his early work on emotions, is that emotional responses are a form of magical thinking, producing unreal solutions to insoluble problems. The second is that we are as much responsible for our emotions as for any voluntary choice. What is wrong with the first idea is apparent as soon as we think of the importance of emotions to us as evolved animals: the biological perspective makes it immediately implausible that a phenomenon as important to our lives as emotion should be devoid of efficient practical utility. Although emotions cause us much trouble, in the main, they represent swift solutions to certain authentic problems that arise from ordinary human situations. As for the other idea promoted by Sartre, it certainly serves as a useful reminder of the fact that we need not be helpless victims of purely passive emotions: emotions can be thought through, emotional responses are often the result of ideological commitments and prejudices which philosophical thinking can overcome.

Andrzej Dąbrowski: Above you mentioned the philosophy of mind. The concept of modularity has played an important role in recent debates in the philosophy of mind. Are emotions modular?

Ronald de Sousa: Emotions 'mimic' the modularity of perception. That is what I meant by saying that they spare us the Frame Problem. But whether they ARE modular is another and rather complicated question. Since I am on record with a paper called *Against Emotional Modularity* (which goes into this in more detail than I can rehearse here), I have to watch what I say. Or at least re-read what I did say... The word is used in a number of ways. In Fodor's book on *The Modularity of Mind* it had a very specific meaning, and applied to the informational encapsulation (the inability to respond to other channels of information, including beliefs) of our senses. Thus we do not dispel the Muller-Leyer illusion even when we know it to be an illusion. It is in this sense that emotions mimic modularity. They are not really modular in the original sense, because unlike sight or hearing, we are not stuck with the separation of channels in perpetuity. When the emotion has subsided, it can allow the person to see other points of view. (Providing the person is willing to allow it: a question of emotional maturity, or perhaps "intelligence", though that is another controverted term). So, emotions are not literally modular in that sense. In the more recent literature (including most of the papers in a volume edited by Faucher and Tappolet, *The Modularity of Emotions*, 2006, to which the article I mentioned above was my contribution), modularity is interpreted as some sort of evolutionarily programmed unit of complex responses designed to deal with some stock situation type. It applies to the "basic" emotions, for those who recognize such things. In this guise, emotions belong to the fast, automatic, low-effort, inexplicit mode of processing identified in the 'two track mind' or what is usually called 'dual processing', as opposed to the deliberate, usually language-based, attention-hungry, 'second track' processing. In that sense, modularity can be understood as implicating something like one of Griffiths' "affect programs", not merely selected but selected for by natural selection; or else it can identify emotions as clusters of

more basic responses varying in a number of dimensions of concern. For my part, I like to stress another aspect of this question – in effect, another sense of modularity. That is the rigidity of the categories in terms of which we understand and classify emotions. The need for such simplified categorization arises from our need to talk about emotions. But the categories in question are far from adequate to render the subjective experience of emotion. That is why some of our emotions are sometimes said to be 'ineffable'. More subtle emotional experiences are vital to all aesthetic experience, and it seems to me vitally important, in order to allow for our full potential for apprehending the aesthetic aspect of our real everyday life, that we should try to free the reality of our multifarious emotional experience from the constraints of rigid categorization.

Andrzej Dąbrowski: Whether human emotions have changed historically? Many scientists (philosophers also like P. Griffths) believe that historical studies contribute nothing to the study of emotions. Do you agree with those researchers?

Ronald de Sousa: I do think emotions have changed historically as ideologies governing interpersonal relations, both public and private, have changed. Some of the virtues identified by Aristotle, in terms of their emotional stance to such things as honor, derived from specific forms of aristocratic values. Without subscribing to Nietzsche's conception of a complete inversion of those aristocratic values under the influence of Christianity, the general lesson that we can retain from Nietzsche is that the genealogy of morals affects their essence, and that emotional stances towards values promoted by warrior societies, say, or in the context of capitalist economies, can be very different from one another. I am not sure what psychologists you allude to when you suggest that many of them think historical studies have nothing to contribute to the study of emotions. The International Society for Research in Emotion (ISRE), and the largest collaborative research institution I know of, the Centre for the Interdisciplinary Study of Affective Science (CISA, in Geneva), for example, both make a point of including historical studies in their overall research plans. My interpretation of Paul Griffiths' view is that he thinks only biologically innate "affective programs", of the sort associated with short lists of basic emotions, can be regarded as forming psychological natural kinds exemplifying homologies with dispositions observed in other primates or mammals; but as I understand him he does not deny that many of the phenomena we call emotions may form complex, socially constructed clusters of response dispositions that have their own psychological reality and sociological significance. That is certainly my view. Moreover confidence in the existence of "basic emotions" has waned in recent years; a more thoroughly constructionist view might appeal to even more basic processes in the construction of what I have called "paradigm scenarios" which define different specific emotions. But that would render them all the more susceptible to alteration on the basis of changes in ideological assumption about what is "normal" or "expected" in a given social context.

Andrzej Dąbrowski: Are cognitions and emotions one or two different worlds? The first books on cognitive psychology and cognitive sciences as such had no chapters on emotions. The situation has recently changed. There are a huge number of books devoted to close connections between emotions and cognitions. Even there was a cognition-emotion debate. Are there any significant philosophical results of this debate?

Ronald de Sousa: The cognitive science revolution, perhaps beginning with Chomsky's notorious review of Skinner's book on *Verbal Behaviour*, was certainly more interested, at first, in belief formation, inference strategies, and memory. But it soon became clear to pretty much every one concerned that emotions played a crucial part in those cognitive processes, both for good and ill. Several scholars have noted the existence of specifically epistemic emotions such as interest, doubt, and the feeling of certainty. Again, these are not really new, since they figure prominently in both Plato's *Meno* and *The Meditations* of Descartes. The precise way in which emotions contribute to

cognition is indeed, as you say, an active area of current investigation. A very interesting recent book by Michael Brady, for example, argues that emotions do not directly afford a knowledge of axiological facts, but rather provide us with motivation for looking into evidence for such facts. Thus fear does not give us evidence for thinking something is dangerous, but does provide a motive for paying attention to what might be a threat. The intertwining of emotion and cognition in current philosophy and psychology is itself, in my view, an important philosophical result. One of the manifestation of the interaction of cognition and emotion is the growing interest shown in "affective computing" by scientists working in Artificial Intelligence. From the philosophical point of view, I would stress the crucial role of emotions in complementing mere logical inference in two essential ways: first, our emotional states set or at least constrain our goals. Emotions therefore set the stage for any rational calculation. Second, by narrowing the field of our attention to immediately relevant concerns, emotions, I have suggested, largely save us from the "philosopher's frame problem". That problem, originally identified as a difficulty for computer science, lies in the importance of knowing what to ignore among the myriad possible consequences of any action. Since it is in practice impossible to give adequate consideration to all of those possible consequences, our narrowing of attentional focus under the influence of emotion allows us to concentrate on only subsets of potential consequences of whatever we undertake.

Andrzej Dąbrowski: Physiological (or somatic) theories of emotion (W. James, R. Zajonc, A. Damasio, J.E. LeDoux, J. Prinz) claim that bodily responses are essential to emotions. Defenders of cognitive theories of emotion (W. Lyons, R. Solomon, M. Nussbaum, J. Neu) argue that cognitive elements – beliefs, judgments, evaluations, or thoughts – are essential to emotions. Is there a chance to reconcile (to unite) these two different perspectives?

Ronald de Sousa: I don't really see much of a conflict. Those who are interested in the physiology of emotions focus on the so-called basic emotions that have homologues in other primates, such as fear, anger, joy and sadness. Insofar as those involve rather primitive responses, they are not, in my sense, yet fully emotions, but rather proto-emotions. Most emotion researchers acknowledge that a full-fledged emotion unites cognitive, experiential, physiological, expressive and conative (or "action tendency") components. The complexity generated by that combination, and the modulation of the physiological elements by thoughts, make for elaborate human emotions that are characterized by a dramatic structure ("paradigm scenarios"). Those thought-driven emotions are, to some degree, susceptible to "re-gestalting": meaning that one can construe a situation in different ways by thinking about it in a new way. Thus, for example, anger might be turned to compassion if one construes offensive behavior as resulting from discomfort or ineptitude rather than hostility. In her book *The Managed Heart*, Arlie Hochschild relates that flight attendants were trained to cope with drunken and irate passengers by gestalting them as babies having a tantrum. The more thought-impregnated emotions are obviously the most interesting from the philosophical point of view. But there is no problem of reconciliation with more basic emotions: unlike Griffiths, I regard all of them as emotions, but I agree with him in seeing them as belonging to significantly different classes.

Andrzej Dąbrowski: Do you think your theory (or proto-theory?) on the rationality of emotions presented first in *The Rationality of Emotion* (1987) seriously evolved?

Ronald de Sousa: From my 1987 work, I still retain a number of convictions. First, that in individual human lives each of us forms a repertoire of emotions that is to some extent idiosyncratic. We each have our own "emotional idiolect". These idiolects, like literal idiolects in language, are constructed relatively early in life, out of a stock of basic human capacities to respond; these are integrated into paradigm scenarios in terms of which we interpret subsequent episodes of our emotional life.

Art, literature, and philosophical reflection are all able to modify those scenarios to some extent; but the tendency to exhibit repetitive patterns that characterizes most people's emotional biography bears witness to the tenacity of the original scenarios as well as underlying temperament. I also still think that the idea of pure rationality makes no sense without the double role of emotions in setting the questions to which we seek rational answers, and the narrowing of our attention so as to constrain the range of relevant considerations that we bring to the search for such answers. Both roles, of course, can work against us on occasion (all virtue is vice, though not vice versa!). The counter-productive nature of many familiar emotions (anger, jealousy, spite) illustrates the way in which the "blinders" imposed by our emotional states can cause us to miss what from a broader point of view might turn out to be the most important things. In my work since *The Rationality of Emotion*, the main change in my perspective is that I have come to think that many of the problems raised by the theory of emotions are illuminated by the relation between the two "levels of processing" mentioned above. Emotions straddle the intuitive and the analytic tracks, which is why it is worth regarding them as having functional aspects in common despite the difference between Griffiths' strictly defined "affect programs" and what he calls the "cognitive emotions". I have not been concerned so much with modifying those ideas, but in seeing how they can be applied to a range of issues of more general philosophical interest. On the theoretical level, these include a more elaborate notion of "generic truth" which can make sense of the ascription of truth to emotions, and I have tried to take a broader view of what we may count as 'cognitive' than the once dominant emphasis on propositional objects. On the practical implications of emotion theory, I have come to regard emotions as crucially significant for ethics and aesthetics. I have also been interested in more specific problems arising from a philosophical perspective on the pursuit of happiness, on love, and on the possibility of taking a more richly aesthetic attitude to our own emotional lives. These issues form some of the topics in *Emotional Truth* (2011). More specifically, I have recently focused on love, which I regard not as an emotion like others but as a syndrome. By that term, I refer to a peculiar feature that distinguishes love from other emotions (though in some cases it is a matter of degree), namely that any emotion at all might be, in some context, a manifestation of love. The role played by ideology in defining what we think of as appropriate feelings, behaviour and attitudes in love is the focus of my most recent book, *Love: A Very Short Introduction*, which takes aim at some of the absurdities and hypocrisies that pervade most people's endorsement of the ideal of monogamous marriage.

Andrzej Dąbrowski: It's an interesting topic. There were many attempts to find the equation of love… Maybe did you do it? What the value of love is?

Ronald de Sousa: I'm not quite sure what you have in mind by the 'equation' of love. As I mentioned, love is not so much an emotion as a 'syndrome'. (This is a view that has now been elaborated in a superb paper by Arina Pismenny and Jesse Prinz, to appear in an Öxford Companion to Love and Sex edited by Chris Grau and Aaron Smuts.) Perhaps, indeed, love should be said to be several different and perhaps even incompatible syndromes. What I find most striking when thinking about love is that although the various sorts of human affiliation that go by that name obviously have deep biological roots, they are strongly influenced by expectations that stem from ideology. Current ideology in the Western world insists, for example, that passion, sexuality, companionship and intimacy are all essential to love, and that all of these must form a psychological conduit to monogamous marriage. But in reality those ways of connecting are very different in their character and particularly in their typical duration. So in real life they fall out of step, leading to bitter disappointments and resentments rooted in unrealistic expectations. In a modern world where we no longer have to associate sex with reproduction, and where rigid social rules no longer require that everybody's life should follow one of a few set patterns of affiliation, the ideal of monogamous marriage (which at best was 'honoured more in the breach than in the observance') could be and is being re-thought more realistically. Such re-thinking, and the

experiments in living carried out by polyamorous communities, aim at allowing those different components (passion, companionship, and specific goals such as the raising of children) to be pursued in relative independence of one another. They are premised on the curiously unpopular notion that pleasure is not intrinsically sinful, and that emotional authenticity and fidelity need not be defined in terms of sexual or affectionate exclusiveness. In this way, philosophical investigations of love bring the practice of philosophy back to its ancient connection with the quest for better ways to be human.