Studia Semiotyczne—English Supplement, t. XXX (2019), s. 5–19
ISSN 0137-6608, e-ISSN 2544-073X
DOI: 10.26333/stsen.xxx.01
Article

Piotr Wilkin *

NATURALIZED REPRESENTATIONS— A USEFUL GOAL OR A USEFUL FICTION?

Summary: One of the key concepts of naturalized epistemology as well as the cognitive sciences that stem from it is the naturalized concept of mental representation. Within this naturalized concept, many attempts have been made to unify (for humans as well as for other living organisms) the notion of representation error. This text makes an attempt to argue against the adequacy of using a naturalized concept of representation error as well as casts doubt on the wide program of naturalizing concepts related to human conceptuality.

Keywords: mental representations, representation error, naturalization.

1. Introduction

In philosophy of mind, the naturalistic approach is becoming more and more popular; it is also a constitutive approach, if not for cognitive science, then at least for some branches of it. One of the fundamental concepts of cognitive science which is often naturalized is the concept of *cognitive representation*. One of the most popular approaches to naturalizing

* University of Warsaw, Faculty of Philosophy and Sociology. E-mail: ilintar@gmail.com. ORCID: 0000-0003-4714-5269.

representations is that of Dretske, which ties representations with certain natural functions of biological organisms (Dretske, 1986). Dretske's solution has two major strengths. From a philosophical perspective, it can tackle many problems that informational or correlational approaches to representations have had problems with (one of the key issues that the abovementioned approaches faced was that of the ubiquity of representations: smoke is often correlated with fire, but is smoke a cognitive representation for fire? Does fire have cognitive representations?). From a methodological perspective, it can provide a universal take on representations for many classes of organisms, which lets us obtain empirical data about representations from studies on simple animal organisms or even bacteria. It also gives us a clean transition from animal cognition to human cognition—as such, it is a strong counter to all dualistic approaches to cognition.

The main aim of this text is to undermine the universality of cognitive representations as seen by Dretske and his successors. Within this text, we shall be assuming the representational approach in cognitive science. We shall not deal with issues of anti-representationism because, while the question of whether cognitive representations are a valid element of the cognitive science landscape is no doubt interesting and valid, it is out of scope here and would only muddle the main points of the argumentation. Therefore, when we discuss the various pros and cons of the naturalistic approach to representations, remember we do so only under the assumption that representations themselves are useful and significant.

2. Naturalized Representations and Misrepresentation

An important feature that Dretske and many of his successors (e.g. Millikan) emphasize in their solution is the ability to analyze representations with respect to their correctness—in particular, to show criteria of *misrepresentation*. In Dretske's approach, a representation of property X is correct when (in normal conditions) it works according to its function, that being indicating the presence of X. A key aspect of this approach is the concept of function, or more precisely, a specific type of function: a *natural* function. If we want to naturalize representations, we cannot simply use a general notion of function, since there are many possible classes of functions and we risk a problem we seem to have just averted—that of the ubiquity of representations. Therefore, we restrict ourselves to the class of natural functions, which are those that guarantee

evolutionary success. Dretske gives the example of bacteria which have a natural magnetic indicator of north, which allows them to move towards less oxidized waters (a surplus of oxygen is deadly for the bacteria). The same bacteria, when moved to the southern hemisphere, will die, since their magnetic sensor will incorrectly direct them away from the pole, towards deadly waters filled with oxygen.

It's worth noting that even Dretske when providing the example cautions against using it as an instant case of naturalized representation. This is due to the fact that the bacteria's indicator simply points them towards magnetic north, not towards oxygen itself. Even if we assumed an evolutionary criteria for selecting natural functions, it's hard to explain why the mechanism actually indicates the presence of oxygen and not the presence of the magnetic north, with the magnetic north being the environment in which the bacteria normally thrive. In other words, the example with the bacteria moved to the southern hemisphere might not be one of misrepresentation, but one of abnormal world conditions. Even if a given representation could be evoked by one of many independent mechanisms, it still wouldn't be enough to tell us that it's a representation of the property that we desire, rather than a disjunction of the immediate triggers (for example, if the bacteria had a light indicator together with the magnetic one, we could still say that they have a representation of the property Light-or-North and not of the property Oxygen). Dretske claims that only organisms that have a set of independent representation-controlling mechanisms and are able to switch them on during their lifecycle have the capacity to misrepresent. In other words, it is only when an organism has a representation of property X which, during various phases of learning, is evoked by different stimuli originating from X (but having the common feature of being caused by X), that one can talk of a representation that can misrepresent.

It should be clear now that, contrary to the promising start (of bacteria having representations), to talk about representation in the Dretskian sense we need more complicated organisms than bacteria. However, it's still a notion of representation that is scientifically attractive—most animals, even the very simple ones, have some capacity to adapt, so we could obtain a lot of empirical examples for representations and misrepresentation from the rich world of animal behavior.

Moreover, Millikan's solution (1995), which was an answer to Dretske, manages to solve even the problem of bacterial misrepresentation. Millikan solves the problem of vagueness present in Dretske's approach by

assuming that for a given organism, its *proper* function (as Millikan calls her extension of the notion of natural function, see [Millikan, 1987]) is indicating a property that is required for the organism to survive and reproduce (in other words, to achieve evolutionary success). In this approach, the bacteria have the representation of Oxygen (instead of Light-or-North) because it's the former that is required for their survival—the latter is strictly accidental.

All of the approaches mentioned are very well developed and show promise when it comes to studying representations in animals. However, do they actually make it easier for us to understand representation in humans?

3. Human Errors and Humans' Natural Function

Let's now look at a typical case of misrepresentation that happens during the human language acquisition process. A child looks at a ripe, red apple, reaches for it and says "tomato"—with the clear intention of eating the apple as a tomato. She hasn't yet learnt that there are other fruits of similar size, shape and color as the tomatoes that she's observed before.

There are two possible explanations for the situation described. One is that the child simply has a wrong representation of tomatoes, i.e. that she has a representation of tomatoes, but it's not *the correct* representation. Another explanation is that the child does have a representation of tomatoes, but it didn't work correctly that time. Let us call the first explanation that of a *general error* and the latter one—a *particular error*. In both cases we now want to ask the question—how would we naturalize such a notion of representation?

Note that if we want to talk about a functional approach to naturalizing representations (whether it be Dretske's approach or Millikan's approach), we want to talk about a *biological* function—one that we could single out in both humans and in simpler organisms (although, as we mentioned before, Dretske seems to believe that to properly determine a representational function, you need a certain level of biological complexity). This function should be somehow connected with the evolutionary (or, more directly, reproductive) success of the organism. It's worth noting here that Millikan speaks about "representation reproduction" instead of "organism reproduction", which opens up the possibility of understanding it in non-biological terms. However, most of Millikan's own research

pertains to biological reproductive success, so we shall assume that is the dominant understanding for now. We shall tackle the other possibility later in the text.

Now let us consider this: can we actually find a biological function of the child's organism that would determine that the proper representation of the tomato should be one of a tomato and not one of a red apple? Before we actually move on to try to answer this question, it's important to understand that a potentially higher level of complexity (and thus, a more complex function) would not be problematic here. If the difference between human representations and simpler organism representations were just one of degree, that would not be a major difficulty for the naturalized theory of representations. One could rightfully hold the view that the level of complexity of a natural function that realizes a given representation is proportional to the complexity of the organism itself. In such a case we should not find it surprising that a human's natural function is much more complex that one of an amoeba or bacteria. Furthermore, for Dretske such a situation would actually be a pro rather than a con—to talk about natural functions and avoid ambiguity, we need a complex system that makes certain choices based on more than one criterion.

Let us therefore assume that we actually managed to discover a natural function that corresponds to the child's representation of a tomato. Let's also assume that the function actually explains the representational error that the child makes when calling a red apple a tomato (or when it reaches for the apple with the intention of eating it as a tomato, to avoid linguistic criteria). Can any such function really be a *natural* function? Of course, we don't want to define the class of natural functions so widely that it loses its intuitive meaning—after all, we wanted to restrict the class of functions to natural functions precisely to avoid some problems with naturalizing representations. Therefore, we want to relate the natural function to the organism's survival. However, it seems that no credible explanation of that sort can be actually found, as I shall now try to show.

Starting with the most direct approach, a proponent of the naturalistic approach might claim that the ability to distinguish apples from tomatoes is critical for survival. For example, take a child that has a deadly allergy to apples (but not to tomatoes); a misrepresentation might turn out to be fatal (e.g. if the child reaches for the apple and eats it before her parents manage to react). This type of analysis might seem promising, since it only deals with biological criteria. Also, one can provide less convoluted examples where distinguishing one organism from another is critical for

avoiding poisoning. Take, for example, the parasol mushroom and the death cap. This example is even better in that it deals with general mechanisms (the death cap is poisonous for humans as a species rather than just for individuals), so it's easier to claim that such a function would be natural in the sense that it correlates with the evolutionary success of the species.

However, our language is too rich to permit such an analysis for all concepts, so this way is doomed to fail sooner or later. We are not be able to find a direct evolutionary function for every single concept, although we can probably find a scenario in which misrepresenting a concept results in an organism's death. However, inventing scenarios is not a good argumentative road—for every scenario one can find a counter-scenario in which having the allegedly incorrect representation ensures success (for example, a scenario taken almost out of Grimms' fairy tales, where Hansel brings a death cap home and feeds it to the witch, who was just about to cook him in the oven). To justify naturalizing a representation, we must have a universal function—one that can be explained on the level of the entire species, not just single organisms. In the literature, one can indeed find many guidelines on how to correctly describe natural functions so that they are indeed natural (i.e. so that they can be properly naturalized; Millikan's analysis is a good example of this).

We shall drop this line of enquiry now mostly because a criticism of a specific approach to natural functions will not be a definitive rebuttal to the idea of naturalizing representations in general. Even if we cannot tell what the evolutionary advantage is of having the representation of a convertible distinct from the representation of a station wagon, the very fact this distinction exists might suggest that it somehow contributes to our survival. The proponent of the naturalization approach to representations might say that we might not know the exact natural function corresponding to more complex concepts, but it is the task of empirical studies to find and describe it.

Therefore, the objection to naturalizing representation must have a more fundamental nature. The question that will lead us to that objection will be the following: how do we assert misrepresentation in humans? What makes us say that someone misrepresents (in both the general and the particular sense) some class of objects (for example tomatoes or convertibles)? And finally: how do we learn to make the relevant distinctions? The answers to those questions will hopefully cast doubt on the validity of the naturalization approach for human representations.

4. Humans and the Natural Error

Humans are a very specific species in the animal kingdom in that a lot of their representations have a social source—they are created and changed not only in response to stimuli connected to the represented object, but also (or, one could claim, mainly) in the process of socialization. This process of socialization is special even among animals who do have a process of socialization—many of our representations are created with the help of language. I do not want to tackle the topic of the relation between social interactions and cognitive representations in this text, as it would be widely out of scope. This is not only true for representations on the personal level (as per the personal/subpersonal distinction due to [Dennett, 1969]), where the relation to language is quite obvious, but also on the subpersonal level. For example, take the notion of attractiveness—it would seem that the representation of a "potentially attractive mate" is something that we share with the rest of the animal kingdom. However, a short historical enquiry is sufficient to discover that the socially prevalent criteria for attractiveness have changed much more often than would be credible for an evolutionary explanation.

Therefore, even if we restrict ourselves to subpersonal representations, we cannot guarantee that they were not formed without the presence of social factors (unless we are talking about inborn representations; as we shall further discuss, the origin of representations is a quite important differentiating factor). Moreover, if an important feature of representations is supposed to be their durability, then the social explanation seems to be more plausible than the evolutionary one—the example of attractiveness suggests that social interactions are more important in determining representations than purely evolutionary factors.

Let us come back to the definition of misrepresentation formulated earlier and fill in some specific objects for the variables: Athanasius is misrepresenting the parasol mushroom if his representation (parasol mushroom) leads him to collect a death cap in the forest (at least in the general case; in the particular case, he mistakenly takes a death cap to be a parasol mushroom). It would seem that, due to the direct biological effects, this would be a paradigmatic case of naturalized representations—a misrepresentation leads, after all, to an organism's death. However, is this really a case of misrepresentation? More specifically: is the correct representation of the parasol mushroom really what we understand by the linguistic concept "parasol mushroom"? After all, we can

imagine a case where the representation itself does not change, but the inclination to eat the mushroom does. This counter-argument could be rebutted by asserting that having distinct representations for a death cap and a parasol mushroom is evolutionarily superior to having just a representation of a parasol mushroom as an inedible one (again: imagine a scenario in which we have a tribe living in a forest where their only potential food sources are either death caps or parasol mushrooms). However, we can also imagine that the very same tribe represents all those mushrooms as parasol mushrooms—just with the distinction that the greener ones are poisonous, while the more brown-tinted ones are edible. In other words, they ascribe the edibility criteria to certain states of a given type of organism rather than to a distinct type of organism (a real-life case of such a distinction is the mushroom commonly known as a "puffball", whose early forms are actually edible).

Perhaps by now an analogy to a famous argument from philosophy of language—Wittgenstein's criticism of "private language errors" (Wittgenstein, 1953) later expanded upon by Kripke (1982). This analogy does not seem to be accidental—I believe that talking about misrepresentation in the context of our cognitive representations (other than the native ones) in the same way we talk about misrepresentation in the case of simpler organisms in relation to their natural functions is a dead end.

Most arguments that Wittgenstein (and later Kripke) use to refute the possibility of a naturalized conceptual error can be adapted to the case of misrepresentation. Take for example the abovementioned case with death caps and parasol mushrooms. Even Wittgenstein's original example (recall that Wittgenstein, and after him Kripke, claimed that we can't determine whether someone, when talking about addition, or the use of the plus sign, really means "plus" instead of "quus", where quus is different from plus in that it behaves differently in very specific conditions which do not obtain in the given case) could be possibly used (if not for the fact that the concepts used are highly abstract, which makes finding the corresponding representations difficult). Note that the gist of the argument is the same in both cases. Wittgenstein (and Kripke after him) says the following: using purely objective criteria, we are not able to determine, which of the two descriptions of the concept is the correct one—similarly, we cannot determine which of the two descriptions of cognitive representations is the correct one other than rationalizing it *ad hoc* after the fact ("weird parasol mushroom" vs "parasol mushroom / death cap").

This argument can also be used in two ways. If our misrepresentation is understood as a general error (having an incorrect representation), the question becomes: how do we determine the correctness of the representation (in other words, how do we select one specific proper function over all others). If we understand it as a particular error instead, meaning a representation is used incorrectly, then we can ask, after Wittgenstein: how do we know that it was an error and not an exception specified in the rule?

However, the naturalization proponents are in a better place than Wittgenstein's opponents in the rule-based concept usage debate—they can still fall back on the concept of natural functions and defend our representations by relating them to biologically proper functions. However, that route seems a dead end as well—even in the case of the death cap, which seems well-suited for naturalization, it's hard to show a clear advantage of the double representation version over the "weird parasol mushroom" version.

To the fundamental arguments one can add empirical arguments as well. Even if we could agree that the idea of naturalized misrepresentations can be defended on theoretical grounds, it would be hard to defend the claim that our cognitive representations are really formed in the way that this idea describes and that we diagnose misrepresentations based on evolutionary consequences. The richness of our conceptual system and, in consequence, of our representational system (since we have already noted that most of our representations have linguistic correlates) is too big compared to the period of potential evolutionary change for this explanation to actually be plausible. One could defend this type of theory when it comes to bees, whose communication does seem to be evolutionarily coded, but in the case of humans, our systems of communication are too short-lived for the evolutionary context to be relevant.

One could claim that the naturalized approach to representations is nevertheless correct also in humans and that the correct representations are those that realize some natural functions (or proper functions, if we prefer Millikan's terminology) and that the social agreement or disagreement towards concept use has no bearing on the notion of misrepresentation. However, that type of approach requires accepting one of the following assumptions—each of which seems problematic for its own reasons.

First of all, we can assume that linguistic concepts and cognitive representations are not directly correlated—that concepts are not rooted in cognitive representations. In text, we tacitly assume that such a ground-

ing exists, but of course a negation of such a claim can be imagined. In its radical version (concepts have completely no connection whatsoever to cognitive representations) it seems completely implausible for anyone who wants to respect the scientific foundations of cognitive science, including the empirical results of developmental psychology. However, one could opt for a weaker version of the negation—for example, accepting the grounding on the level of types (cognitive representations overall are grounded in cognitive representations), but refusing it on the level of particulars (specific concepts are not grounded in specific cognitive representations). It's hard to see, however, how this type of negation helps alleviate any of the problems mentioned above.

A second option is to assume that the current state of language is not an adequate measurement of the correctness of cognitive representations. Such a solution requires assuming a Leibnitzian view of a perfect language which would best suit our evolutionary needs and which would be the one according to which we should judge representations. However, metaphysical problems notwithstanding, there is a fundamental problem here: is such a solution actually naturalistic? How do we scientifically verify the correctness of representations with a postulated ideal language best suited for our evolutionary success?

The third option is to go holistic—instead of evaluating particular representations as correlated with particular concepts, we evaluate representations based on their role in an entire linguistic system. However, this type of holism only masks the problem—since now we are no longer suited to judge particular representations, instead, we need to evaluate an entire system which the given representation is tied to. This does not seem like a naturalistic approach at all and, moreover, seems to direct us towards an antirepresentational approach which we agreed not to discuss in the introduction.

Besides the problems with the abovementioned three options, the solution that ignores the linguistic side of cognitive representation does not seem to be well reflected in empirical studies. It ignores our actual mechanisms of evaluating representations as erroneous in favor of an idealized concept of misrepresentation that is different from what is commonly understood as a representational error. This seems to be similar to certain solutions within the semantic contextualism debate which, to avoid contextual dependence for some propositions, gives them a literal meaning that comes out as false in virtually all circumstances where we would assert them as true and explains this assertion using a system of complex

implicatures. These types of solutions, while formally correct, do seem dubious in terms of their explanatory power.

5. Representations and Society

Since we have provided the arguments against a fully naturalized solution to representations, a further task remains: to provide an alternative solution to full naturalization. We explicitly refused the antirepresentational solution at the start, so now we are tasked with providing another positive option.

Let us employ a classic tool of analytical philosophy: linguistic use case analysis. When do we say that someone is misrepresenting an object? In our case: when do we say that Athanasius is misrepresenting the parasol mushroom?

We said that, similarly to Wittgenstein's solution, the social consensus seems important here—Athanasius is misrepresenting the parasol mushroom if his representation does not match what society has established as the proper representation. Should we, however, understand this as pure social consent, i.e. Athanasius has the correct representation if and only if society agrees that his representation is correct?

This solution has many benefits and simplifies a lot of matters when it comes to the conceptual side of representations. It's also quite antiscientific—under this approach, we would have to drop all attempts to reduce representations to objects described by empirical sciences. However, that by itself is not a critical problem—after all, we consider many sociological phenomena to be fully emergent, and we don't posit their reduction to the biological layer. This solution has another problem, however—it makes it impossible for us to positively resolve the "individual vs society" dilemma.

Let's consider an archetypal story of a brilliant lone scientist. In this story, an individual comes across a breakthrough discovery, she's shunned by the majority of the scientific community, but then we discover she was right all along. We might try to apply this scenario for example to the real-life historical discussion regarding black holes (assuming that the discussion in this case really was of the "individual vs society" type, since in reality those cases rarely happen in their pure form). If the representation is decided purely by society's consensus, there is no possible case in which the scientist is actually right—she will never be able to prove the correctness of her representation. Even a version of the scenario in which

she gradually convinces the community of her approach isn't well described in this case—since she is constantly wrong when she does the convincing and only starts being right once she's actually convinced the majority. This description seems wrong—we would surely prefer to claim that the scientist was right all along and the majority had the erroneous representation. How can we save this intuitive description?

The best approach seems to be to combine the functionalist approach towards representations, which has a very respectable intellectual history and has developed many useful and precise concepts, with the social approach. To do that, we only have to abandon... naturalization. We would still want to say that having a certain representation is realizing a certain function—what changes is the nature of that function. It would no longer be a natural function—it would instead be a socially-regulated function, in a manner similar to how Wittgenstein understood the way language-meaning rules are governed by society, namely that the meaning of a word is what the linguistic society currently enforces as its meaning.

In such a theory, the scientists who single-handedly maintains the existence of black holes might still be correct—as long as his representational function follows the rules that are enforced by the society. He might still differ with the rest of the society as to what exactly corresponds to the object of those representations (in the same way that I can agree with others that "the fastest man in the world" means "the person who just got the fastest time in the men's 100m sprint at the Olympics", but due to a lack of information I could be convinced that this refers to Justin Gatlin (since that's what the first reports might have indicated), while a later analysis of the photo-finish showed that the fastest one was actually Usain Bolt. This is a bit similar to how Kripke describes necessary truths that are known *a posteriori*—from the fact that the society agrees (explicitly or implicitly) on the meaning postulates regarding a certain concept (in our solution, that would mean they agree with respect to the representation function that realizes the concept), it does not follow that they have knowledge about all true propositions which the concept is part of, as some of those propositions can be only known by empirical research and not just by conceptual analysis. Our token scientist might therefore agree with other scientists on the ostensive definition of black holes (e.g. "black hole" = "that which constitutes the center of known galaxies"), while disagreeing on the essential physical properties of those objects (e.g. their ability to capture light or alter the gravitational field).

If someone still remains unconvinced by the analysis above, here's an alternative argument showing that the naturalistic approach to human representation is not plausible—one grounded in the results of cognitive sciences (some elements of that argumentation can already be found above). Let us consider what is the subject of inquiry of cognitive science when it comes to humans and compare it to the subject of inquiry when it comes to animals and other living organisms. Assuming that we can provide a common metaphysical description of representations in both cases, we have to ask how the respective representations are formed. It seems that while in the case of animals almost all representations are inbred and have an evolutionary source, that's not the case with humans—our representations, judging e.g. from their linguistic correlates, seem to be contingent and have a social ground. If we aim at providing a common characteristic for human and non-human representations by using natural functions, we obtain easy empirical data, but the data will not necessarily be adequate for human cases (since it's hard to provide a credible evolutionary explanation for humans in the same way that it's possible for simpler organisms). Therefore, this solution picks data accessibility from the accessibility / credibility pair, which of course is better if we want an easy influx of superficially convincing examples, but raises concerns from a methodological perspective.

Let us come back to the solution offered by Millikan that we mentioned earlier and see if we can recover the naturalistic approach by assuming a broader approach to the concept of representation reproduction. Can we understand "reproduction" in a social way here and assume that representations are persistent if they are socially reproducible? Of course, we could do that, but it seems that for a naturalist that would actually be a pyrrhic victory. While it seems quite obvious that biology is the science that is suitable for describing evolutionarily stable mechanisms, it would be quite a stretch to assume that biology is likewise suitable for describing socially stable mechanisms (such as linguistic concepts, language systems or cultural norms). We can refer to sociology, psychology or economics to fuel us with theories that handle those concepts, however, it will be hard to assume that such a solution will still be naturalistic. Usually, by "naturalistic" we understand a reduction to the results of natural science—we would either have to use an unusually broad notion of naturalization or assume that sociology, psychology or economics can be completely naturalized—which would be defending the naturalization

of representations by assuming an even stronger and more controversial claim.

Therefore, if we want to defend the functionalist approach to representations, we shall have to modify many assumptions that usually underlie this approach. Most importantly, we shall have to get rid of the "natural" teleology and the corresponding approach to natural functions which ties them with evolutionary stability (an approach common in Millikan's writing). An in-depth analysis of the argumentation provided in this paper allows us to go even further—we should get rid of teleology completely. An analysis that takes into account both representation data from human and non-human examples suggests that we might be better off by instead considering proper functions with respect to their causes instead of their purposes. This would also help explain the teleological approach present in the research on representations—in the evolutionary approach, cause and purpose are almost indistinguishable (it is very hard to tell "has function X because his genes survived" apart from "has function X to allow his genes to survive"). However, the two categories are very sharply distinct when it comes to human representations—we can talk more easily about representations that have a linguistic origin and distinguish them from ones which have an evolutionary origin (note that under such an approach, we do not assume that there are no evolutionarily-driven representations in humans—again, taking into account the results of cognitive science, such an assumption would be quite controversial, as many subpersonal representations, especially of the simple perceptual variety, do seem to have an evolutionary origin).

6. Summary

In light of all the argumentation presented, it seems that a completely naturalistic approach to human representations is hard to defend. Not only are there good philosophical reasons to refute it, there are also strong methodological reasons for the refutation related to the origin of representations in humans. On the other hand, the hybrid social-functional approach sketched here, which uses the origins of representational functions instead of their purposes, seems to be better suited for explaining the differences in representations between humans and simpler non-human organisms, as well as for dealing with the problem of misrepresentation. It remains to be seen how much of the research on functionalism with respect to representations can be ported to such an ap-

proach—however, I believe that such a hybrid solution would be effective and have the added benefit of bridging the gap between naturalistic and anti-naturalistic approaches to cognitive representations.

REFERENCES

Dennett, D. (1969). *Content and Consciousness*. London: Routledge.

Dretske, F. I. (1986). Misrepresentation. In: R. Bogdan (Ed.), *Belief: Form, Content, and Function* (pp. 17–37). Oxford: Clarendon Press.

Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Cambridge, Mass.: Harvard University Press.

Millikan, R. G. (1989). Biosemantics. *The Journal of Philosophy*, *86*(6), 281–297.

Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, Mass.: MIT Press.

Millikan, R. G. (1995). Pushmi-Pullyu Representations. *Philosophical Perspetives*, *9*, 185–200.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell Publishing.

Originally published as "Reprezentacje znaturalizowane – użyteczny cel czy użyteczna fikcja". *Studia Semiotyczne*, *31*(1), 91–108, DOI: 10.26333/sts.xxxi1.06. Translated by Piotr Wilkin.