Homunculi, the mereological fallacy and crypto-dualism.
Two dilemmas for the intentional stance

GLORIA DEL CARMEN BALDERAS ROSAS

Universidad de Navarra, Pamplona
gcbalderas@gmail.com

Abstract. Neuroscientist Maxwell Bennett and philosopher Peter Hacker defend the need to eradicate the mereological fallacy of cognitive neuroscience. This fallacy attributes to the parts of an animal psychological predicates that make sense only when applied to the whole animal. In contrast, philosopher Daniel Dennett argues that it is possible to explain behavior and cognitive abilities by applying the Intentional Stance (IS) to the brain, a stance that attributes increasingly simple psychological capacities to increasingly less intelligent homunculi. So, among other things, taking the intentional stance requires i) attributing psychological predicates to the brain and its parts, and ii) gradually replacing psychological predicates with non-psychological predicates.

If the criticism of Bennett and Hacker is accepted, these requirements lead to two dilemmas. According to the first, the intentional stance would be inapplicable or fallacious: because (i) it implies incurring in the mereological fallacy, and to deny (i) is equivalent to rejecting the intentional stance. The horns of the second dilemma are dualism and explanatory vacuity: (ii) can be interpreted in a dualistic sense because it presupposes the distinction between psychological and non-psychological predicates; however, failing to respect (ii) generates an infinite regress.

In this article I intend to show that both dilemmas are resolved by focusing on the details of the relationship between the intentional stance and the design stance.

Keywords: homunculus fallacy; mereological fallacy; intentional stance; design stance; prediction and explanation in cognitive sciences.

Introduction

In *Philosophical Foundations of Neuroscience* (Bennett and Hacker 2003) neuroscientist Maxwell Bennett and philosopher Peter Hacker defended the need to eradicate the mereological fallacy from cognitive neuroscience[1]. One incurs in this fallacy if psychological predicates that correspond to the whole animal are attributed to its parts. Bennett and Hacker state that this kind of illicit predication is not a fallacy in itself, but it leads to fallacious inferences and arguments. Their proposal has given rise to a heated debate, particularly with the positions of Daniel Dennett and John Searle, two philosophers to which Bennett and Hacker had extensively criticized in their book.

In what follows I assume that is possible to accuse Dennett of incurring in the mereological fallacy in order to state two dilemmas—the dilemma of predication and the ontological dilemma—to which I intend to respond. My paper begins with some terminological clarifications (Section 2), as a background for explaining the mereological fallacy (Section 3) and the dilemmas (Section 4). The ontological dilemma is discussed in Section 5 and the dilemma of predication in section 6. Finally, I present some concluding remarks.

1. Three homunculus fallacies

The so-called "homunculus fallacy" is a precedent to the mereological fallacy. The term was introduced by Anthony Kenny (Kenny 1987) for the purpose of advising caution in the application of predicates that are attributed to human beings, particularly in the field of cognitive science. The fallacy is committed when predicates proper to human beings are attributed to objects that are not sufficiently similar to human beings. Kenny explained that this inappropriate attribution is not strictly a fallacy, but leads to fallacious arguments. An example of this type of fallacious argument would be using (as a premise) the attribution of a predicate to

---

[1]   For some critical approaches see (Janzen 2008; Keestra and Cowley 2009; Pöyhönen 2014).

a human being in order to conclude that it is permissible to apply the same predicate to his or her brain. This type of attribution is a logical mistake, and raises methodological and pedagogical problems (Kenny 1987, 132).

Kenny describes the most naive version of the fallacy, which involves postulating a "little man" within each person to explain human experience and behavior (Kenny 1987, 125; 1989, 106–107). This homunculus creates a problem of infinite regress because its capabilities demand an explanation in their turn. It could also be said, though it is less common, that the homunculus leads to a vicious circle or to begging the question. Whatever the description of the difficulty is, the problem is that appealing to the homunculus is explanatorily empty[2]. In this case the logical problem of predication becomes a methodological problem about explanation, which Dennett has called the "homunculus problem".

Kenny recalls that, without using the term, Descartes rejected the homunculus fallacy. In his studies on vision Descartes warns his readers against the tendency to assume that there is a pair of eyes in the brain seeing the retinal images. However, Descartes also suggested that the soul could look at those pictures. This last statement, consistent with his substantial and interactionist dualism, continues to influence the debate on mental representation, a central notion in cognitive science used to explain capabilities such as perception, knowledge and language. Theories of representation assume at least that representation is something that has a content or meaning. Dennett, among others, criticizes the notion of representation because meaning requires an interpreter: "something is a representation only *for* or *to* someone; any representation or representations system requires at least one *user* of the system who is external to the system" (Dennett 1978a, 101). In this case an ontological problem arises because the representation requires a homunculus, and homunculi imply a commitment to dualism (Kenny 2010, 111). Of course, this internal "spectator" also generates problems of infinite regress, circularity and begging the question.

---

[2] See (Sober 1982) for a different analysis about homunculi and explanatory emptiness.

In short, homunculi in the literature on cognitive science entail logical, methodological and ontological problems, which allow distinguishing three "homunculus fallacies". Following the more accepted terminology, I will apply the label of the *mereological fallacy* to the problem of illicit predication; that of the *homunculus problem* to the explanatory vacuity caused by the regressus and that of the *Cartesian theater*[3] to a description of cognitive abilities committed to dualism. In what follows the ontological and methodological assumptions of the mereological fallacy are discussed.

2. The mereological fallacy

Bennett and Hacker follow Aristotle in proposing five clearly anti-Cartesian theses (Bennett and Hacker 2007b, 132, 209n8). First, it is a mistake to attribute to the *psyche* of an animal predicates that can only properly be attributed to the animal as a whole; Hacker calls this antecedent of the mereological fallacy the Aristotelian principle. This principle differs from the mereological fallacy, because the brain is literally a part of the body, while the *psyche* is not. Second, that the psyche is a set of capabilities. Third, that capabilities are identified by the activities they capacitate an agent to carry out. Fourth, that capacities are observable in the activities of organisms. Fifth, it is incoherent to ask whether mind and body are the same.

Hacker calls his ontological position neo-Aristotelian monism (Smit and Hacker 2013). According to neo-Aristotelian monism, there are important differences in the notions of human being, mind, body and person. A human being is a substance. The mind and body are not substances, nor can they be considered at all as parts of an organism. Furthermore, the notion of person differs from that of a human being, because the person is not a substance, rather, the term identifies the moral and legal status of human beings.

Hacker criticizes Descartes for turning the mind into the subject of psychological predicates (Hacker 2009). Since the mind is not a substance, it makes no sense to speak of the mind as an agent or a patient. Neither does it make sense to ask about the causal powers of the mind or whether the

---

[3] This term was coined by Dennett.

body is identified with it. The proper functioning of the brain is necessary for the exercise of human psychological faculties, but this does not mean that the brain is the possessor of such powers.

According to neo-Aristotelian monism it is correct to say that the human beings *have* a mind or a body; and also that the human being *is* a body. Saying that the human being *has* a mind or a body derives from the fact that humans have psychological and corporeal properties, but this does not mean that humans are in some kind of relationship with their bodies or their minds. Besides, the human is a body in the quasi-technical sense that every human being is a material space-time continuum, but one should avoid exporting this way of speaking to extra-philosophical contexts. Psychological predicates are attributed to human beings and, by extension, to other organisms. Bennett and Hacker maintain that humans differ from other animals by the evolution of language. The rational faculties of intellect and will are dependent on the use of language[4].

With regard to philosophical methodology, Bennett and Hacker (Bennett and Hacker 2007a, 4–11) distinguish sharply between empirical and conceptual questions. Cognitive neuroscience must do empirical research to explain the neural conditions that make possible the exercise of various cognitive abilities (Bennett and Hacker 2007b, 128). In contrast, the philosopher's task is to describe the logical relationships between concepts and to examine the structural relationships between different conceptual fields. The formulation of the mereological fallacy depends on this way of doing philosophy, and coincides with a passage from Wittgenstein's *Philosophical Investigations* quoted by Bennett and Hacker:

> Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind;·hears; is deaf; is conscious or unconscious (Wittgenstein 1953, I, § 281).

As has been said, the *mereological fallacy* consists in attributing to the parts of an animal the psychological predicates that correspond to the living acting

---

[4] "We are *Homo loquens* and therefore *Homo sapiens*, the wise ape" (Smit and Hacker 2013, 11).

human or animal as a whole. Therefore, attributing psychological predicates to the brain involves falling into the mereological fallacy. Also, identifying the mind with the brain amounts to treating them as substances, but they are not.

Cognitive neuroscience runs the risk of committing the mereological fallacy because it is on the frontier between neurophysiology and psychology and those disciplines employ concepts from very different categories. In general, one falls into this fallacy if a term—ordinary, technical or quasi-technical—is used in a context in which it is inapplicable or leads to equivocation. This lack of conceptual clarity can lead to fruitless questions or to undermining the strength of arguments, the utility of experiments and the interpretation of results. The solution for avoiding the mereological fallacy is to explain the logical-grammatical character of the concepts that are at stake.

Furthermore, Bennett and Hacker (Bennett and Hacker 2007b, 131) note that the most notable representatives of the first two generations of cognitive neuroscience were openly Cartesian and that the third generation subscribes to a crypto-Cartesian position, insofar as it retains a structural dualism or brain-body dualism that separates thought and action. In this sense, the practice of ascribing psychological predicates to the brain would be crypto-dualist.

Researchers can avoid the charge of committing the mereological fallacy if psychological predicates are used in a sense different from ordinary use. There are four "escape routes" from the fallacy: employing i) a derivative sense, ii) an analogous (or otherwise) extension of regular use, iii) homonymy and iv) figurative or metaphorical senses (Bennett and Hacker 2003, 74; 2007b, 149). If any of these are resources available, the intentional stance will be able to avoid the dilemmas presented in the next section.

3. Two dilemmas for the intentional stance

This section outlines what the intentional stance is, what it is for, and when it applies, in order to present the dilemma of predication and the ontological dilemma.

14

The intentional stance is a strategy that derives from what Dennett calls "folk psychology" (Dennett 1998/1987, 43–57). *Folk psychology* is a theory that explains and predicts human behavior by attributing mental states such as beliefs and desires. "[*B*]*eliefs* are information-bearing states of people that arise from perceptions and that, together with appropriately related *desires*, lead to intelligent *action*" (Dennett 1998/1987, 46). The problem is that mentalistic terms do not meet the requirements of scientific language. Scientific explanations are formulated in terms of cause and effect, which requires defining terms precisely. The terms of *folk psychology* often do not have clear criteria for their application; in particular, it is not clear how to identify these terms with causal conditions or with specific brain processes. For example, it is not possible to establish which brain processes or causal conditions are shared by people who believe that human beings have landed on the moon. This belief is compatible with a large number of neurophysiological conditions in different individuals and even in the same individual at different times (Dennett 1988, 501).

In order to preserve the benefits of *folk psychology* for explaining and predicting behavior, Dennett proposes to replace it with two theories: the intentional stance and the design stance.

Adopting the intentional stance (IS) allows predicting and explaining the behavior of a complex system, which can be a person, an organism or an artifact. This strategy has two requirements: first, the attribution of intentional predicates—mainly beliefs and desires—to the system under study; and second, the assumption of rationality (Dennett 1971). This presupposed rationality is a revisable idealization that guides the interpretation of the person who adopts the intentional stance (Dennett 2007a, 251). According to Dennett, attributing intentionality to psychological states is equivalent to ascribing content or meaning to it (Dennett 1998/1987, 341). Intentionality is characterized by the logical property of referential opacity.

A belief is referentially opaque because, in the sentence that expresses it, it is not possible to substitute one word for another that has the same reference, while preserving truth value. When terms are substituted it may be that the first sentence is true and the second is false (or vice

versa). In such contexts the coincidence in truth value of both sentences is accidental. For example, for most of Sophocles's tragedy, the sentence "Oedipus believes he has married Jocasta" is true, but the sentence "Oedipus believed he has married to the mother of Oedipus (i.e. his mother)" is false, although the terms "Jocasta" and "mother of Oedipus" refer to the same character.

According to IS, the subject of intentional attributions is the whole complex system (Dennett 1998/1987, 58). The attributions are made assuming the interaction between beliefs and desires, taking into account the data about system actions or changes in the environment. Dennett (Dennett 1983, 345–346) distinguishes between first-order intentional systems (able to have beliefs and desires) and second-order intentional systems (able to have beliefs and desires about beliefs and desires). Moreover, for Dennett (Dennett 2002, 91) it is obvious that language is what distinguishes humans from other species.

The design stance predicts the behavior of a complex system, assuming that it has been constructed in the best possible way to perform a task. The explanation consists in dividing the system into subsystems whose functions enable the capacity of the system as a whole. This analysis can be repeated at different levels as many times as necessary; additionally, the functions of the subsystems must be simpler and of a different type than the *explanandum* (Dennett 2007b, 88; Drayson 2012; Cummins 1975). Dennett emphasizes the teleological nature of this presupposition of optimality in design and justifies it by its heuristic power, i.e., its capacity to guide interpretation (which can be retrospective) and prediction about the behavior of the system (Dennett 1998/1987, 278–279).

When the issue is explaining the behavior of people, IS is a kind of explanation at the personal level while the design stance is an explanation at a subpersonal level (Dennett 1969; Drayson 2012). IS describes in mentalistic terms what the system has done or what it will do. The design stance must explain how he or she does it, i.e., which subsystems are required and which functions must be satisfied in order to have as the outcome the behavior described in intentional terms.

The joint application of IS and design stance is characteristic of research in artificial intelligence. An artificial intelligence program is designed by a top-down strategy: from the specifications provided by IS the system task is divided into simpler tasks until the description matches the function of some mechanism that the designer can achieve or build up. Dennett metaphorically described this strategy as the division of the intentional system into homunculi. To avoid the problem of regressus, Dennett added the requirement that the homunculi should not completely reproduce the ability of the system as a whole; rather, they should be increasingly "dumb" until the activity that they do can be executed mechanically (Dennett 1978b, 123–124)[5].

It is important to note that Dennett uses the term "intentional stance" in two ways, related to different objects of study: behavior and cognitive abilities. In the first case, IS is a version of *folk psychology* that can be applied in other disciplines, such as ethology (Dennett 1983, 349). In the second case, IS functions as the first step of the top-down strategy that attempts to determine how to carry out a cognitive capacity; therefore, IS is inseparable from the design stance (hereinafter IS-D). The possibilities of design are restricted by the way in which IS-D specifies the task that the system must perform. IS-D can be applied more than once at any level of analysis of the design stance.

IS-D is applied in neuroscience, where i) the subject of attribution is not the entire system—the person—but the brain and its parts; and ii) psychological predicates are gradually replaced by non-psychological predicates. These two assumptions lead to the ontological dilemma and the dilemma of predication.

According to the dilemma of predication IS-D is fallacious or inapplicable. If psychological predicates are attributed to the brain IS-D incurs in the mereological fallacy, but avoiding these ascriptions prevents applying the intentional stance.

If the difference between psychological and non-psychological predicates is interpreted as an ontological distinction between the physical and

---

[5] Lycan (Lycan 1981) calls Dennett's proposal "homuncular functionalism". Dennett agree (Dennett 1978c, xx; 1998c, 362). Famously, Searle argues against this strategy.

the mental—or as Dennett says, between the mechanical and the non-mechanical (Dennett 1969, 91; 2007b, 76)—it seems that IS and IS-D are crypto-dualist and the replacement of predicates is not legitimate. If, to avoid dualism, one rejects the thesis that psychological predicates are replaced by non-psychological predicates, then IS-D loses its explanatory power. The reason is that the attribution of psychological predicates to subsystems obviously is not an explanation, but a redescription of the explanandum.

The next two sections are dedicated to the discussion of these dilemmas, beginning with the ontological dilemma.

4. The ontological dilemma

Dennett describes his philosophical activity as a reflection on the implications of the standard scientific image, particularly with regard to intentionality and consciousness (Dennett 1998c). He assumes that science cannot be free of metaphors, but it is possible to replace some bad metaphors by better ones. He describes himself as a kind of engineer designing instruments that can be improved or discarded if they fail to accomplish their function. (Dennett 1993, 203–204).

Unlike Hacker, Dennett is reluctant to propose any ontological thesis (Dennett 1969, 189; 1991, 29, 51; 1993, 210–214). Therefore, this section does not intend to discuss the ontology of Dennett, but rather it relies on some of his statements to determine whether the difference between psychological and non-psychological predicates should be considered as marking an ontological distinction. If the distinction is not ontological, and no use of the intentional stance implies dualism, then IS and IS-D can legitimately move from one type of predicate to another.

It is possible to tentatively construct Dennett's position regarding the nature of mind on the basis of his criticism of various –isms, together with the infrequent positive characterizations of his own position.

Usually Dennett is considered as an instrumentalist, a position for which the validity of intentional attributions depends on its utility, not on its truth. Instrumentalism can be understood as assuming that such

adscriptions are neither true nor false, or that they are useful fictions (fictionalism). Dennett has insisted that he is not instrumentalist in any of these ways (Dennett 1998/1987, 69–81). Nor is he an interpretationist if this position implies that ascriptions of belief do not have an objective basis but are entirely dependent on the observer. For Dennett (Dennett 1988, 496) beliefs are objective phenomena, although his position can be called "interpretivist" in the sense that in order to recognize these beliefs one must adopt IS.

On the other hand, Dennett has rejected substantial dualism:

> The problem of mind is not to be divorced from the problem of a person. Looking at the "phenomena of mind" can only be looking at what a *person* does, feels, thinks, experiences; minds cannot be examined as separable entities without leading inevitably to Cartesian spirits (Dennett 1969, 189)[6].

Realism in the philosophy of mind is a materialistic view that shares with Cartesian dualism the conviction that the mind is some kind of entity. Dennett rejects realism: thoughts cannot be identified with physical processes in the brain because the mind is not a "thing" that can be identified (or not) with another. For the same reason Dennett refuses to identify intentional states with logical or functional states of the nervous system (Dennett 1969, 189). However, Dennett has suggested, without admitting it, that somehow his position can be seen as a *sort* of dualism (Dennett 1993, 210). So, is Dennett a crypto-dualist?

Dennett calls his position *mild-realism* (Dennett 1991, 29–30). To explain it, Dennett focuses on the predictive power of IS. Predictions function if *there is* an order or pattern in the world. In the case of intentional attributions, the pattern "is discernible in agents' (observable) behavior" when IS is adopted, that is, when it is assumed that the beliefs attributed are appropriate for a rational agent to achieve its purposes in a given context.

---

[6] Used by Dennett, "person" refers to human beings. Hacker's notion of person would be comparable to the Dennettian forensic notion of personhood.

Although the agent's behavior is the necessary "data" for adopting IS, Dennett it is not a behaviorist, because IS presupposes intentionality and behaviorism denies it. A passage from Anscombe quoted by Dennett (Dennett 1991, 43) suggests what the character is of the patterns captured by the intentional stance. Anscombe says that the interest of the Aristotelian explanation regarding the practical syllogism is that this relationship "describes an order which is there whenever actions are done with intentions" (Anscombe 1957, 80). From Dennett's exposition it can be inferred that the order or pattern of intentional attributions is analogous to the order that is discernible in organisms and that results from selective pressure. These patterns are not mere regularities, but rules of good design. They are present in behavior in the way that the rules of aerodynamics can be present in the wings of birds. Patterns are abstract, but real.

Design and intentionality are in the same boat. Dennett claims that IS can be seen as a "limiting case" of the design stance, because the predictions of IS presuppose rationality and rational design is an optimal design. This affinity between attitudes depends on their teleological character (Dennett 1998a, 312; 1998/1987, 73). In connection with this teleology it is worth transcribing a passage where Dennett rejects epiphenomenalism, in which it is clear that intentional patterns are real, even causally efficacious:

> Several interpreters of a draft of this article have supposed that the conclusion I am urging here is that beliefs (or their contents) are epiphenomena having no causal powers, but this is a misinterpretation traceable to a simplistic notion of causation. If one finds a predictive pattern of the sort just described one has ipso facto discovered a causal power—a difference in the world that makes a subsequent difference testable by standard empirical methods of variable manipulation. [...] The fact that the regularities on which these successful predictions are based are efficiently capturable (only) in intentional terms and are not derived from "covering laws" does not show that the regularities are not "causal"; it just shows that philosophers have often relied on pinched notions of causality derived from exclusive attention to a few examples drawn from physics and chemistry. Smith has pointed out to me that here I am echoing Aristotle's claim that his predecessors had ignored final causes (Dennett 1991, 43–44n22).

Dennett adds that there may be rival intentional interpretations without it being possible to establish which one is correct. This is because IS is an idealization and also because interpreters who adopt IS differ in their interests and experience.

It is worth highlighting a difference between IS and the design stance that is relevant to the extent that IS is applied on a personal level to explain the behavior of humans and other living beings.

When a device is designed, the subsystems are constructed in relative isolation from each other to avoid interference and other unwanted side effects. In contrast, in nature the "designs" are not isolated. Sometimes their interaction produces innovative and beneficial side effects: these effects *emerge* through interaction. To show that this type of emergency does not imply an ontological gap, Dennett gives the example of a chess program. According to a designer "[the program] thinks it should get its queen out early" (Dennett 1978a, 107). The program in question does not have explicit rules for that movement, the movement is a non-intended outcome of its global functioning, i.e., from the interaction of the rules that are part of the project of the designer. It is possible that the patterns of behavior that IS captures have emerged during the evolutionary process in a similar way, such that they are not an ontological novelty.

In light of this notion of emergence Dennett is understandably skeptical about reductionism. Dennett (Dennett 1998/1987, 68) thinks that the reduction is not about eliminating intentional terms to replace them by physiological terms. Rather, the "reduction" would be to show how a system described in physiological terms allows or justifies an intentional characterization. In this sense, the distinction between the physiological and the psychological is not a border or dividing line between two entity types, but the difference between a holistic view of the system in relation to its environment and the job description of the subsystems presumably involved when certain behavior is presented.

If this interpretation of Dennett's position is correct, then the distinction between the personal and the subpersonal in IS coincides with the distinction between the behaving organism (the whole) and the brain (a part). Smit

and Hacker (Smit and Hacker 2013, 9) have rejected this claim. At first glance one might say that they are right because Dennett (Dennett 1969, 91; 2007b, 76) uses the terms "non-mechanical" and "mechanical" as synonyms for "personal" and "subpersonal". Literally and out of context, this identification is misleading because it suggests an ontological difference. However, from a more general perspective, the identification makes sense because there is no genuine ontological distinction between "mechanical" and "non-mechanical", but rather a distinction between the observable patterns in the behavior of the whole and the functions of a relevant subsystem (the brain).

Dennett cannot be accused of crypto-dualism because there is no ontological difference between the psychological and the non-psychological. Solving this horn of the ontological dilemma permits avoiding the *regressus* by replacing intentional predicates with mechanical predicates in both IS and IS-D. However, it remains to clarify the meaning of intentional attributions at the subpersonal level.

5. The dilemma of predication

Bennett and Hacker (Bennett and Hacker 2007b, 147) affirm that a competent speaker knows the meaning of a word and its proper use just as well as a competent mathematician knows how to calculate or a competent chess player knows how to move his or her pieces. In addition to ordinary use, they admit four types of predication that do not incur in the mereological fallacy: metaphor, homonymy, analogy and derived senses. In what follows I examine whether Dennett applies the intentional predicates to the subpersonal level in any of these ways.

Dennett avoids the fallacy when he uses "homunculus" metaphorically (Dennett 1978b, 123–124). However, Dennett also describes his use of psychological terms at the subpersonal level in other ways. I will focus on two: literal (Dennett 2007b, 89) and attenuated (Dennett 2007b, 87).

Dennett cites a passage in which Hacker says that it would make sense to ascribe psychological predicates to brains if it were possible to literally attribute psychological predicates to machines. This requires accepting

that machines behave like humans when simulating human cognitive abilities and that brains and machines are similar in terms of the functional structure of these capabilities.

Dennett accepts this reasoning (Dennett 2007b, 89). But it might be better to reject it because it is not clear that psychological predicates apply to machines *literally*. Dennett himself has questioned whether the machines of the artificial intelligence projects behave like humans and whether their programs effectively simulate human or animal capacities. In addition, Dennett has proposed considering artificial intelligence programs as "thought experiments" rather than as models of such capacities (Dennett 1978b, 117; 1998d, 231). These programs often oversimplify tasks and system requirements. To overcome this difficulty Dennett has suggested attempting the simulation of the activity of simpler, whole organisms, rather than trying to reproduce human micro-skills (Dennett 1978d; 1998/1987, 257).

Dennett argues that the attribution of psychological predicates to the brain is a legitimate extension of the ordinary uses of these terms because some parts of the brain are involved in certain processes very similar to behaviors at the personal level, such as deciding or believing. That is, they "behave as" a human. Accepting or disavowing this use depends on the data of empirical research that confirms the similarity between what the brain does and what people do. Dennett (Dennett 2007b, 87) adds that these intentional attributions to the brain are not made in the full sense of the terms, but in an attenuated sense that prevents incurring in the mereological fallacy.

Bennett and Hacker (Bennett and Hacker 2007b, 211n23) accept the analogical extension of the use of the terms, but reject the "attenuated sense" of Dennett, appealing to the behavioral criteria that are used on a personal level to make psychological attributions. They note, for example, that it is not clear why it would be legitimate to attribute beliefs to a brain, because the brain cannot display the verbal and nonverbal behavior that helps to determine what a person believes (Bennett and Hacker 2007b, 141). This argument is sound, but I think that Dennett can still avoid the fallacy.

First, it is debatable that the rules for the ordinary use of the terms are as accurate as, for example, the rules of chess or mathematics (Dennett

2007b, 85). Ordinary language is probably more flexible than Bennett and Hacker seem to assume, and certainly analogy is more common than univocity.

On the other hand, it can be noted that Dennett never uses the intentional terms in their ordinary senses. Even in his folk psychology the terms are used while eliminating many of their habitual connotations, so this use is already analogous. In *folk psychology* terms as "belief" and "desire" are used with interpretive and predictive purposes, but it is true—as Bennett and Hacker say—that when we learn to use the terms it is not always or even primarily to predict behavior (Bennett and Hacker 2003, 425). Actually, *folk psychology* is a didactical tool for introducing IS and the design stance. The role of this theory in Dennett's proposal is consistent with his conviction about that explanations of behavior in ordinary terms are only indirectly involved in scientific discourse (Dennett 2007b, 202n1).

Since the use of psychological predicates in IS and IS-D is also subordinate to the explanatory and methodological interests (Dennett 2007b, 84, 86, 90), it could be said that these intentional predicates do not satisfy the criteria for ordinary use, but rather that the terms have a technical (or quasi-technical) use that fits with three assumptions: optimality (rational or designed), teleology and holism. It can be added that, perhaps, in Dennett's view optimality and teleology are equivalent.

When it comes to explaining behavior, attributions of IS are dependent on the assumption of rationality and on the relations between one another of the beliefs and desires of the intentional system, as well as the context in which the subject operates and the previous actions of the system.

When it comes to explaining a capacity, each psychological predicate attributed to IS-D homunculi is an ascription of function. It is assumed that the systems are organized hierarchically to carry out certain activities and processes that enable that capacity.

Furthermore, the use of psychological predicates is different in IS and IS-D. IS can be applied without appealing to the design stance, but IS-D is at the service of the design stance.

It has already been said that IS-D specifies the task to be performed by the system. But IS-D is also required if the designer does not know how

a subsystem should work to perform its function. If the designer encounters a difficulty in constructing a subsystem, he or she characterizes it intentionally, *as if* the subsystem was an intelligent subsystem (Dennett 1998a, 312). This strategy is equivalent to assuming that the subsystem at issue must be well designed to perform its function. In neuroscience homunculi are abstract and hypothetical analyses of the roles of different structures or processes in the brain presumably involved in the exercise of a given capacity. The explanation progresses to the extent that these imaginary homunculi are replaced with neural structures and activities (Dennett 2007a, 14). In this case it is also assumed that the brain is like a perfect machine, however, this assumption is revisable; any hypothetical analysis made in intentional terms is also revisable and correctable. Once you understand what the brain (or a brain's part) does, psychological predicates are discardable.

Bennett and Hacker question the usefulness of this strategy while Dennett defends it. Bennett and Hacker hold that intentional predicates block scientific progress. In contrast, Dennett trusts that these predicates facilitate an explanation by offering hypotheses and giving provisional labels to unresolved problems. This article will not seek to settle this debate. In any case, it corresponds to neuroscientists to decide whether the strategy is useful for their purposes. The aim here is to determine whether IS and IS-D commit the mereological fallacy.

IS can avoid the fallacy because there is no ontological difference between psychological and non-psychological predicates; rather, there is a difference between the consideration of certain patterns in the behavior of the system and the perspective of the functions of the parts. When IS-D is used in neuroscience, it also avoids the intentional fallacy because attributions serve as hypothetical descriptions of system functions or as an interim strategy to defer (not impede) the explanation.

Concluding remarks

In this article I have distinguished two uses of the intentional stance. The first one, IS, is a strategy to interpret behavior at the personal level.

The second one, IS-D has three functions in relation to the design stance: specifying the task performed by the system, supplying hypothetical analysis to distribute functions among subsystems and providing labels for functions not yet analyzed.

Both IS and IS-D presuppose the distinction between psychological and non-psychological predicates. This distinction is not between two types of entities or things, but between two approaches to the intentional system. So IS and IS-D can avoid the ontological dilemma.

In addition, IS and IS-D can avoid the dilemma of predication because the use of psychological terms has, in principle, a technical character. However, one should proceed with caution to avoid undue conflation of this sense and ordinary usage.

References

Anscombe, G. E. M. 1957. *Intention*. New York: Blackwell.

Bennett, Maxwell R., and Peter M. S. Hacker. 2003. *Philosophical foundations of neuroscience*. Oxford: Blackwell.

—. 2007a. "Selections *Philosophical foundations of neuroscience*." In *Neuroscience and philosophy: Brain, Mind, and Language*, edited by Maxwell R. Bennett, Daniel Dennett, Peter M. S. Hacker and John Searle; introduction and conclusion by Daniel Robinson, 3–48. New York: Columbia University Press.

—. 2007b. "The Conceptual Presuppositions of Cognitive Neuroscience: A Reply to Critics." In *Neuroscience and philosophy: Brain, Mind, and Language*, edited by Maxwell R. Bennett, Daniel Dennett, Peter M. S. Hacker and John Searle; introduction and conclusion by Daniel Robinson, 127–162. New York: Columbia University Press.

Cummins, R. 1975. "Functional analysis." *The Journal of Philosophy* 72 (20): 741–765. DOI: http://dx.doi.org/10.2307/2024640.

Dennett, Daniel C. 1969. *Content and consciousness*. London: Routledge & Kegan Paul.

—. 1971. "Intentional Systems." *The Journal of Philosophy* 68 (4): 87–106. DOI: http://dx.doi.org/10.2307/2025382.

—. 1978a. "A Cure for the Common Code?" In *Brainstorms: Philosophical essays on mind and psychology*, 90–108. Bradford Books.

—. 1978b. "Artificial Intelligence as Philosophy and as Psychology." In *Brainstorms: Philosophical essays on mind and psychology*, 109–126. Bradford Books.

—. 1978c. "Introduction." In *Brainstorms: Philosophical essays on mind and psychology*, ix-xxii. Bradford Books.

—. 1978d. "Why not the whole iguana? (Commentary on Pylyshyn)." *Behavioral and Brain Sciences* 1 (01): 103–104. DOI: http://dx.doi.org/10.1017/S0140525X00059859.

—. 1983. "Intentional systems in cognitive ethology: The 'Panglossian paradigm' defended." *Behavioral and Brain Sciences* 6 (03): 343–355. DOI: http://dx.doi.org/10.1017/S0140525X00016393.

—. 1988. "Précis of The Intentional Stance." *Behavioral and Brain Sciences* 11: 495–546. DOI: http://dx.doi.org/10.1017/S0140525X00058611.

—. 1991. "Real Patterns." *The Journal of Philosophy* 88(1): 27–51. DOI: http://dx.doi.org/10.2307/2027085.

—. 1993. "Back from the Drawing Board." In *Dennett and his Critics*, edited by Bo Dahlbom. Oxford: Basil Blackwell.

—. 1998a. "Cognitive Ethology: Hunting for Bargains or a Wild Goose Chase." In *Brainchildren Essays on Designing Minds*, 307–322. Cambridge, MA: Bradford Books/The MIT Press.

—. 1998b. "Cognitive Science as Reverse Engineering: Several Meanings of 'Top-Down' and 'Bottom-Up'." In *Brainchildren Essays on Designing Minds*, 249–259. Cambridge, MA: Bradford Books/The MIT Press.

—. 1998c. "Self-portrait." In *Brainchildren Essays on Designing Minds*, 355–366. Cambridge, MA: Bradford Books/The MIT Press.

—. 1998d. "The Logical Geography of Computational Approaches: A View from the East Pole." In *Brainchildren Essays on Designing Minds*, 215–234. Cambridge, MA: Bradford Books/The MIT Press.

—. 1998/1987. *The Intentional Stance*. Cambridge, MA: MIT Press.

—. 2002. "Reply to Clark." In *Philosophy of Mental Representation*, edited by Hugh Clapin, 91–93. Oxford: Clarendon Press.

—. 2007a. "Heterophenomenology reconsidered." *Phenomenology and the Cognitive Sciences* 6 (1): 247–270. DOI: http://dx.doi.org/10.1007/s11097-006-9044-9.

—. 2007b. "Philosophy as naïve anthropology: Comment on Bennett and Hacker." In *Neuroscience and philosophy: Brain, Mind, and Language*, edited by Maxwell R. Bennett, Daniel Dennett, Peter M. S. Hacker and John Searle; introduction and conclusion by Daniel Robinson, 73–95. New York: Columbia University Press.

Drayson, Z. 2012. "The uses and abuses of the personal/subpersonal distinction." *Philosophical Perspectives* 26 (1): 1–18.

Hacker, Peter M. S. 2009. "On having a Mind, having a Body, and being a Person." Философия. Социология. Политология. Вестник Томского Государственного Университета 2 (6): 147–160.

Janzen, G. 2008. "Bennett and Hacker on neural materialism." *Acta Analytica* 23 (3): 273–286.

Keestra, M., and S. J. Cowley. 2009. "Foundationalism and neuroscience; silence and language." *Language Sciences* 31 (4): 531–552.

Kenny, Anthony. 1987. *"The Homunculus Fallacy." The legacy of Wittgenstein*. Oxford: Basil Blackwell. Originally published in Marjorie Grene, ed., *Interpretations of life and mind: essays around the problem of reduction* (Routledge & Kegan Paul, 1971).

—. 1989. *The metaphysics of mind*. Oxford: Clarendon Press.

—. 2010. "Concepts, brains, and behaviour." *Grazer Philosophische Studien* 81 (1): 105–113.

Lycan, William G. 1981. "Form, Function and Feel." *The Journal of Philosophy* 78 (1): 24–50. DOI: http://dx.doi.org/10.2307/2025395.

Pöyhönen, S. 2014. "Intentional concepts in cognitive neuroscience." *Philosophical Explorations* 17 (1): 93–109. DOI: http://dx.doi.org/10.1080/13869795.2013.742556.

Smit, H., and P. M. Hacker. 2013. "Seven Misconceptions About the Mereological Fallacy: A Compilation for the Perplexed." *Erkenntnis* 79 (5): 1–21. DOI: http://dx.doi.org/10.1007/s10670–013–9594–5.

Sober, E. 1982. "Why must homunculi be so stupid?" *Mind* 91 (363): 420–422. DOI: http://dx.doi.org/10.1093/mind/XCI.363.420.

Wittgenstein, L. 1953. *Philosophical Investigations*, edited by G. E. M. Anscombe and R. Rhees, translated by G. E. M. Anscombe. Oxford: Blackwell.