Jakub Bożydar Wiśniewski

Irreducible Holism

Diametros nr 30, 76-92

2011

Diametros № 30 (December 2011): 76-92

IRREDUCIBLE HOLISM

– Jakub Bożydar Wiśniewski –

In this paper I would like to explore some issues surrounding ontological reduction, as advocated by the friends of physicalism, and its relation to conceptual reduction, which the said persons generally seem to eschew. It is my intention, and my hope, that the remarks contained herein will shed some new light on those scientific and philosophical programmes that include ontological reduction[1] as one of their foundational assumptions. More specifically, I want to hint at certain general methodological and ethical implications of the research projects (both in the natural sciences and in the philosophy of mind) that aspire to explain the nature of complex phenomena within a physicalist setting, and offer their brief analysis.

Furthermore, I am going to argue against the feasibility of an all-reaching physicalist framework for accomplishing the above-mentioned task by attempting to show that a certain set of concepts bridges the ostensible gap between ontological and conceptual reduction; these concepts, which I shall call 'irreducibly holistic', appear to me to be unamenable to a physicalist treatment in which the metaphysical and semantic components are supposed to be clearly separable. Within the ambit of irreducible holism, as I shall argue, ontological and conceptual reduction are inextricably bound together, hence if the latter is impossible, then so is the former; and since I take irreducibly holistic concepts to be insusceptible to conceptual reduction[2], physicalism with respect to them (and thus physicalism *per se*) turns out to be untenable.

---

[1] From now on, I shall use the term "ontological reduction" as a shorthand for "ontological reduction as endorsed by physicalists" and the phrase "ontologically reducible" as a shorthand for "ontologically reducible in the physicalist sense". The reason why I allow myself to treat the above pairs of terms as almost synonymous is the fact that I am not familiar with any relevant modern literature propounding any form of ontological reduction free of physicalist assumptions, or, for that matter, any form of not purely naturalistic ontological reduction. If there is a respectable version of idealist or spiritualist monism on the present-day marketplace of philosophical ideas, then I apologize for not acknowledging it, but in any event my arguments are not concerned with any such view.

[2] As the present paper is predominantly concerned with conceptual reduction within the context of discussions of physicalism, my use of the term "conceptual reduction" can be safely treated as a shorthand for "conceptual reduction to physics". This, however, should not lead one to assume

An alternative to embracing that conclusion is to revise one's conceptual framework by eliminating the troublesome elements from it. Nothing in principle forbids that step, but some desirable modifications in one place oftentimes result in undesirable changes elsewhere; I will attempt to show, for instance, that the elimination of what I take to be irreducibly holistic ethical concepts precludes the possibility of living what is ordinarily considered as a coherent and contentful moral life.

There are yet other ways of approaching irreducible holism from within a physicalist setting, and I will concisely assess their viability before concluding that none of them, in my view, succeeds in reconciling the two. Hence, the upshot of my arguments may be seen as an endorsement of a form of ontological pluralism.

The currently prevailing construal of physicalism appears to be that of a programme directed at "a system of knowledge within which all aspects of reality have a place and are related to physics in certain specifiable ways" (Poland [1994] p. 10). In other words, it is a construal that postulates the existence of a multi-level physical reality, whose distinct echelons are related to one another, as well as to the rock-bottom foundation of physics, by the principle of supervenience. The concept of supervenience has been variously defined and used to articulate numerous distinct claims, so perhaps it is still not entirely free from a certain degree of ambiguity, yet I believe that a sort of uncontroversial core can be extracted from it. Following Stalnaker ([1996] p. 4), I take supervenience to consist of an ontologically reductive element (supervenient facts or properties are nothing over and above their subvenient base – if the subvenient base is physical, then the supervenient facts or properties have to be physical as well) and an explicit denial of a potentially corresponding conceptually reductive element; in Stalnaker's words, "the concept of supervenience is supposed to be a concept that helps to isolate the metaphysical part of a reductionist claim – to separate it from claims about the conceptual resources and explicit expressive power of theories we use to describe the world".

There are clear examples of areas in which supervenience accomplishes its intended purpose: for instance, earth sciences such as geology, oceanography and paleontology are conceptually irreducible to physics (since each of them focuses on certain specific high-level phenomena, whose effective explanation requires studying certain specific high-level arrangements of physical forces and particles,

---

that irreducibly holistic concepts are insusceptible only to that kind of conceptual reduction; on the contrary, I take them to be conceptually irreducible *simpliciter*.

not the low-level nature of what is a physical force or a physical particle) but the facts and properties studied by these disciplines share the same ontological building material with the facts and properties studied by physics. Irreducibly holistic concepts, on the other hand, do not supervene on the physical, since their conceptual irreducibility to physics implies that they do not share the same ontological building material with the phenomena studied either by the said discipline or by any branch of science that is ontologically rooted in it.

Examples of concepts that I consider to be irreducibly holistic are: dignity, conscience, empathy and inspiration; I would like to refer to some of these in order to elaborate a bit further on the nature of the doctrine espoused in this paper. At the same time, I hope to dispel some of the possible misunderstandings of what I have said so far.

Firstly, it must be made clear that although I think of irreducibly holistic concepts as immune to conceptual reduction, I certainly do not think of them as immune to conceptual analysis. Let us take the first two examples from the beginning of the previous paragraph. One could describe them as a pair of complex ethical and mental concepts – that is, concepts analyzable in terms of ethical predicates and propositional mental states. Thus, for instance, dignity could perhaps be analyzed as the self-judgment that one is worthy of respect and honor or as the belief that one has an inalienable negative liberty. I regard these as successful and illuminating analyses, yet the concepts employed in them do not seem to be of a lower level than the corresponding *analysanda*. On the contrary, it appears to me that when it comes to conceptual analyses of irreducibly holistic concepts, the *analysantia* and the *analysanda* turn out to be interdefinable on a single level.

Secondly, to some it might seem that holism implies irreducibility, thus making 'irreducible holism' sound like a tautology. I do not believe that the above implication holds – for instance, various systems that include emergent properties are clearly holistic (since these properties cannot arise unless the more fundamental properties of a given system are all at work simultaneously), but their holism is not irreducible (since emergent qualities in these systems are directly traceable to their elemental constituents, and their complex, novel character can be fully accounted for in terms of the more fundamental qualities that give rise to it). In short, if a given system is to be treated as irreducibly holistic, its emergent qualities may be susceptible to conceptual analyses, but they cannot be conceptually reducible (or even explicitly traceable) to the low-level processes that presumably underlie or cause their appearance. Furthermore, as I have stressed several times already, conceptual irreducibility of these qualities must be associated with their

ontological irreducibility; or, if one prefers to put it in other terms, with their ontological distinctness from the qualities (or properties, phenomena, etc.) that partake in the above-mentioned low-level processes.

Thirdly, one might claim that the doctrine of irreducible holism appears to be just another name for the doctrine of the so-called "strong emergence". The reasons for that could stem from conceiving of irreducible holism as the conjunction of the familiar dictum that the whole is more than the sum of its parts and the thesis that in some cases no amount of knowledge and understanding of the nature of the parts can explain the essential characteristics of the emergent whole. I see the above interpretation as largely inaccurate and thus concealing real differences that exist between irreducible holism and strong emergence, which I shall try to elucidate by referring to Chalmers' [2006] account of the latter doctrine in question.

According to Chalmers, "a high-level phenomenon is strongly emergent with respect to a low-level domain when the high-level phenomenon arises from the low-level domain, but truths concerning that phenomenon are not deducible even in principle from truths in the low-level domain". With regard to this point, the considered views are alike, but I feel that this is just about where their explicit similarities end. For one thing, I take the category of irreducibly holistic phenomena to be much more restricted than the category of strongly emergent phenomena (which already implies a divergence from Chalmers' position, since he thinks that there is only one strongly emergent phenomenon, namely, consciousness). It is because, in my view, besides being non-deducible even in principle from the underlying low-level processes, irreducibly holistic phenomena are ontologically distinct from these processes; this is not the case with strongly emergent phenomena, which, although non-deducible even in principle from the underlying low-level processes, can nonetheless share the same ontological building material with the phenomena that constitute these processes. In other words, unlike irreducible holism, strong emergence is compatible with ontological reduction, and hence with physicalism, provided that one takes physical reality to be sufficiently anomalous and unpredictable.

Moreover, strongly emergent phenomena, as their name suggests, need some more primitive base to emerge from, whereas irreducibly holistic phenomena can themselves be ontologically basic. Perhaps it is appropriate to say here that if I were not discussing irreducible holism in opposition to physicalism, I could safely mention the fundamental concepts of physics as additional examples of concepts that I regard as irreducibly holistic. However, since the endorsement of physicalism requires an intellectual confrontation with the totality of irreducibly holistic concepts, not only the unproblematic ones, I decided to

leave the only category that consists solely of the latter out of my discussion of the title doctrine. If, on the other hand, I were arguing against idealism, I would point to the irreducibly holistic character of the basic concepts of physics without any hesitation whatsoever; that is, I would contend that since the physical is conceptually irreducible to the mental, it has to be ontologically irreducible to it as well, and therefore reality cannot be wholly mental.

Fourthly and lastly, it needs to be said that those irreducibly holistic phenomena which are not ontologically basic are of course not precluded from having partially physical causes – after all, all that their nature requires in this regard is that they be caused by the processes from which they are ontologically distinct and to which they are conceptually irreducible, and physical processes satisfy both of these criteria. However, it immediately has to be added that the connections between the said phenomena and the physical part of their causes are not necessary, hence they are not subject to the relation of supervenience – as a slightly modified slogan would have it, in this case there can be an *A*-difference without a *B*-difference. To give an illustrative example, there is no logical impossibility (or even a metaphysical impossibility) involved in the notion that a non-physical entity can possess dignity and conscience just as well as a physical entity (or a psychophysical entity) can. Here one might conjure up the age-old doubts concerning the notion of inter-substantial causation, but such a move would seem to me to be based on the equally age-old prejudice of thinking that causation is always a matter of reason, not a matter of fact, and that causes should always partake in the same substance as their effects. Consequently, I shall not pursue this issue further, except to refer the reader to some very helpful remarks on it by Crane and Mellor ([1990] p. 8).

Now, having hopefully made the definition of irreducible holism sufficiently clear, let me sketch once again the possible ways of approaching the concepts characterized by it. One is to regard them as corresponding to ontologically *sui generis* objects. Another is to eliminate them from one's conceptual scheme. The third is to try to accommodate them in one of the ways subsumed under the label of "preservative irrealism" (Horgan [1993] p. 27) – that is, continue using them on the grounds of their explanatory utility or communicational indispensability without thereby incurring any realist commitments. My sympathies obviously lie with the first option, and I shall argue for it by means of attempting to show that the remaining alternatives are untenable.

To start with, let us recall that physicalism can be uncontroversially defined as a programme directed at a system of knowledge within which all aspects of re-

ality supervene on the physical. Thus, since propositions involving what I consider to be irreducibly holistic concepts, such as "you violated my dignity" or "my conscience tells me this", clearly do convey some sort of knowledge (they express something distinctive about certain aspects of reality), the working assumption for the physicalist must be that this knowledge is in some sense ultimately grounded in physical facts (and hence ontologically reducible). It does not seem, however, to be so grounded in the same manner as the knowledge conveyed by the earth sciences mentioned earlier (or any other higher-level sciences) – unlike it, it does not appear to have any explanatory power with respect to the nature or workings of either high-level or low-level arrangements of physical forces and particles, nor does it appear to perform any normative role in any "science of behaviour".

Consequently, the physicalist seems to be confined to choosing either between concluding that the aforementioned propositions do not express any genuine knowledge (and therefore the troublesome concepts employed in them should be eliminated as meaningless) or finding some way to anchor them ontologically (but not necessarily conceptually) in propositions belonging to either physics or any of the higher-level natural sciences. Of course, neither of these options is compatible with the acceptance of irreducible holism; on the contrary, they envisage that the supposedly irreducibly holistic concepts are either vacuous and redundant or supervenient on concepts that are themselves not irreducibly holistic.

As far as I can tell, there is currently no known way of implementing the latter alternative – no preservative irrealism capable of showing that the phenomena of dignity and conscience are supervenient on the realm of the physical looms on the horizon of contemporary scholarship. Even if it were, however, it would not be of any help to the physicalist – it could perhaps strip the above phenomena of their allegedly irreducibly holistic character, but in order to do so, it would have to help itself to irreducible holism elsewhere. Now let us see why that would be the case.

Suppose, for the sake of argument, that one succeeded in demonstrating that the phenomena of dignity and conscience supervene on the facts and properties studied by neuropsychology. Now, since neuropsychology is rooted in the domain of biochemical mechanisms, which in turn is ultimately governed by the fundamental physical processes, the demonstration in question would amount to revealing that dignity and conscience are supervenient on, say, the sets of physical states *P1* and *P2* respectively.

Yet having said that, it remains to be explained what is the reason for retaining complex ethical terms in the conceptual framework that supposedly describes the total physical reality, which is claimed to be all there is, given that these concepts have no special explanatory power or analytical utility with respect to any of its constituent levels, be they high or low. In other words, a tool for successfully implementing the programme of preservative irrealism remains to be found. Let us now survey some possible candidates for this role.

If one wishes to be a consistent physicalist, yet finds it uncomfortable to cease to talk about seeking inspiration, listening to one's conscience and respecting the dignity of others, one could perhaps try to avoid the predicament by introducing the distinction between 'formal', 'scientific', 'objective' language and 'informal', 'familiar', 'subjective' language, subsequently attempting to use them alternately, depending on the context. Thus the agenda of preservation would consist in emphasizing the legitimacy and importance of the expressive diversity of communicational activities. This putative solution, however, creates more problems than it solves. Besides opening the seemingly unbridgeable gap between the subjective and the objective view of reality, a gap whose existence no coherent physicalism can afford to admit[3], it allows for treating the aforementioned distinct types of language themselves[4] as irreducibly holistic concepts.

The physicalist might nonetheless try to follow this dubious escape route and suggest that the 'informal', 'familiar' type of language used by him is not, strictly speaking, a distinct type of language, but only a more expedient form of expressing 'formal', 'scientific' statements, a form that employs certain 'mental shortcuts' and 'abbreviatory metaphors' (e.g., saying "I feel the pangs of conscience" instead of "the set of moral states *M* supervenient on the set of physical states *P* obtains").

As can be expected, such a reformulation of his position makes the physicalist vulnerable to a whole new cluster of objections. Let us note that the use of such so-called shortcuts and metaphors is, granted the feasibility of ontological reduction, as unwarranted and inefficient as describing the phenomenon of lightning in terms of the wrath of Zeus rather than in terms of electrical discharge.

---

[3] I am not going to pursue this problem here. Although I do find it dangerous to physicalism, it is not vital to my line of argument; there is a voluminous literature on the subject, a representative selection from which could be said to start from Nagel (1974) and end with Jackson and Chalmers (2001).

[4] Or at least one of them – the informal, familiar, subjective type. There seems to be no problem in providing an ontologically reductive analysis of the formal, scientific, objective type, but then it is possible to charge the physicalist with inconsistency for introducing the distinction in the first place.

Again, the physicalist cannot resort to introducing another distinction, e.g., the distinction between 'apparently physical' phenomena (insusceptible to abbreviated or metaphorical description) like electrical discharge and 'apparently mental' phenomena (susceptible to abbreviated or metaphorical description) like empathy or inspiration – such a move would involve either the by-now familiar problem of introducing irreducibly holistic categories of language (a dead-end for the physicalist) or a paradigmatic example of intellectual dishonesty (if it is known that the mental supervenes on the physical, and the former is not irreducibly holistic, then one should not use a different type of language or mode of description for each, unless one wants to engender confusion and obfuscate the truth).

Likewise, the physicalist is not allowed to distinguish between different ways of describing phenomena as serving different 'purposes' (whatever they might be), since then he has to fall back on the irreducibly holistic concept of purpose – and even if one grants that 'purpose' can be construed as one of the fundamental properties of a purely physical universe (perhaps as something akin in meaning to 'workability', i.e., the property of every physical state being able to cause other physical states, or something similar), on this view there is place for only one such purpose[5], which does not allow for the introduction of the aforementioned distinction.

Finally, one might attempt to secure for oneself the right to use the language of complex concepts (ethical, mental, emotional or otherwise), while at the same time denying that they are irreducibly holistic, by resorting to a form of ontological reduction that follows a certain taxonomic order, analogous to the one used in describing the evolutionary relationships within a given biological group. The underlying strategy is this – each complex phenomenon at a given level of organization can be described as supervenient on the phenomena being one level lower on the taxonomic tree, but it is important to note that the respective organizational levels must stand in direct relationship; that is, they must directly follow one another – an attempt to describe a given phenomenon as supervenient on the phenomena being two or more levels down the taxonomic tree will be correspondingly faulty and incomplete. Thus, for instance, DNA can (and should) be described as supervenient on genetic instructions, but not as supervenient on simple physical causation. Likewise, conscience and dignity can

[5] Still, I think that even such a diluted notion of purpose retains much of its irreducibly holistic character, and I would expect a consistent physicalist to unpack it in terms of an explicitly physical account.

(and should) be described as supervenient on relevant neuropsychological states, but not as supervenient on genetic instructions. The general upshot is that one is justified in using different descriptions for the same phenomenon without having to admit that any of them denotes an irreducibly holistic concept.

The model delineated above appears to be more appealing than the non-taxonomic physicalist frameworks considered so far, but when it comes to the alleged ontologically reductive explanations of complex ethical and mental concepts (or, for that matter, any concepts that the physicalist might agree to call 'not explicitly physical'), I see it as equally untenable. The point is that phenomena such as conscience and dignity do not seem to enter the aforementioned taxonomic hierarchy at any level whatsoever – the very introduction of such a hierarchy seems to be founded on a physicalistic basis which subsumes many organizational levels of phenomena (molecular motion, genetic replication, brain activity, etc.), but excludes many others (e.g., moral feelings, complex emotions). There is no question that there is a common underlying substance responsible for such phenomena as molecular motion, genetic replication and brain activity, namely physical substance, thus one can consider them all as branches of the same taxonomic tree. In the case of, e.g., dignity and conscience, however, one encounters an obvious discontinuity, an essential gap, whose denial would involve an article of unwarranted monistic faith.

An important thing to note is that the physicalist's problem with bridging this gap consists not so much in the need of accounting for the existence of subjective faculties in terms of their alleged supervenience on objective physical processes, but in finding a way to save the efficacy and practical significance of these faculties without having to admit that the are irreducibly holistic.[6] In the next section of the paper I shall concentrate on the difficulties surrounding attempts to find such a way, moving steadily towards the proposed conclusion that there is none and that concepts like dignity and conscience can do their work only under a robustly realist interpretation.

As an aside, if one wishes to treat dignity and conscience as primarily mental concepts, then the above considerations of ontologically reductive taxonomy clearly bring out the difference between having an irreducibly holistic character and being subject to what has been termed 'the anomalism of the mental' (Davidson [1980]).

[6] It is noteworthy that the subjective-objective distinction need not enter here at all, since many people consider their right to dignity or the prescriptions of their moral intuitions to be as objective as the fundamental laws of physics.

The latter is markedly similar to being subsumed under the kind of taxonomic order discussed in the preceding paragraphs – it claims that propositional mental states are grounded in the general principles of human psychology, which in turn are grounded in the underlying neural mechanisms and eventually in the fundamental workings of physics, but that does not mean that it is possible to make a direct leap from the realm of propositional mentality to the realm of basic physical causation – these two are simply too different and need to retain their differences in order to remain what they are. On the contrary, irreducible holism claims that in order to remain what they are, certain complex mental concepts precisely cannot enter the aforementioned taxonomic hierarchy (at any level whatsoever), since their emergent properties not so much do not have any law-like connections with their grass-roots constituents (although this is probably true as well), but are something essentially over and above these constituents (that is what defines their emergent character).

In short, for the time being, the taxonomic approach to ontological reduction appears to fare no better than its non-hierarchical counterpart in handling the language of complex, non-scientific concepts. Consequently, the physicalist ends up being impaled on either horn of the dilemma – asserting that physicalism is true, but there are some irreducibly holistic concepts, leads to a blatant contradiction; asserting that physicalism is true, but it is preferable to speak as if there were genuine irreducibly holistic concepts, leads to blatant intellectual dishonesty.

This concludes my discussion of preservative irrealism and the solutions it offers with regard to approaching the concepts characterized by the title doctrine. Since, as I see it, it turned out to be incapable of achieving its intended goal, the only remaining choice open to the physicalist is to, if possible, eliminate irreducibly holistic concepts from his conceptual framework. According to the classic statement of eliminativism, the development of the natural sciences (most notably neurosciences) will eventually displace the commonsense language of propositional mental states, such as beliefs, desires and judgments (Churchland [1981]). Given what I have said so far, a natural extension of the above mentioned linguistic transformation would be the elimination of such non-scientific ethical concepts as dignity and conscience; in the consistent physicalist's[7] ideal future,

[7] Since I have already assessed the programme of preservative irrealism as untenable, and at this point I see eliminativism about irreducibly holistic concepts as the only alternative to prevent the failure of physicalism, I shall henceforth use the term "physicalist" as a shorthand for "physicalist who embraces eliminativism about irreducibly holistic concepts" and the term "physicalism" as a shorthand for "physicalism based upon eliminativism about irreducibly holistic concepts".

they would have to be replaced by exhaustive descriptions of certain sets of physical states, compatible with ontologically reductive taxonomy discussed earlier.

How could such a future be brought about? Perhaps by appeal to the explanatory powers of mature natural sciences; it is plausible to assume that the physicalist could simply, and quite bluntly, declare to his opponent that what she terms 'conscience' is, admittedly, a complex emergent concept, but natural sciences can unweave its complexity and fully account for its emergence – given an exhaustive scientific analysis, there just remains no place left for its supposed irreducibly holistic, morally realist character.

The reply might go: "Well, so much the worse for exhaustive scientific analyses. Notions such as conscience and dignity derive their essence and indispensability precisely from their irreducibly holistic character. If you deny it from the outset, then you cease to talk about conscience and dignity; your subsequent analysis deals with some superficially similar, but in fact fundamentally different concepts, severely impoverished simulacra of their original counterparts".

Upon hearing something like the above, the physicalist could shrug his shoulders and perhaps legitimately ask why he should feel troubled by what he sees as nothing more than queer and unsubstantiated intuitions of his objectors. After all, he knows the truth and it is their loss if they refuse to accept it. The problem with such a reaction, however, is that what he sees as nothing more than clinging to a groundless fiction turns out to be a fact to be reckoned with. First of all, one might claim that unless the physicalist finds some way to eliminate the intuition of his opponent, he should consider himself defeated, since that intuition itself can be treated as irreducibly holistic. This, though, is perhaps not a very serious objection – it is not entirely implausible to assume that the physicalist could come up with a comprehensive, multi-level analysis that would explain such intuition in terms supervenient on relevant psychological phenomena, underlying neural occurrences, etc.

A much more substantial difficulty consists in the fact that explaining (or even explaining away) a given intuition does not entail explaining away its object; even if I can account for my friend's belief that there is a bandit behind the corner as a result of a childhood phobia, it certainly does not imply that there actually is no bandit behind the corner. Hence, even if I can show that one's intuition about dignity and conscience is not irreducibly holistic, it does not follow that dignity and conscience themselves are not irreducibly holistic.

Another route the physicalist might try to follow is that of brute-force persuasion; he might say to his objector: "Let us stop theorizing. Instead, I will just show you that there is no mysterious, integral whole called 'dignity' attached to your physical constitution. I shall proceed by performing an appropriate brain surgery, the result of which will be the disappearance of your misleading intuition and your acceptance of the truth that I have tried to convince you of. Thus, I shall prove that your self is nothing over and above your physical makeup, since manipulating the latter in an appropriate manner entirely determines the shape and content of the former, including its reflective beliefs and insights. You may see this operation as harsh, but if you consent to it, you will realize that up to now you have lived under a profound misapprehension." Let us assume that the objector agrees to undergo the surgery and, just as the physicalist anticipated, he subsequently recants his views and admits to having lived under a grave misconception. Should we now conclude that his opponent succeeded in proving his point?

Not necessarily. On the contrary, it seems that the situation brought about is precisely one of those in which the irreducibly holistic character of dignity manifests itself in full force. After all, it makes good sense to say that by bringing his interlocutor to the understanding that there is no mysterious, integral whole called 'dignity' attached to his physical constitution, the physicalist violated precisely his dignity, thus bringing out its robust significance.

An immediate protest might be: "But the operation was carried out only after his explicit consent; therefore, it is impossible to say that the fact that his physicality was tinkered with caused him any unwanted moral damage". This is hardly convincing, though. It seems reasonable to believe that there is a fairly broad class of things that we cannot do unto others even if they consented. Even if irreversible tampering with someone's brain does not obviously fall into this class (although I think it does), it is plausible to assume that some other actions do, which confirms the inalienable nature of dignity – it cannot be done away with at will. In fact, I find it quite probable that inalienability is a constant feature of all irreducibly holistic concepts.

At this stage, the physicalist could say: "This whole reasoning may well be correct, but only given the presupposition of irreducibly holistic dignity. This is not fair to me. When one accuses me of arriving at viciously circular inferences (demonstrating that irreducibly holistic dignity does not exist by means of encroaching upon it), one puts the cart before the horse. In reality it is my objector's reasoning that is disturbingly circular; first, he groundlessly asserts the existence of a certain uncanny object, and then, when I put forward an

experimental argument against that assertion, he claims that I actually just confirmed it. In other words, he maintains that by questioning or denying his otherwise unfounded theory, I unwittingly prove its correctness. Not much deliberation is needed to realize how shaky the above train of thought is."

I believe that this is a legitimate protest, but there is a good answer to it. The crucial point to notice is that with regard to accepting certain presuppositions, the physicalist and his opponent are on a par – the former's implicit endorsement of ontological reduction is very much like the latter's appeal to intuition about irreducibly holistic concepts. However, what perhaps gives an advantage to the latter's view is that it fits, and to a large degree constitutes, the framework of our ordinary commonsense moral beliefs and customs. On the contrary, embracing physicalism would require us to revise radically or abandon altogether the established conceptual structure of our moral theory and practice. The presupposition of irreducibly holistic concepts is justified insofar as they support and reinforce one another in an interlocking structure, whose self-sufficiency derives from the mutual dependence and complementariness of its elements.[8] As far as I can see, no similar justification is available for the thesis of ontological reduction.

Let us conclude the paper by granting that the above vindicatory asymmetry is irrelevant to the issue at hand and looking at what positive programme the physicalist can offer us in place of our age-old intuitions and related practices. To survey the said programme with sufficient caution and foresight, let us fast forward into the future and invoke a variation of the familiar 'experience machine' thought experiment (Nozick [1974] p. 42-45).

The machine is able to give one whatever experiences he wishes. Moreover, it can make him forget that the experiences he receives are illusory – as soon as he plugs in, his single wish can make the machine-generated world indistinguishable from the real world, thus eliminating any potential remorse arising from trading

[8] For a discussion of a related point, used to define holistic systems, see Michael Esfeld (1998). According to Esfeld, a holistic system can be analyzed in terms of "generic ontological dependence among the things which are its constituent parts in so far as these things instantiate some of the properties that make something a constituent of a system of the kind in question, given a suitable arrangement". He illustrates his view with examples of holistic systems such as systems of beliefs and social communities, but it seems to me that the framework of our ordinary commonsense moral theory and practice would apply to it equally well (think, for instance, about the dependence between such concepts as 'crime' and 'punishment', 'conscience' and 'empathy'). After all, it should not come as a surprise that a system that includes irreducibly holistic concepts is itself holistic. It is important to note, however, that holism (as understood by Esfeld) should be carefully distinguished from irreducible holism, since the former, unlike the latter, is neutral with regard to the prospects of ontological reduction.

reality for fiction. Finally, the machine is completely self-sufficient, immune to malfunction and able to sustain the life of its user for the period of the average human life expectancy, so questions such as "what happens if it breaks", "how long can one survive hooked up to it" and "who will operate it if everyone plugs in" should not arise at all.

Now, the choice for the physicalist is clear: either he plugs in or he does not. Since I presume that his wish is to present us with some account of a good and meaningful life, I am bound to conclude that one of the above options offers exactly that. However, since I shall also argue that, given the abandonment of irreducibly holistic concepts, neither of them allows for a coherent moral life, my other conclusion is that the physicalist has no viable proposals on offer.

To start with, it seems clear to me that, provided the assumptions that, first, the rational agency of human beings consists of nothing over and above the physical constitution of the brain and, second, it answers to no external, irreducibly holistic authorities (since there are none), the physicalist ends up in a situation where finding any plausible reason for not plugging in becomes alarmingly difficult. He cannot claim that he cares about the 'authenticity' of life, nor can he contend that, in addition to having certain experiences, it is essential to be a certain kind of 'person' (e.g., a person concerned with the 'dignity' of others), since both 'authenticity' and 'person' (not to mention 'dignity'), if they are not treated as mere fancy labels of certain complexes of experiential states[9], constitute paradigmatic examples of irreducibly holistic concepts. In fact, given the two assumptions listed at the beginning of this paragraph, together with the claim that a good life is a life of satisfied desires (excluding those that clearly harm others [even illusory others] and hinder the satisfaction of their desires) and the assertion that satisfying a desire is just a question of proper stimulation of brain-matter, it appears that a consistent physicalist has no choice but to hook up to the machine.

Plunging into the world of permanent illusion may be seen as an appalling perspective by many of those who were perhaps too quick to declare themselves physicalists, but if they want to maintain their views and remain consistent, then it seems to be the only perspective available. Is it, however, really available? One might rightly say that it is extremely difficult, if not outright impossible, to construct an illusory world in which literally all of one's desires are fulfilled, since it is almost sure that some of these desires are incompatible with others. For instance, one cannot hope that the machine will allow him to become both a Don

[9] If they are treated as such, then obviously the machine can produce them without any difficulty, hence ruling out their pursuit as an available excuse for not plugging in.

Juan and a Christian saint, though he might certainly entertain such simultaneous desires. In other words, the set of worlds available to the machine user is constrained by the limits of his imagination and by the need of coherence of his preferential makeup.

Thus, the best that the physicalist can hope for is one of the worlds that are maximally consistent with his overall system of desires. There is no reason to suppose that there is any such single world, since different worlds can be compatible with different parts of the said system. Hence the crucial question: how shall the machine user choose one of them; what criteria shall he use? In this situation, the only way that occurs to me of making a well-founded decision is to specify the value or the personality trait one attaches most weight to (e.g., respect, love, courage, erudition, etc.) and choose the world that, in addition to being one of those maximally consistent with his overall system of desires, is more conducive to the promotion of that particular value or personality trait than the rest of them.

Here, I believe, the physicalist reaches the ultimate dead end; the point, familiar by now, is that he denies the existence of irreducibly holistic concepts, but decides to live as if they existed, thus running afoul of intellectual honesty. What is the reason to claim that he treats his most cherished value or personality trait as if it were irreducibly holistic? It is the fact that he needs the machine to provide him with the essential belief that the illusory world he lives in is actually the real world; were this belief absent, he could not seriously contend that the surrounding phantasmal environment truly satisfies his most significant desires. Hence the implication that, even for our alleged physicalist, systems of robust evaluative criteria cannot be sustained without something more than purely physical experiences, and genuine moral needs cannot be met simply by means of stimulation of brain-matter. This concludes my final argument for the truth of irreducible holism set forth in this paper.

As a last aside, we could obviously imagine a machine-user who has no moral needs of the kind discussed above; for instance, his desires might be fixed solely on sunbathing, drinking expensive wines and pursuing bodily pleasures. The machine is clearly able to create the world that he would consider ideal, but I seriously doubt whether the majority of those who lean towards physicalism would share his preferences.

On a closing note, let me use the inferences drawn from our thought experiment to comment briefly on a position incompatible with irreducible holism, but fitting neither the label of eliminativism nor, as I see it, that of preservative irrealism. What I have in mind is Blackburn's quasi-realism (Blackburn [1998]

ch. 3). Arguing on the basis of the view that we project properties and qualities onto the objects of our thought and experience, Blackburn suggests that we tend to see these objects through the prism of ostensible, self-made facts. Consequently, he claims that in many contentious areas (e.g., aesthetics, morality, modality, etc.) we are entitled to treat our discourse within the realist framework, even though it is in fact ultimately based on our subjective projections. Hence, it seems that a physicalistically inclined quasi-realist could issue a challenge to irreducible holism by arguing that perhaps we cannot help attributing an irreducibly holistic character to phenomena such as dignity and conscience, but that does not imply that they do actually possess such a character.

I believe that the challenge can be met by reflecting a bit more thoroughly on our experience machine story. Let us note that the existence of such a machine would amplify the tools of the quasi-realist to an immense extent; in fact, it would carry his programme to its very limits, since it would allow for literally transforming his projections into reality (at least experientially speaking, but that is the most the physicalist can and should expect). Hence, given that using the machine to generate one's ideal world exhausts the demands of the quasi-realist and that making that world capable of satisfying one's genuine moral needs (still) requires at least an implicit acceptance of irreducible holism (as follows from our original thought experiment), quasi-realism turns out to be neither an alternative nor a threat to the title doctrine. This concludes my discussion.

To sum up, I have tried to argue that the affirmation of ontological reduction, seen by many as a potent tool for better and more precise understanding of reality, might plausibly lead to a dangerous distortion of our deliberative and moral attitudes, eventually leaving us in the cul-de-sac of eternal stagnation. By their nature, the presented arguments aim at highlighting the shadows of a whole family of views, including smug scientism, many forms of naturalism[10] and all forms of monism.[11] My alternative is ontological pluralism, an

[10] I should note that I find the term 'naturalism' ambiguous and misleading, especially because it is often grounded in what I see as an artificial and semantically confused distinction between the 'natural' and the 'supernatural'. To my understanding, the semantic difference between the natural and the supernatural is that the latter is not exclusively natural (it is more than natural). So the supernatural is natural, but not purely so. A term denoting the opposite of natural is thus not 'supernatural', but 'non-natural', though I confess I remain unsure as to what it is supposed to mean.

[11] See Crane (2000) for an argument in favour of distinguishing between physicalism and what the author calls „physical monism". The author contends that the essential characteristics of physicalism include not only its denial of dualism, but also "the epistemological and ontological authority it gives to physical science". Physical monism, on the other hand, is taken to be the less demanding conviction that all particulars are exhaustively physical. Since, however, I consider my arguments to address both of these claims, I believe that agreeing with the distinction in question makes no substantial difference to the conclusions of the present paper.

open-minded acceptance of the variegated nature of reality, which I consider to be the only view fully compatible with all the diverse commitments of different areas of our theoretical and practical activity. If adopted under a robustly realist interpretation, it should clarify and deepen our understanding of phenomena falling under a number of disciplines, ranging from moral psychology to the philosophy of science, as well as help us avoid seeing conflicts between them where there are in fact none.

References

Chalmers [2006] – D. Chalmers, *Strong and Weak Emergence*, [in:] P. Clayton, P. Davies (eds.), *The Re-emergence of Emergence*, Oxford University Press, 2006, p. 244-256.

Blackburn [1998] – S. Blackburn, *Ruling Passions: A Theory of Practical Reasoning*, Clarendon Press, Oxford 1998.

Churchland [1981] – P.M. Churchland, *Eliminative Materialism and the Propositional Attitudes*, "Journal of Philosophy" (78) 1981, p. 67-90.

Crane, Mellor [1990] – T. Crane, D.H. Mellor, *There is no Question of Physicalism*, "Mind" (99) 1990, p. 185-206.

Crane [2000] – T. Crane, *Dualism, monism, physicalism*, "Mind and Society" (2) 2000, p. 73-85.

Davidson [1980] – D. Davidson, *Mental Events*, [in:] *Essays on Actions and Events*, Clarendon Press, Oxford 1980.

Esfeld [1998] – M. Esfeld, *Holism and Analytic Philosophy*, "Mind" (107) 1998, p. 365-380.

Horgan [1993] – T. Horgan, *From Supervenience to Superdupervenience: Meeting the Demands of a Material World*, "Mind" (102) 1993, p. 555-586.

Jackson, Chalmers [2001] – F. Jackson, D. Chalmers, *Conceptual Analysis and Reductive Explanation*, "The Philosophical Review" (110) 2001, p. 315-361.

Nagel [1974] – T. Nagel, *What is it like to be a bat?*, "The Philosophical Review" (83) 1974, p. 435-450.

Nozick [1974] – R. Nozik, *Anarchy, State, and Utopia*, Basic Books, New York 1974.

Poland [1994] – J. Poland, *Physicalism: The Philosophical Foundations*, Oxford University Press, Clarendon Press, Oxford 1994.

Stalnaker [1996] – R. Stalnaker, *Varieties of Supervenience*, "Philosophical Perspectives" (10) 1996, p. 221-241.