

# The Conceptual Inexhaustibility of Personhood

Andreas Kemmerling (安德里亚·凯莫林)\*

**Abstract:** Some leading neuro - scientists recently proclaimed an obviously false view that a human person is his/her brain. This falsity arises partly from the conceptual difficulties concerning personhood/a person. By revealing inexhaustible richness of the characteristics of this concept of a person, this essay explains why the concept is so utterly puzzling. The author contrasts Descartes' concept of a person with Locke's. For Descartes, the concept has four features: (1) it is the concept of the mind/body - union; (2) it is innate and a primitive (i. e. unanalysable) concept; (3) it is not clear - and - distinct (because it is primitive so that it cannot be reduced to clear - and - distinct concepts); (4) it is of enormous value (because it helps us to dissolve, not to solve, the mind/body - problem). Correspondingly, Locke's concept of a person has another four features: (1) it is the concept of an entity which is justly rewarded/punished for its doings (including mental doings); (2) it is neither innate (no so - called innate concepts in Locke) nor primitive, but rather a complex idea which is reducible to the concepts of consciousness and memory; (3) it is clear and distinct (because it is made of clear and distinct simple concepts); (4) it is of e-

---

\* Andreas Kemmerling, Professor, Department of Philosophy, Heidelberg University, Heidelberg (安德里亚·凯莫林, 海德堡大学哲学教授).

normous importance (not because it helps us to solve or dissolve the mind/body – problem, but because it is central to our conceptions of justice and self – care). Through the two examples, this article argues that the concept of a person indeed has inexhaustible richness that is historically caused because this concept is used by many thinkers as a conceptual tool or as a technical or semi – technical term for answering different questions in various distinct and incommensurable theoretical frameworks, so that quite different features have been introduced as characteristics of a person. Thus, we do not have a shared intuitive grasp of this concept, but only a common learned tradition which has bequeathed to us a blend of quite diverse conceptual features that were never meant to go together. The characteristic inexhaustibility of the concept of a person may result from the following reasons: (1) this concept is anthropocentric in its actual application; (2) it is not at all anthropocentric in its intension; (3) it is part of the concept of a person that persons are distinctively valuable.

**Keywords:** personhood/a person; brain; personality; Descartes; Locke; inexhaustibility

## “一个人格化者”概念的不可耗竭性

**摘要:** 最近一些顶级的神经科学家主张了一个明显错误的观点: 一个人类的人格化者 (a human person) 就是他或她的大脑。这个错误主张部分地源于“一个人格化者 (personhood/a person)”这个概念本身的困难, 因为很难去说清楚什么是“一个人格化者”。本文通过揭示“一个人格化者”这个概念的特征的无限丰富性去解释为什么这个概念是如此地让人头疼。本文对比笛卡尔和洛克两个大哲学家的“一个人格化者”的概念。笛卡尔的“一个人格化者”概念有四个特征: (1) 它是一个心灵—身体的统一体的概念; (2) 它是一个天赋的

和原始性（即不可分析性的）的概念；（3）它不是一个清楚分明的概念（因为它是如此原始的以至于不能被还原为清楚分明的概念）；（4）它是极其有用的概念（因为它帮助我们消解而不是解决了笛卡尔的身心关系问题）。相应地，洛克的“一个人格化者”概念对应有四个不同的特征：（1）它是一个因为其行为（包括心灵行为）而被奖惩的实体的概念；（2）它既不是天赋的观念（洛克没有所谓的天赋观念）也不是原始性的概念，而是一个可以还原为意识和记忆的复杂观念；（3）它是清楚分明的概念（因为它由清楚分明的简单概念所构成）；（4）它是极其重要的概念（不是因为它帮助我们解决或消解了身心关系问题，而是因为它是正义和利己概念的核心）。通过对比本文表明“一个人格化者”的概念拥有不可耗尽的丰富性，这个丰富性是由历史所造成的。因为这个概念被许多思想家作为概念工具或作为技术性或非技术性术语去回答那些在非常不同的甚至不可公度的理论框架下被提出的各种问题，从而许多不同的特征都被引入到这个“一个人格化者”的概念之中。我们对于这个概念其实并没有一个统一的直觉性的把握，而仅仅只有一个共同的传统，这个传统遗留给我们混杂着很多不同特征的“一个人格化者”概念，这些不同特征在过去从未打算被协调地整合在一起。“一个人格化者”概念的特征的无限性可能是由如下三个原因所造成：（1）这个概念在其概念应用上是以人类为中心的；（2）但是这个概念在其内涵上却完全不是以人类为中心；（3）这个特征 [人格化者的集体（persons）是极其有价值的] 也是“一个人格化者”概念的一部分。

**关键词：**一个人格化者；大脑；人格/性格；笛卡尔；洛克；不可穷尽性

The concept of a person is a vexing one.

There is ample evidence for this claim, both in time-honoured works and in recent publications. Before I concentrate on some of the old stuff, let me briefly turn to recent examples. The following sample of quotations from leading neuro-

scientists may illustrate how deep the confusion about what a person is can go among the educated, even for today. Francis Crick stated his *Astonishing Hypothesis* as follows:

“You” [...] are in fact no more than the behaviour of a vast assembly of nerve cells and their associated molecules. As Lewis Carroll’s Alice might have phrased it: “You’re nothing but a bunch of neurons.” This idea is so alien to the ideas of most people alive today that it can truly be called astonishing. <sup>①</sup>

A few years later, this ‘idea’ seemed not anymore astonishing to Michael Gazzaniga who prefers to put it this way: Some simple facts make it

... *clear* that you are your brain. The neurons interconnecting in its vast network [...] – – that is you. <sup>②</sup>

It required the brilliancy of a German professor of psychiatry and neuro – dactics to take it to a further extreme. He found a way to expand Crick’s and Gazzaniga’s point by enriching it with a homespun piece of congenial ludicrousness. In a German radio – broadcast in November 2006, Manfred Spitzer declared:

You don’t *have* your brain, you *are* your brain.

Maybe this is a world record. Is it humanly possible to display more fundamental confusion in less than ten syllables? One is almost inclined, with respect to someone who says such a thing, to believe at least the first part of his *dictum*.

Notice that in these three quotations we are addressed directly by use of the word “you”. As who or what could we consider ourselves so addressed (given that we are, in the same breath, straightforwardly identified with our brains)?

---

<sup>①</sup> Crick (1994: 3).

<sup>②</sup> Gazzaniga (2005: 31); italics mine.

Clearly not as human beings. Human beings aren't just brains. Almost all of them have one.<sup>①</sup> And some of them use it, before they make grand claims. Let's assume that this much is known even to those who would make, or agree to, such claims as the ones I quoted. It's unlikely that they simply confuse a human being with one of his or her organs.

So assuming that we are not addressed, in the statements quoted above, as members of the species *homo sapiens sapiens*, the question remains: As whom or what do Crick, Gazzaniga and Spitzer presume to address us, when they say "you"? Well, I guess, we are meant to be addressed as *persons*. What these neuro - scientists seem to want to say is:

You, the *person* you are, are your brain.

A human person is nothing but his or her brain? The negative answer is obvious again. And again it is known, by most, to be obvious. I shall not go into this once more.<sup>②</sup> Instead I shall address in the following a different—an aetiological—kind of question: How can it happen that people who can read and write (well, on second thoughts, I'd rather say: people who can write) get so confused as to identify persons with their brains? Part of the explanation seems to me to be this: Our very idea, or concept, of a person is utterly baffling. And I shall investigate some of the reasons why this is so.

---

① The pitiable exceptions include anencephalics, microcephalics, hydrocephalics, and some brainless adult human beings, occasionally mentioned in the literature, for whom there seems to be no scientific label yet.

② For arguments against the "thesis" of person/brain - identity, cf. Kemmerling (2000: 221—243). But let me warn you. You'll probably find nothing in this paper which you do not know anyway. Trying to point out the obviously obvious almost inevitably results in dull papers. What excuse is there for a philosopher to engage in this sort of business nevertheless? Well, as J. L. Austin once put it: "Besides, there is nothing so plain boring as the constant repetition of assertions that are not true, and sometimes not even faintly sensible" (Austin, 1962: 5).

\*

If the concept of a person is a vexing one, what is it that makes it vexing?

There are various ways in which a concept may perplex us. First, there are concepts which may strike one as inherently unthinkable – or, to put it less sloppily: concepts as such that the items which they purport to be concepts of seem unthinkable. Infinity may serve as an example. Second, there are concepts which are, or seem, analysis – proof in a very peculiar way. They are, or seem to be, innocent, well – functioning non – primitive concepts which we, as normal speakers, have fully mastered; moreover, we are perfectly in the clear about what we consider as their most important ingredients. Nevertheless, there is a conceptual ingredient which consistently resists our attempts to make it explicit. Knowledge is an example. It is fairly uncontroversial that knowledge entails truth, belief and justification, and it is also clear that knowledge is not merely justified true belief – but nobody has been able to pinpoint what else is required for knowledge. The concept of knowledge contains at least one component, that vexing ‘last bit’, which seems inexplicable. Third, there are concepts which are, or at least seem to be, paradoxical, although they appear to be well – functioning, some of them even indispensable, concepts. Take the concept of being uninteresting for example. It lends itself to the comparative and the superlative form. But isn’t the most uninteresting event of all times *ipso facto* an interesting one? I, for one, would be anxious to be informed about it. Or take the concept of a belief for example. One holds each of one’s beliefs to be true (this is what believing is, after all), but at the same time, a sane person believes that some of his beliefs are false. Or take truth itself for example. The so – called Liar – paradox is known and unsolved since ancient times: “What I hereby say is not true”. Or, for that matter, take any of those countless concepts for which a paradox of the *Sorites* type can be construed, like famously for the concept of a heap itself.

The conceptual difficulties concerning personhood seem to be of an altogether

er different kind. Prima facie, personhood is nothing inherently unthinkable; there's no problem with a deeply hidden conceptual 'last bit' (we'd be happy to get hold only of the uncontroversial first bits); and we have no compelling reason to think that the very concept itself is paradox – ridden. ①

On the one hand, the word “person”, as it is commonly used, seems to be not much more than a singular form of the word “people”; it serves to denote human beings like you and me. In normal conditions, as soon as we have recognized an adult human being, we have recognized a person; we don't need any extra information about special features of this particular human being in order to draw the ‘further’ conclusion that he or she is a person. In the absence of very weighty counter – evidence or of compelling reasons to withdraw judgment, the presumption, concerning any human being, that he or she is a person, is not only epistemically admissible or reasonable, but rather morally obligatory. ② The *application* of the concept of a person, in familiar standard cases, does not appear to involve problems which are harder than those involved in recognizing people: normal members of the human race.

But *the concept itself* is problematic. At least it is difficult to say, in plain words or, for that matter, more refined ones, what a person is – even given the most basic and austere sense of the word “person”.

---

① One may think that *person* clearly is a vague concept (*i. e.*, allows for border line cases) and that therefore at least a paradox of the *Sorites* type can be construed. But I am not sure about it. The sad fact is, I think, that *person* is a concept so extremely indeterminate that we cannot even definitely say whether it is vague or not.

② Note that a presumption is not just an assumption, however plausible. As Whately once put it magisterially: “According to the most correct use of the term, a ‘Presumption’ in favour of any supposition, means, not (as has been sometimes erroneously imagined) a preponderance of probability in its favour, but, such a *preoccupation* of the ground, as implies that it must stand good till some sufficient reason is adduced against it; in short, that the *Burden of proof* lies on the side of him who would dispute it” (Richard Whately, 1841: 120, original italics). For an attempt at an outline of a theory of presumption, see Scholz (1999: 148—159).

## Person vs. Personality: Separating the Ontological From the Psychological

I would like to give one word of clarification, because it seems necessary in the light of the best recent discussions concerning personhood I am aware of.<sup>①</sup> When I talk in the following of “the concept of a person”, or exchangeably of “(the concept of) personhood”, I do not have a psychological concept in mind. *Person*, as I shall consider it, is an ontological concept. It is meant, by me here, to pick out a special category of entities – a category which is worth considering when the question is raised: “What sorts of particulars are part of the ultimate furniture of the world as we know it?” As an answer, I’d mention, with no attempt at originality: physical bodies, fields of gravitation, events, abstract particulars (sets, numbers, propositions, and maybe others), and . . . persons. I do not mean to be making a big claim here, I am not saying that persons are particulars which *do*, in the final analysis, belong to the ultimate furniture of the world as we know it, *i. e.*, cannot be reduced to (combinations of) more basic particulars. I would simply like to rank them among those entities which should be considered carefully as candidates. (Descartes for example, as we shall see, considered them as candidates, but decided not to assign to them the ontological status of being irreducible to more basic entities.) – Now, that’s what I’d like to emphasize, the “ontological” concept of a person should be kept as pure and austere as possible. In particular it should be kept distinct from any psychological notion, however seemingly close, like, *e. g.*, the concept of a personality. A personality is, I take it, something a person *has* (and presumably it is not a particular, but some universal which, at least in principle, different persons may share; but even if personalities would have to be accepted as particulars, they’d be par-

---

<sup>①</sup> I am thinking here of authors like, *e. g.*, Bernard Williams, Robert Nozick, Derek Parfit, David Lewis and Martine Nida – Rümelin.



ticulars different in kind from persons). What I'm trying to draw your attention to is not that *person* and *personality* are distinct concepts (this is banal), but the less obvious point that the tight and rigid connections between these concepts run only in one direction. Personality conceptually requires personhood; but not *vice versa*.<sup>①</sup>

The sparse “ontological” concept of a person I shall consider in the following is psychologically neutral, or noncommittal, in a thoroughgoing way: It does not exclude, for example, the conceptual possibility of one and the same person's changing his or her personality abruptly and completely. Psychological similarity, continuity, or conscious self – accessibility over time is not a *conceptual* ingredient in personal identity. It is, indeed, a *factual* ingredient in the human persons – over – time we are acquainted with. And, indeed again, the absence of this ingredient may make us wonder whether we are really dealing with the same person. But what I am trying to bring to the fore is that there is a basic “ontological” concept of personhood which does not by itself compel us to deny personal identity in cases of abrupt and vast psychological discontinuity. That's what I mean by calling the concept psychologically noncommittal; it is as it were silent about these cases. In focusing on this basic concept, I don't mean to deny that there are other legitimate concepts of personhood —concepts which are ‘more psychological’, in the sense just adumbrated. I shall not be concerned with any of those.

Now one thing that is deeply vexing about personhood is this: Even the fundamental and utterly austere “ontological” concept of a person seems inexhaustibly rich. And it is quite unclear which of its features are core components and which are peripheral, which of its aspects should be considered as being fundamental and which as being derived. It is this deplorable phenomenon that I shall

---

① One may be tempted to assume that at least the more specific concept of a *human* person involves having a personality. But I am not so sure about this either; maybe only our concept of a *normal* (or a non – deficient) human person contains personality as a feature.

be concerned with in the following. I shall try, very briefly, to present evidence that it is a fact; I shall venture an explanation for it, which I shall try to support, less briefly, by two examples from the history of philosophy.

Consider the following random list of characteristics of personhood which have been emphasized by various thinkers as they have employed that concept in their theorizing.

A person, it is said,

1 is an individual capable of rationality;

2 is responsible for what it does;

3 has dignity;

4 is not a something ( “*quid*” ) but a someone ( “*quis*” );

5 is free;

6 is a unity of a body and a mind ( soul );

7 is anything to which words and actions of human beings are attributed;

8 is an intelligent agent, capable of a law, happiness, and misery;

9 is an end in itself and an object of respect;

10 is an entity to which both mental and physical properties can be ascribed;

11 is capable of treating others as persons;

12 is capable of verbal communication;

13 is conscious and self – conscious;

14 is capable of second – order intentionality ( in particular, is capable of second – order volitions which are a precondition of having a free will ).

Many of these features themselves do not seem conceptually less demanding than personhood; many of them are somewhat vague. Some of them may appear controversial. ( As to #11, for example, there are forms of autism, or so I am told, which disable people from treating others as persons; but we would not be ready to accept without reservation, I presume, that anyone who suffers from such

a disease is *ipso facto* not a person.<sup>①</sup>) Arguably, not all of these features go together. (For example, #10 is so wide that it seems to allow for persons who don't exemplify several of the other features.) Clearly, several of these features seem to be dependent on others and so, maybe, this list needs to be reduced. But even more clearly and most importantly for our purposes, there is nothing about this list which gives us reason to assume that it is complete. The list is heterogeneous and it is open; and for all we know, it is essentially open.

Therefore, on the one hand, personhood appears to be a straightforward matter; as a matter of fact, we can, in normal circumstances, tell a person from anything else with remarkable ease. On the other hand, we do not have a clear idea of what the crucial marks of personhood are. The features which come to mind when we think about it are too many as it were, to elucidate what we really mean by "person"; and we are prepared to admit that even more features may turn out to be conceptually relevant, as we keep on thinking about it. Moreover, there is no reason to think that the word "person" is ambiguous. It would be absurd to claim that the features listed above specify distinct meanings of the word. "Per-

---

① This is not to say that *person* is not a QS - concept, *i. e.*, a concept which essentially involves a certain standard of quality in the following sense: It is part of the mastery of such a concept that one acknowledges, concerning the items subsumable under it, that they can be classified according to how good they are as items falling under this concept. Roughly speaking, if *C* is an QS - concept, then it is fully mastered only by someone who has also mastered a family of concepts such as "an excellent *C*", "a good *C*", "a middling *C*", "a lousy *C*", etc. An example for such a concept would be *argument*; you don't really know what an argument is, as long as you have no idea of how to classify arguments according to their quality as arguments. But you may very well know what a logical proof is, without even being willing to classify such proofs as good or bad ones. So *proof* is not a QS - concept. Three more remarks on QS - concepts: First, they do not need to be evaluative themselves, although their mastery essentially requires the ability to draw value distinctions concerning the members of their extensions. Secondly, it is characteristic of the natural sciences (at least of the more fundamental ones, and clearly of physics) that their theoretical terms do not express QS - concepts. Third, QS - concepts are not reducible to concepts which do not involve standards of quality. I am not sure what to say about *person*. But I think it is an interesting question whether it is a QS - concept or not. If it is, or were, one, then it may be difficult to stick to the view (which I have taken here) that there is a 'psychologically non - committal' concept of personhood. That's why I am inclined to assume that *person* is not a QS - concept.

son” clearly is not like “bank” ( “ground beside a river” / “institution offering financial services” ). It is exactly the fact that “person” is *not* a homonym which makes the essential openness of any collection of its conceptual features an embarrassing richness.

How is this richness of the concept of personhood to be explained? One answer to this question is historical. Over the centuries, the concept has been used by many thinkers as a conceptual tool for answering quite different questions: metaphysical, theological, and moral; and in reaction to these problems, quite different features have been introduced as characteristics of a person. So the word “person”, for a very long time, has been a technical, or semi – technical, term in various quite distinct theoretical frameworks; and it has been used in these frameworks for the solution of various quite distinct theoretical problems.

I shall try to illustrate this by two examples from the history of philosophy which I take to be quite telling. I hope that they reveal some aspects of the complexity and heterogeneity of our inherited concept of a person which has been partly formed ( reformed and, arguably, deformed ) by thinkers like Descartes and Locke.

### Descartes’ Concept of a Person

Let us consider, as a first example, the use Descartes makes of the concept of a person. According to his metaphysics, any human being consists of two entities which are really distinct: the body and the soul. They are really distinct, because the body is a physical substance and the soul ( or mind ) is an immaterial substance, and these two substances could exist without each other. It should be noticed that what Descartes calls *areal* distinction between substances is not a factual separateness, but a possible one: two substances are really distinct if they are capable of being separated, “at least by God”.<sup>①</sup> In the *Sixth Meditation*

---

① AT 7: 78.

Descartes presents (the definitive version of) his famous proof that his body and his soul are really distinct. The crucial point of the proof is this. Descartes claims that one can clearly – and – distinctly think of oneself insofar as one is only a thinking thing and not a material (or extended) thing; and one can clearly – and – distinctly think of one's body insofar as it is merely an extended thing and not a thinking thing.<sup>①</sup> Whenever anyone can clearly – and – distinctly understand one thing apart from another, God could have created these things in that way. And this is to say: his or her soul and his or her body are really distinct things. One can exist without the other.

This is, for Descartes, a fact of metaphysics. But metaphysics is not everything there is in life, not even for Descartes. As he says during a conversation with the theologian Frans Burman: “A point to note is that one should not devote so much effort to the *Meditations* and to metaphysical questions, or give them elaborate treatments in commentaries and the like. Still less should one ... dig more deeply into these questions than the author [*i. e.*, Descartes himself] did; he has dealt with them quite deeply enough. It is sufficient to have grasped them once in a general way, and then to remember the conclusion. Otherwise they draw the mind too far away from physical and observable things, and make it unfit for studying them. Yet it is just these physical studies that it is most desirable for people to pursue, since they would yield abundant benefits for life.”<sup>②</sup> And in a letter to the Princess Elizabeth he puts this point as follows: “I believe that it is very necessary to have properly understood, once in a lifetime, the principles of metaphysics, since they are what gives us the knowledge of God and of our soul. But I

---

① The hyphenation of “clear – and – distinct” is meant to remind you that this is a technical term of Descartes'. An idea, or an perception, is *clear*, if it is vivid (like the idea of pain when you suffer from one); it is *distinct*, if it is sharply separated from all other ideas (as the idea of pain is not distinct, according to Descartes, since we have a tendency to mix up the sensation of pain itself with something painful in the cause of the pain). – But the term “clear – and – distinct” has a very special meaning for Descartes: it is reserved for those ideas about which it cannot be assumed, on pain of manifest absurdity, that they are misrepresentations.

② AT 5: 165.

think also that it would be harmful to occupy one's intellect frequently in meditating upon them. ”<sup>①</sup>

The metaphysical conclusion that our minds and our bodies are distinct entities is hard to bring into unison with how we experience ourselves. This conclusion is true; it is even shown to be absolutely certain by a metaphysical proof, Descartes insists. But he recommends leaving it at that. The way we experience ourselves, he concedes, is not as consisting of two distinct entities; rather we experience ourselves as the union of our soul and our body. But this union is, in reality, not something really existing. There is no third entity, over and above our body and our soul. <sup>②</sup> When it comes to taking stock of the *really* existing entities, then, strictly speaking, there are only the two substances of body and soul which are distinct however intimately they may be interrelated. So, in a sense, when we experience ourselves as a mind/body – union, the way we experience ourselves is not true to the metaphysical facts.

It is exactly this union of body and soul that Descartes denotes by the concept of a person. “Everyone feels that he is a single person [*une seule personne*] with both body and thought [*i. e.* , soul] so related by nature that the thought can move the body and feel the things which happen to it. ”<sup>③</sup>

But, as he makes it clear, particularly in his correspondence with Elizabeth, Descartes is prepared to concede that this way of experiencing ourselves as

① AT 3: 695.

② There are some attempts at terminological appeasement. In a letter from January 1642 to his follower Regius, a professor of medicine at the university of Utrecht who later caused severe trouble for him, Descartes recommended, as Regius' ghost – writer in his dispute with the Dutch theologian Voetius, the following formulations: “... human beings are made up of a body and a soul ... by a true substantial union [*perveram unionem substantialem*] ... If a human being is considered in himself as a whole [*homo in se totus*] ... he is a single *Ens per se*, and not *per accidens*; because the union which joins a human body and a soul to each other is not accidental to a human being, but essential, since a being without it is not a human being” (AT 3: 508). This is intended to sound soothing, but the plain fact remains: mind and body are distinct substances, while their union, even if a “true substantial” one, is not a substance.

③ AT 3: 694.

persons is not just due to some sort of negligence or other kind of avoidable mistake. He says, surprisingly, that among our primitive notions which are innate and “can only be understood through themselves”, there is not only the notion of body and the notion of mind, but also the notion of their union.<sup>①</sup> This is surprising, since —in the final analysis— there is, as we have just seen, nothing to which this notion applies in reality, and therefore the notion of a person is, metaphysically speaking, at least a misleading one. Whereas both the soul and the body can be conceived by the pure intellect, their union, Descartes says, “is known only obscurely by the intellect alone … but it is known very clearly by the senses.”<sup>②</sup> Let me add that this means for Descartes: Although we have very strong and vivid ideas of the senses concerning the union of the body and the soul, these ideas never amount to genuine knowledge, since our senses can *never* give us ideas which constitute knowledge, not even when they are clear (*i. e.*, strong and vivid). Genuine knowledge consists in the intellect’s perceiving clear – and – distinct ideas. It is only such clear – and – distinct ideas of the intellect that God guarantees to be true. However, to repeat, the ideas we have of the mind/body – union, Descartes insists, are not clearly – and – distinctly perceived by the intellect. So when Descartes says: The union of mind and body is “known very clearly by the senses”, we must not forget that the knowledge in question is at best second class knowledge, or strictly speaking, not knowledge at all. What we do have, when we experience ourselves as *persons*, is nothing but vivid ideas of the senses, but not clear – *and* – *distinct* ideas of the intellect.

As soon as the intellect, in a metaphysical effort, has brought the ideas both of body and of soul to clearness – and – distinctness, and has achieved the insight that body and soul are really distinct entities, it faces what we nowadays call Descartes’ mind/body – problem: How can there be a causal interaction between these two entities, one of them material, while the other immaterial? When Frans

---

① AT 3: 665.

② AT 3: 692.

Burman asked him in 1648: “But how can this be, and how can the soul be affected by the body and *vice versa*, when their natures are completely different?”, Descartes replied: “This is very difficult to explain; but here our experience is sufficient, since it is so clear on this point that it just cannot be gainsaid.”<sup>①</sup>

So here is why the use Descartes makes of his concept of a person is important for him: Although we have no clear – and – distinct idea of a person, this idea is a primitive innate notion which cannot be reduced to clear – and – distinct notions. It is as persons that we experience ourselves quite naturally, as long as we do not philosophize about our nature. As long as we experience ourselves in this natural way, the mind/body – problem simply does not arise.

That is why people who never philosophize and use only their senses have no doubt that the soul moves the body and that the body acts on the soul. They regard them as a single thing, that is to say, they conceive their union; because to conceive the union between two things is to conceive them as one single thing. Metaphysical thoughts, which exercise the pure intellect, help to familiarize us with the notion of the soul; and the study of mathematics ... accustoms us to form very distinct notions of body. But it is the ordinary course of life and conversation, and abstinence from meditation ... that teaches us how to conceive the union of the soul and the body.<sup>②</sup>

Descartes seems to suggest here, and in other passages,<sup>③</sup> that metaphysics (pure thinking, performed by employing clear – and – distinct notions of the intellect) does not and cannot give us the solution to the mind/body – problem. The

① AT 5: 163.

② AT 3: 692.

③ *E. g.*, in a letter to Arnauld (July 29, 1648), where he writes: “That the mind, which is incorporeal, can set the body in motion is something which is shown to us not by any reasoning or comparison with other matters, but each and every day by the surest and most evident experience [*certissima & evidentissima experientia*]. It is one of those things which are known by themselves and which we only make obscure when we try to explain them” (AT V 222).



way to deal with this problem is rather to *dissolve* it, by recognizing that it simply does not arise as long as we experience ourselves in the way which is most natural for us: as persons. So the concept of a person is used by Descartes as a philosophical tool for the dissolution of a problem—indeed, a mystery— arising in his metaphysics.

According to the Cartesian account, the concept of a person is not a concept which could help us to gain metaphysical insights into the ultimate structure of reality. It is not clear – and – distinct; it is not one of those concepts by which we can reach genuine knowledge. In the letter to Elizabeth from which I have quoted extensively, Descartes says: “It does not seem to me that the human mind is capable of forming a very distinct conception of both the distinction between the soul and the body and their union; for to do this it is necessary to conceive them as a single thing and at the same time to conceive them as two things; and *this is absurd.*”<sup>①</sup>

Taking all this together, I suggest that Descartes’ thought is this: When you do metaphysics, when you inquire into the ultimate structure of what there is, you are bound to accept that your soul and your body are really distinct; and then, as long as you are engaged in nothing but pure metaphysics, you cannot conceive of yourself as a person (*i. e.*, of the union of your body and your soul). Strictly metaphysically speaking, this is not just too difficult; it would be simply absurd. At the end of the day, the concept of a person is not just confused, but also in principle so for a simple reason: The very concept is in tension with an irrefutable metaphysical fact. Nevertheless, this concept (which God was kind enough to put into our souls) is of enormous value. It captures an important aspect of our worldly existence, “which everyone invariably experiences in himself without philosophizing.”<sup>②</sup>

Let me list a few salient features which are characteristic of Descartes’ con-

① AT 3: 693; my italics.

② AT 3: 694.

cept of a person as I have just sketched it:

- (1) The concept of a person is the concept of the mind/body – union;
- (2) This concept is innate and a primitive, *i. e.* unanalysable, concept;
- (3) It is not clear – and – distinct; and since it is primitive, it cannot be reduced to clear – and – distinct concepts. So we may say that it is essentially not clear – and – distinct;
- (4) Nevertheless, it is of enormous value. Not because it helps us to solve the mind/body – problem, but because it helps us to dissolve it.

### **Transtemporal Personal Identity: a Blank Area in Descartes' Metaphysics**

Assuming for a moment that this sketch of Descartes' doctrine is on the right track, there is little wonder that he never cared to raise questions of transtemporal personal identity. One wonders why Descartes, otherwise a most subtle thinker on topics concerning the metaphysics of the mind, was apparently never puzzled by the problems about fission and fusion, the body – hopping of minds (or the mind – hopping of bodies, if that makes a difference) and all that kind of weird stuff which seems to spring immediately from his substantial mind/body – dualism.

So why was Descartes, of all thinkers, never puzzled by these questions which have occupied metaphysicians ever since Locke's *Essay*, and which seem to be taken bitterly seriously in recent metaphysics—indeed, today seem to be considered more urgent and important than ever? A tempting answer goes as follows: Because, for him, these are all pseudo – problems. A problem which wears its *principled* insolvability on its sleeves is a pseudo – problem. To put it in a bunch of slogans: “There's no *puzzle* of transtemporal personal identity. If the relevant questions could be framed at all, they could be framed clearly – and – distinctly; and then they could be answered. But they can't be framed clearly – and – dis-

tinctly, since they essentially involve the concept of a person.<sup>①</sup> A question which *on principle* cannot be phrased clearly – and – distinctly is a pseudo – problem; it simply has no answer.”

This, I gather, was not Descartes’ reason for avoiding issues of transtemporal personal identity. The problems in question would be pseudo – problems for him only if the concept of a person were a ( “materially” ) false idea, *i. e.*, one which is “such as to provide subject – matter for error”<sup>②</sup> by not representing anything real, but representing what they represent as something real.<sup>③</sup> But *person* is not a false idea. What it represents is something real ( the mind and the body as a union ), so whatever is wrong with it is not that it represents something as real which is not real. What is cognitively inferior about it, in comparison to concepts like *mind* and *body*, is that it essentially represents its *repraesentatum* indistinctly ( or as Descartes would put it: “*confuse*”, which is his technical term for the opposite of “distinctly” ). Yet this, by itself, is not a stain on its conceptual credentials. For its rationale is exactly to represent two – things – considered – as – one. Its appropriate realm of application is *outside* metaphysics. ( Within metaphysics, mind and body demonstrably are to be considered as two distinct things. But as I said: Metaphysics is not all that there is in life, not even for Descartes. )

For Descartes, the concept of a person is a fine concept, for *the conduct of life*. It is of utmost importance within this realm. It is a concept we live by ( if this is English ). And it would betray a grave intellectual misunderstanding to sneer at it because of its lack of distinctness. A concept’s lack of distinctness is not, *per*

① All these puzzling questions ( *e. g.*, “ Would somebody, let’s call him E. P., who enters, on Earth, a Parfitian Teletransporter be the same *person* as the one who, on Mars, leaves the teletransporter, given that the brain and the body in the cubicle of the Earthian Teletransporter were destroyed in due time? ”, “ If E. P. were teletransported twice over and subsequently destroyed, would any of the two duplicates be the same *person* as E. P. ? ” ) involve the concept of a person *essentially* — *i. e.*, they could not be rephrased without this concept.

② AT 7: 231.

③ AT 7: 44.

se, a conceptual deficiency. This sort of lack is the hallmark of many perfectly good concepts. In fact, the vast majority of the concepts on which we have to depend in order to lead our humble human lives are indistinct in not separating their bodily and their mental components: hunger, thirst, love, pain, sweet, soft, red – to mention but a few.

Nevertheless, *person* is merely a second – class concept when it comes to *the contemplation of truth*.<sup>①</sup> The contemplation of truth to the conduct of life is like a move in a game of Blitz chess to its analysis without time – limit. A perfectly good move in one context may not live up to the standards of the second.

Therefore, given that the concept of a person can, on principle, not be brought to distinctness, questions about transtemporal personal identity, for Descartes, are fated to imperfect answers (all of them, not only those bizarre cases which are characteristic of our contemporary debate). No answer could possibly possess genuine certainty. True knowledge, *scientia* in the emphatic Cartesian sense, is restricted to the realm of our most clear – and – distinct thoughts. A important philosophical fact about transtemporal personal identity is that no knowledge *sensu stricto* is to be had on the topic – and that therefore, in a sense, personal identity is not a metaphysical topic at all. The only ‘knowledge’ that could be hoped for would be epistemically second – class, knowledge merely “in the moral sense [*moralis sciendi modus*] which suffices for the conduct of life”.<sup>②</sup>

In a nutshell, Descartes’ view might well have been that the questions of transtemporal identity aren’t pseudo – problems, but neither are they questions to

---

① For the Cartesian distinction between the conduct of life and the (metaphysical) contemplation of truth see AT 7: 149.

② AT 7: 475. In this passage of his *Seventh Replies* to Bourdin’s objections, Descartes adds: “I frequently stressed that there is a *very great difference* between this type of knowledge and the metaphysical knowledge ...” (AT 7: 475; my italics). The very great difference lies in the following: Only metaphysical knowledge has God’s truth – guarantee; He would, *per impossibile*, have to be a deceiver, if our alleged metaphysical knowledge turned out to be false beliefs. But concerning our alleged knowledge in the mere moral sense, God’s benevolence does not guarantee the truth of what we believe. Moral certainty inextricably contains an element of epistemic risk.

which a philosophical answer could be given. We would have to try to find answers (or rather, practical decisions about how to deal with the situation), if we were confronted, in practice, with a problem – case. For such cases, we could not bring to bear moral certainty and not even practical knowledge [*connoissance en pratique*], since the latter would at least require a firm habit of belief,<sup>①</sup> which we could not have acquired concerning novel and extravagant situations (body – hopping of souls, etc.). Our guidance would have to be good common sense [*sens commun bon*;<sup>②</sup> *sensus communis*, in the non – technical sense<sup>③</sup>], which Descartes mentions occasionally, but does not theorize about.

Now suppose we were to actually confront such a case, *e. g.* one in which “the Soul of a Prince, carrying with it the consciousness of the Prince’s past Life, enter [s] and inform [s] the Body of a Cobbler as soon as deserted by his own Soul”<sup>④</sup>, and had to face the question whether the cobbler now is the same person as the prince. From a Cartesian point of view, no answer could be given with certainty, not even with moral certainty.

A narrow – minded straight application of the criterion for transtemporal personal identity suggested by Descartes’ concept of a person would yield the negative answer: No, the cobbler – now is not the same person as the prince – then, for personal identity, according to Descartes, obviously would have to be identity of the mind/body – union; and the prince’s mind and the cobbler’s body clearly constitute a union very different from the prince’s original mind/body – union. But the strategy of, *first*, concluding that the cobbler is not the ex – prince and *then* drawing whatever consequences from this result as if it were a theorem proven, presumably would not be what our good common sense recommends.

It would display more common sense to take into account what concrete prac-

① AT 4: 296.

② AT 11: 386.

③ AT 10: 518, 527.

④ *Essay*, II. 27. 15.

tical consequences are at issue. ( For example, is there a large sum of money the prince – then owes to somebody, and are we facing now the question whether the cobbler – now or the prince’s wife should pay the debt? Or is the question whether the cobbler – now ought to be hanged for a crime, committed by the prince – then? etc. ). Get clear about what, *in concreto*, is at issue in this particular situation, and in the light of this and of all that you know, if only with moral certainty; and discern the best solution<sup>①</sup> to this concrete problem with all of its contingent features. This may sound convoluted, as a piece of advice delivered by common sense. But then again, common sense may be more refined than the scoffers would concede. Its maxims may not be confined to what can be expressed in six – word sentences without hypotaxis. Descartes thought very highly about common sense— where it belongs. For him, it indispensably belongs to all the matters, where problems of personhood are concerned.

Let’s turn to something else. It is worth emphasizing that, for Descartes, mental identity is not sufficient for personal identity. Descartes has not explicitly formulated a criterion of transtemporal personal identity, but given his concept of personhood, his doctrine would obviously yield the following:

Person *A*, at *t*, is the same person as *B*, at *t*’, if and only if (i) the mind of *A* at *t* is the same mind as the mind of *B* at *t*’ and (ii) the body of *A* at *t* is the same body as the body of *B* at *t*’.<sup>②</sup>

It is a common mistake to assume that Descartes is implicitly committed to a purely mental criterion of personal identity. The reason for this mistake, presumably, is this: According to the Cartesian doctrine, I could exist without the body I happen to have; I could even exist without a body; but I could not exist without my mind; and this is to say, my essence is my mind and nothing physical is part of my essence. Therefore, if one’s mind is one’s total essence, then mind – identi-

---

① Or rather, “one of the best solutions”, for there may be more than one optimal solution.

② Note that transtemporal body – identity need not be strict “atom – to – atom” identity.

ty is that which (completely) constitutes personal identity.

But this last step is *anon – sequitur*: More specifically, it is a fallacy of equivocation. For Descartes, there are two ways of using the word “I”. If it is used, as almost always, in the common, “personal” (or “total”) way, it refers to the speaker as a mind/body – union. But if it is used in the “philosophical” (or “precise”) way—as Descartes’ thinker frequently does in the *Meditations*, then the user refers to himself as an entity whose existence he can prove with utmost certainty. Descartes sometimes cares to distinguish between these two uses by applying phrases like “*ego totus*”<sup>①</sup>, in contrast to “*ego quem novi*”.<sup>②</sup> The metaphysical result that my mind is my complete essence is a truth exclusively in the second sense of “my”. From this nothing can be inferred to the effect that my mind is my complete *personal* essence.

Descartes is committed to the criterion for human personal identity just mentioned (same mind & same body). But how this criterion would have to be applied to the enormous variety of bizarre possibilities discussed as problems of transtemporal personal identity, is a matter about which he, at least in published writings, simply remained silent. For this, as we have seen, he may have had very good reasons: first, these problems do not have a strictly philosophical or otherwise *a priori* justifiable answer; second, as long as we do not encounter these problems, there is no practical reason for dealing with them; third, as long as we do not know the practical consequences of our answers, there is not much which could guide our good common sense when we attempt to come up with an answer. Common sense is all that we could rely on in such cases.

## Locke on Personal Identity and Personhood

Let’s turn to Locke. For Locke, the concept of personal identity is an impor-

① *E. g.*, AT 7: 81, where he adds “insofar as I am composed of a body and a mind”.

② See for example in AT 7: 27.

tant one because the justice of all reward and punishment, whether performed by us or performed by God, depends on whether the one who did it is the same person as the one who is rewarded or punished. Our best clue of what *wereally* consider personal identity to consist in does not come from metaphysics ( “the same immaterial thinking substance” ), physics ( “the same material body” ), or biology ( “the same human being” ) but rather from how we proceed in applying our laws. The fact that we do not punish ( and would not consider it just to punish ) “the *Mad Man* for the *Sober Man*’s actions, nor the *Sober Man* for what the *Mad Man* did”<sup>①</sup> is of utmost importance. For Locke, this shows that when serious practical decisions need to be made, we treat the sober man and the mad man as different persons ( his actual wording is “thereby making them two Persons” ). If they are two persons, this is so in spite of the fact that, physically speaking, they are ( approximately ) the same body, in spite of the fact that they are, biologically speaking, the same human being and in spite of the metaphysical presumption that their immaterial thinking substance is one and the same.

This fact about what we consider just ( namely, not punishing somebody, *e. g.*, the sober man, who is physically, biologically and mind – substantially identical with the wrong – doer ) is, I suggest, Locke’s primary observation about transtemporal personal identity. Certainly, his conclusion ( concerning personal non – identity ) is not inevitable. ( We may, instead of jumping to Locke’s conclusion, prefer to say that in certain circumstances we do not punish the very person who did the deed. ) But I am not concerned here with the feasibility of Locke’s theory, but with trying to bring to notice what sort of problem he is actually dealing with, how the concept of personhood is meant to be serviceable to a solution, and what is supposed to motivate the proposed solution.

If I am right, Locke’s primary target – concept was transtemporal personal identity, not personhood. Where do we actually, and in a serious and responsible manner ( not just in the context of idle metaphysical speculation ), apply this con-

---

① *Essay*, II. 27. 20.



cept? What is characteristic of this particular sort of application? The answers to these questions pave the way to our best understanding of what personal identity is. Once we have reached such an understanding, the subsequent clarification of a fitting concept of personhood will be light work. Two equations have to be solved in the right order. So I suggest the following as the Lockean agenda:

1. Personal identity = that relation, whatever it is, that makes it just to reward/punish someone for something that was done in the past
2. Person = that entity, whatever it is, which is a proper relatum of this relation

Famously, Locke offers consciousness (or more specifically, conscious memory) as the solution for the first equation. The rough idea is this: Person *A*, at *t*, is the same person as *B*, at *t'*, if and only if *B*'s consciousness at *t'* could contain a memory of an action consciously performed, or a thought (consciously) had, by *A* at *t*. If *A* committed crime *c*, then the relation between *A* and *B* which makes it just to punish *B* for *c* is *B*'s (potential) memory of having done *c*—or somewhat more complicated: *B*'s being able to remember a thought *θ* such that *θ* was a thought of *A* at *t* in virtue of which *A* was conscious of committing *c*. The crucial point is that the relation in question is a psychological relation obtaining between conscious states: one particular conscious state *θ* of *A* at *t*, *e. g.*, *A*'s awareness of doing *c*, and another particular conscious state of *B*, *θ'*, which is *B*'s memory of *θ*. Given this psychological relation, *A* and *B* are “by the same consciousness ... united into one Person”<sup>①</sup>.

Equally famously, in solving the first equation, Locke starts with what he presents as an uncontroversial specification of personhood:

... what *Person* stands for ... , I think, is a thinking intelligent Being, that has reason and reflection, and can consider it self as it self, the same

<sup>①</sup> *Essay*, II. 27. 10.

thinking thing in different times and places ...<sup>①</sup>

What is needed, for a solution of the second equation which is satisfactory in the light of the proposed solution of the first equation, is a close connection between this concept of a person and the concept of consciousness (which is all that constitutes transtemporal personal identity). Locke makes the desired connection as close as possible: As consciousness “unites” persons over time, it unites simultaneous mental states into the same person’s mental states.

[consider it self as it self] ... which it does only by that consciousness, which is inseparable from thinking, and as it seems to me, essential to it: It being impossible for any one to perceive, without perceiving, that he does perceive. When we see, hear, smell, taste, feel, meditate, or will anything, we know that we do so. ... For since consciousness always accompanies thinking, and 'tis that, that makes every one to be, what he calls *self*; and thereby distinguishes himself from all other thinking things, in this alone consists *personal Identity*, *i. e.* the sameness of rational being.<sup>②</sup>

Beneath the surface of Locke’s account of personal identity, both of the momentary and the transtemporal sort, something is at work which deserves our attention. Locke has a strong dislike for the concept of a substance. He scolds traditional philosophy for “the promiscuous use of so doubtful a term”<sup>③</sup>; in using the word “substance”, he says,

... we talk like Children; who, being questioned, what such a thing is, which they know not, readily give this satisfactory answer, That it is *something*; which in truth signifies no more, when so used, either by Children or Men, but that they know not what; and that the thing they pretend to know, and talk of, is what they have no distinct *Idea* of at all, and so are perfectly

---

① *Essay*, II. 27. 9.

② *Essay*, II. 27. 9.

③ *Essay*, II. 13. 18.

ignorant of it, and in the dark.<sup>①</sup>

So there is at least one completely different sort of *desideratum* for Locke's account: A person should not turn out to be a substance of some sort. It is this, I presume, that makes consciousness so irresistible to Locke: The concept of consciousness, for him, is not the concept of a substance, neither a material nor an immaterial substance which, allegedly, is the permanent, indivisible underlying *substratum* of all mental activities. (Whereas Descartes takes the concept of a substance to be metaphysically inevitable and crystal clear, but the concept of personhood to be essentially obscure and merely practically helpful, Locke takes the concept of substance to be metaphysically inevitable and hopelessly obscure, but the concept of personhood to be perfectly faultless.)

Let's take stock of some of our findings in Locke; for the sake of perspicuity I arrange them in an order which gives us the Lockean echo to the four Cartesian tenets listed above:

(5) The concept of a person is the concept of an entity which is justly rewarded/punished for its doings (including mental doings);

(6) This concept is neither innate (there are no such concepts, according to Locke) nor primitive, but rather a complex idea which is reducible to the concepts of consciousness and memory;

(7) It is clear and distinct, since it is made of clear and distinct simple concepts; (Whereas our use of the word "person" creates obscurity.)<sup>②</sup>

(8) It is of enormous importance. It is not because it helps us to solve or dissolve that matter, the mind/body – problem, but because it is central to our conceptions of justice and self – care.

It would be rash to explain the remarkable clash between these theses and

① *Essay*, II. 23. 2.

② See *Essay*, II. 27. 28. It is exactly this alleged linguistic obscurity ("ill use of Names") which makes it necessary for Locke, following a suggestion of Molyneux', to include a separate chapter on these topics into the second edition (1694).

those listed above as a manifestation of the fact that the two thinkers are not really addressing the same topic (*i. e.*, the same concept of personhood). What I'd rather suggest is something else, namely that the concept of person, in virtue of its indeterminate richness, lends itself to wildly different accounts; and that the accounts which have been developed by various influential thinkers, sometimes within incommensurable theoretical frameworks and inspired by unrelated philosophical motivations, additionally have left discordant marks on what we, today, dubiously consider as our 'intuitions' about personhood.

As I said, and as the two examples in the excursion are meant to demonstrate, the word "person", for a very long time, has been a technical, or semi-technical, term in various quite distinct theoretical frameworks, and it has been used in these frameworks for the solution of various quite distinct theoretical problems. Moreover, again for a very long time, the word "person" has been in common use as a non-technical term which is not connected to any particular theory or problem, but which has nevertheless surreptitiously incorporated in its meaning an indefinite amount of the semantical complexity just indicated. Maybe what manifests itself as abundant richness inherent in our concept of personhood is only a reflection of the fact that we do not have a shared intuitive grasp of it, but only a common learned tradition, which has bequeathed to us a blend of quite diverse conceptual features that were never meant to go together.

### **An Immanent Conceptual Discordancy**

Even if this is true, there may be another explanation for the conceptual richness. It has to do with a certain tension right at the core of our concept of a person:

- (1) The concept of a person is anthropocentric in its actual application. The only clear cases of persons we are *familiar* with are human beings.
- (2) It is not at all anthropocentric in its intension. The concept of a person is not supposed to be the same as the concept of a human being. There is concep-

tual leeway both for the possibility of non – personal human beings (members of our species lacking exactly those features which are constitutive of personhood) and for the possibility of non – human persons. Any kind of creature could be, or could turn out to be, a person, if it only had that special something, whatever it is, that makes us persons. Fairy tales, novels and movies keep reminding us of this non – anthropocentric aspect: Hauff’s stork is a person, Shelley’s monster is a person, Mathison’s E. T. is a person, we are pretty sure that some of the replicants in Dick’s *Bladerunner* are persons, and we are supposed to wonder whether Clarke&Kubrick’s computer HAL is a person. If we try to specify what this conceptual leeway comes to, presupposing as we should that any normal human being is a person, we might look at the following two identifications:

Personhood = that, whatever it is, *without* which a common human being would be only biologically speaking a human being

Personhood = that, whatever it is, *with* which any being whatever would be, at least, of the same standing as a common human being

These equations may look funny at first sight, but I gather that they capture an important aspect of our concept of a person. And they may explain the embarrassing conceptual richness we found vexing: The list of features by which these two equations can be “solved” may be essentially open.

The first equation makes it quite clear that “common human being” is not to be taken in a biological sense. It is an honorific term *forus* (who happen to be common human beings) and for every possible being which is of the same standing. It is built into the very concept of a person that there is something valuable about common human beings (an accidental feature which each of them may lack) in virtue of which they are, as it were, not *merely* members of the human race. So we should consider the following as another conceptual core fact about personhood:

(3) It is part of the concept of a person that persons are distinctively valuable.

since there is so much about us which can be considered specifically and distinctively valuable, this again may explain why the concept of a person is inexhaustibly rich.

\*

So much about my rather sketchy attempt at a diagnosis of what is vexing about the concept of a person and what may account for its characteristic inexhaustibility. Now what would have to be done in order to tidy up a bit the conceptual mess? I call it a mess because (as a result of its richness) we have too many ‘intuitions’ about personhood and almost nothing to give them a structure. There are too few universally accepted constraints on this concept in order to make it possible to accept some of our *a priori* assumptions as valid and central and others as questionable or peripheral. I suspect that something like conceptual *analysis* is not what we need in order “to regain a complex concept of human personhood”. If we just stare at the concept and brood over its richness, we will drown in a bottomless pit. Rather something like conceptual construction, or re - construction, is needed. And for this purpose it is necessary to get clear about what theoretical work we want the concept of a person to do. (Think back to the two examples given above: Descartes knows what theoretical aim he was after. He tries to solve the problem: Given that in reality mind and body are categorically distinct, how come we do not experience ourselves as consisting of two separate entities? He employs the concept of a person in his attempt to answer this specific question. Locke tries to solve a different problem: What is the appropriate subject of punishment and reward? He uses the concept of a person for this particular theoretical purpose. Both thinkers have quite determinate ideas of what the concept of a person was supposed to effect within their theories; and this allows them to attach a determinate sense to it.) So my suggestion is this: Only if we get clear about what kind of theory we are striving for, and what role the concept of a person is supposed to play in it, we can get, or regain, a less vexing concept of personhood.

A final *caveat*. In this theoretical, constructive endeavour of getting clearer about personhood, we ought not to expect much help from the natural sciences. The best we can hope for is corrective cooperation. The natural scientist may warn us, for example, that given a certain conception of personhood, persons so conceived could not be members of the natural world. But we must not forget that from a strictly naturalist point of view, *person* is just not a category. (*Nota bene*, this is not to say that personhood cannot be accounted for in a naturalist way. David Lewis, for example, has presented an ingenious naturalist account of personhood and transtemporal personal identity, which is based on a psychological concept of person.<sup>①</sup>) The crucial point here is this: We would have to have reached, independently, considerable conceptual clarity about personhood, before we could reasonably hope for a naturalist characterization of the entities which exemplify it. We cannot ask the natural scientist “What is a person?”, in the same state of almost complete conceptual ignorance and with the same hope for conceptual elucidation, in which we may ask: “What is a magnetic moment?”.

Natural science can teach us what human beings (considered exclusively as members of a certain biological species) are and what storks, computers, and, if there are or were any, extraterrestrials and replicants are—science can inform us about their physical and functional similarities and differences. But we must not hope that among the distinctions drawable in naturalist terms, there is one—already drawn, as it were—between those human beings, storks, computers, extraterrestrials, and replicants which (or who) are persons and those which (or who) are not. This would be silly. The natural sciences, with good reason, attach importance to providing no methodological space for value – concepts. This, of course, is not a frivolous narrow – mindedness on their part, but a well – considered delimitation of what does, and what does not, fall within their cognitive realm.

---

① Lewis (1976: 17–40).

## References

- AT: *Oeuvres de Descartes*, 12 Volumes ( revised edition ), eds. by C. Adam and P. Tannery, Paris: Vrin/C. N. R. S. , 1964—1976.
- Meditations; Objections and Replies; Meditations on the First Philosophy* ( 1641 ), by René Descartes.
- Essay: An Essay Concerning Human Understanding* ( 1690 ), by John Locke.
- Austin, J. L. 1962: *Sense and Sensibilia*, reconstructed from the manuscript notes by G. J. Warnock, Oxford: Oxford University Press.
- Crick, Francis. 1994: *The Astonishing Hypothesis - The Scientific Search for the Soul*, New York: Charles Scribner's Sons, Macmillan Publishing Company.
- Gazzaniga, Michael. 2005: *The Ethical Brain*, Chicago: University of Chicago Press.
- Kemmerling, Andreas. 2000: "Ich, mein Gehirn und mein Geist - Echte Unterschiede oder falsche Begriffe?", in: N. Elsner/G. Lüer ( eds. . ), *Das Gehirn und sein Geist*, Göttingen: Wallstein, 2000, 221—243.
- . 2014: "Why is personhood conceptually difficult?", in: Michael Welker ( ed. ), *The Depth of the Human Person. A Multidisciplinary Approach*, Grand Rapids, Michigan/Cambridge: Wm. B. Eerdmans Publishing Co. , 15—44.
- Lewis, David. 1976: "Survival and Identity", in: Amélie O. Rorty ( ed. ), *The Identities of Persons*, Berkeley: University of California Press, 17—40; reprinted with a postscript in: D. Lewis, *Philosophical Papers*, vol. 1, Oxford University Press ( 1983 ), 55—77.
- Scholz, Oliver. 1999: *Verstehen und Rationalität; Untersuchungen zu den Grundlagen von Hermeneutik und Sprachphilosophie*, Frankfurt a. M. : Klostermann.
- Whately, Richard. 1841: *Elements of Rhetoric*, Oxford: W. Baxter. ( This edition is the sixth revised and enlarged edition of the original edition published in 1828 ).