

Hume's Methodology and the Science of Human Nature

In this paper I try to explain a strange omission in Hume's methodological descriptions in his first Enquiry. In the course of this explanation I reveal a kind of rationalistic tendency of the latter work. It seems to contrast with "experimental method" of his early Treatise of Human Nature, but, as I show that there is no discrepancy between the actual methods of both works, I make an attempt to explain the change in Hume's characterization of his own methods. This attempt leads to the question about his interpretation of the science of human nature. I argue that his view on this science was not a constant one and that initially he identified this science with his account of passions. As this presupposes the primacy of Book 2 of his Treatise I try to find new confirmations of the old hypothesis that this Book had been written before the Book 1, dealing with understanding. Finally, I show that this discussion of Hume's methodology may be of some interest to proponents of conceptual analysis.

Every Hume scholar knows that in his *Treatise of Human Nature* Hume constructs a "science of human nature", or "science of man", which, as he believes, is a basis of other sciences; and that he is sure that "the only solid foundation we can give to this science itself must be laid on experience and observation"¹. Indeed, according to the subtitle of the *Treatise*, it was "an attempt to introduce the experimental method of reasoning into moral subjects". Hume confirms this position in his *Abstract* of the *Treatise*. Here he writes that in the *Treatise* its author "proposes to anatomize human nature in a regular manner, and promises to draw no conclusions but where he is authorized by experience" (T Abstract. 2; SBNT 646).

So, Hume's position seems to be clear enough. And it is quite natural to expect that this is the end of story. Of course, Hume published other books dealing with his science of man, but it is widely held that in them, and, in particular, in *An*

¹ David Hume, *A Treatise of Human Nature*, 2nd ed., ed. L. A. Selby-Bigge and P. H. Niddich (Oxford: Oxford University Press, 1978), XVI (hereafter SBNT XVI). I will also point out in the parentheses the book, part, section, and paragraph of the *Treatise* according to another edition: David Hume, *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton (Oxford: Oxford University Press, 2000) (hereafter T). I will cite Hume's first *Enquiry* in parentheses according to the two editions: David Hume, *An Enquiry concerning Human Understanding*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 2000) (hereafter EHU – by section and paragraph), and David Hume, *Enquiries concerning Human Understanding and concerning the Principles of Morals*, 3rd ed., ed. L. A. Selby-Bigge and P. H. Niddich (Oxford: Clarendon Press, 1975) (hereafter SBN – by page); I will refer to his second *Enquiry* in the same way according to the two editions: David Hume, *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 1999) (hereafter EPM – by section and paragraph), and David Hume, *Enquiries concerning Human Understanding and concerning the Principles of Morals*, 3rd ed., ed. L. A. Selby-Bigge and P. H. Niddich (Oxford: Clarendon Press, 1975) (hereafter SBN – by page). I will cite Hume's *Dissertation on the Passions* in the same way according to David Hume, *Dissertation on the Passions; The Natural History of Religion*, ed. Tom L. Beauchamp (Oxford: Oxford University Press, 2007) (hereafter DP). And I will cite Hume's letters according to J. Y. T. Graig's edition: *The Letters of David Hume*, ed. J. Y. T. Greig, 2 vols. (Oxford: Oxford University Press, 1932) in parentheses by volume and page as (HL).

Enquiry concerning Human Understanding (1748), he adopts the same methodology.²

In this paper I will try to show that the real situation is much more complicated. Let us start with what might be called a kind of terminological anomaly, which may be found in Hume's first *Enquiry*. This work has a loose composition³, but its first section is an obvious analogue of the Introduction to the *Treatise*. In the Introduction Hume tries to defend metaphysics, that is, "abstruse" reasonings, especially in the science of human nature, and also draws a picture of this science as regards its relationship with other sciences, its scope and methods. The same topics he discusses in the first section of his *Enquiry*. He defends "true metaphysics"⁴ as a way to deal with human nature and says a few words about its methods.

But – and this is a striking difference – in Section 1 of the first *Enquiry* Hume *does not* say that the science of human nature is based on experience. The word "experience" is not even used in this section. Compare the Introduction to the *Treatise*: this word is used here four times. He also writes here in a methodological context about "experiments" (five times) and "observation" (three times). The word "experiments" is absent in the first section of the *Enquiry*; as for "observation" and derived words – so they are used in that section (three times), but all the cases of their usage have no relation to Hume's methodological reflections in this text.

Before trying to solve this terminological puzzle, we should realize what Hume actually claims in that section of his *Enquiry* concerning methods of his "true metaphysics". We read about "exact analysis", "accurate and just reasoning", "accurate scrutiny into the powers and faculties of human nature" and "certainty" of such a science (EHU 1. 12-14; SBN 12-14). Taken as such, these characteristics are not very informative, and they can hardly be a direct evidence that Hume abolishes experimental methods in favor of some rationalistic methodology. Indeed, in earlier works Hume said similar things about experimental reasonings. For example, in his *Abstract* he suggests that "accurate disquisitions of human nature" might be founded "entirely upon experience" (T Abstract. 2; SBNT 646).

² Some authors, for example, Nicholas Capaldi, *David Hume: The Newtonian Philosopher* (Boston: Twayne, 1975), Stephen Buckle, *Hume's Enlightenment Tract: The Unity and Purpose of An Enquiry concerning Human Understanding* (New York: Oxford University Press, 2001), and Tom L. Beauchamp, "Introduction: A History of the Enquiry concerning Human Understanding", in: David Hume, *An Enquiry concerning Human Understanding* (Oxford: Oxford University Press, 2000), are quite explicit on this point. But many others rather tacitly accept this view – by not stressing differences between the *Enquiry* and *Treatise* in this respect. A good example is a recent work of Robert J. Fogelin, *Hume's Skeptical Crisis* (Oxford: Oxford University Press, 2009), 139–44. And while some scholars pointed to a few differences in methodology between the *Treatise* and the first *Enquiry* – as, for example, Antony Flew, who claimed in his *Hume's Philosophy of Belief: A Study of His First 'Inquiry'* (London: Routledge, 1961), 108, 14, that the first *Enquiry* was characterized by "intensified methodological interests" and "greater emphasis on questions about the nature, presuppositions, and limitations of various sorts of investigation", the common view, nonetheless, might well be expressed by the following phrase of P. B. Wood: "Methodically, therefore, the *Treatise* and the *Enquires* seem to be of a piece, because in these works Hume sought to discover the mechanisms of the mind through the use of induction and the classificatory and descriptive methods of the natural historian", see P. B. Wood, "Hume, Reid and the Science of the Mind", in M. A. Stewart and John P. Wright (eds.), *Hume and Hume's Connexions* (University Park, PA: The Pennsylvania State University Press, 1994), 122–123.

³ Cf. M. A. Stewart, "Two Species of Philosophy: The Historical Significance of the First *Enquiry*", in Peter Millikan (ed.), *Reading Hume on Human Understanding: Essays on the First Enquiry* (Oxford: Oxford University Press, 2002), 79–81.

⁴ Rejecting at the same time a traditional metaphysics: in this respect both works do not differ as well.

But we should not forget that in Section 1 of the first *Enquiry* Hume simply does not refer to experience. Instead of talking of inner experience he says, for example, that mental objects “must be apprehended... by a superior penetration⁵, derived from nature, and improved by habit and reflection” (EHU 1. 13; SBN 13). And *this* fact, in connection with his words of “accuracy” of his “true metaphysics” gives us reason to believe that he might consider it as a rationalistic enterprise.⁶

In the next sections of this paper I am going to evaluate this idea. We will see that there is quite a strong evidence in favor of it in the first *Enquiry*. Moreover, I will try to show that Hume really uses a kind of rationalistic method in his investigations. In other words, his results are based on deductive inferences, at least partly. The problem is that he used such deductions in his earlier works too. Realizing this fact will lead us to the following question: why Hume had not said the same things – which he says in Section 1 of the first *Enquiry* – in his Introduction to the *Treatise*? And this question will help us to respond to another one: what did he actually mean by the “science of human nature” in his early years?

I

So, let us suppose that Hume was inclined to approve a kind of rationalistic or deductive methodology in his first *Enquiry*. To support this hypothesis we just have to look at Section 12. Here, among other things, Hume discusses skepticism of Descartes. One of the details of this discussion is important for our purposes. Hume claims that the method of Descartes (a kind of quintessence of rationalism) – “To begin with clear and self-evident principles, to advance by timorous and sure steps, to review frequently our conclusions, and examine accurately all their consequences” – is the *only way* “by which we can ever hope to reach truth, and attain a proper stability and certainty in our determinations” (EHU 12. 4; SBN 150). Hume mentions here first, third, and fourth rules of Cartesian method. It is impossible to understand the meaning of the third rule without reference to deduction: according to Descartes, it is deduction by which we “advance” in such a manner.⁷

If we take into account that Hume talks here about “study of philosophy”; that he is sure that these rules “are the only methods, by which we can ever hope to reach truth”, and that one of these rules is a rule of deduction, we can make quite a plausible conclusion that he approves deductions as a way of doing metaphysics.

Of course, it is possible to understand those passages in a weak sense, allowing room for “timorous and sure steps” along the lines of experimental method,⁸ as

⁵ Or “by a superior subtilty and penetration”, as in early editions of his first *Enquiry* under a different title.

⁶ By “rationalistic enterprise”, I understand a research based on a method by which it is possible to be fully certain of obtaining some states of affairs even if the facts of their obtaining are not directly and clearly available to us. This certainty can be based only on some kind of *a priori* reasonings from some safe premises. If, however, we can be certain of states of affairs only if they are directly given to us in a clear way we must follow the way of experience.

⁷ Cf. *The Philosophical Writings of Descartes*, ed. J. Cottingham, R. Stoothoff, and D. Murdoch, vol. 1 (Cambridge: Cambridge University Press, 1985), 120, 15. Hume might be aware of these details, of course. So, it is difficult to agree with Peter Millican’s claim that Hume was “radical” in his “rejection of the whole Cartesian project” – see his “The Context, Aims, and Structure of Hume’s First Enquiry”, in Peter Millican (ed.), *Reading Hume on Human Understanding*, 29. At the same time, most of his contrasting comparisons cannot be disputed, of course.

⁸ Cf. Hume’s story of Boyle and Newton in David Hume, *History of England*. 8 vols. (London, 1826), vol. 8, 293.

well as of deductive one. But even if such an interpretation is a right one, it is obvious that those passages (1) make a sharp contrast to methodological descriptions in the *Treatise*; (2) can be treated as a continuation of rationalistic tendency of Section 1 of the first *Enquiry*; and (3) give us reason to conclude that now Hume admits that “true metaphysics” can be a deductive science, at least partly.

But if Hume admits this why would not he liken his metaphysics to mathematics, which, as he believes, is a deductive science? It may sound strange to some ears, but this is what he in fact does do, in Section 7 of his first *Enquiry*. He says that “advantages and disadvantages” of “moral or metaphysical sciences” and mathematics “nearly compensate each other”: “If the mind, with greater facility, retains the ideas of geometry clear and determinate, it must carry on a much longer and more intricate chain of reasoning, and compare ideas much wider of each other, in order to reach the abstruser truths of that science. And if moral ideas are apt, without extreme care, to fall into obscurity and confusion, the inferences are always much shorter in these disquisitions, and the intermediate steps, which lead to the conclusion, much fewer than in the sciences which treat of quantity and number» (EHU 7. 2; SBN 61). Of course, Hume emphasizes here some differences between mathematical and metaphysical sciences, but doing so he seems to presuppose that they have something in common, namely – as we may reasonably assume – the way they make their inferences. And when he says after that that there are no Euclid’s propositions which have less parts than any sound moral reasoning (ibid.), he makes this tacit supposition much more explicit.

It might be objected that this evidence is not quite conclusive, because Hume adds that our tracing of the “principles of the human mind” (related to metaphysics) is connected with “enquiries concerning causes” (ibid.), and we can find causes only by experience (EHU 12. 29; SBN 164). Hence, our inferences in metaphysical sciences might be of the same nature as experimental inferences in natural sciences. In fact, however, Hume claims that metaphysical sciences differ from natural philosophy in that metaphysicians must spend much more time in clarifying ideas (EHU 7. 2; SBN 61), and it is quite possible that their inferences *in the course of such a clarifying* are more similar to inferences made in mathematics, than that in natural philosophy. In fact, he clearly indicates in the above passages that in metaphysics, as well as in mathematics, we have to “compare ideas”, and presumably make inferences in order to find their relations, which is obviously not a way of natural philosophy. And this is all we need to make our case.

In sum, if Hume really inclines now to approve a deductive method in metaphysics, we can predict that he would make an explicit attempt to compare metaphysics with mathematics (another deductive science), and that in such comparison he would not mention any differences as regards the nature of their inferences. And the above fragment exactly corresponds to these predictions. So it can be treated as confirmation of the hypothesis of Hume’s “rationalistic turn” in the first *Enquiry*. Now it is time to search for further confirmations by looking at those parts of the *Enquiry* where he tries to obtain real results concerning human understanding. It is natural to expect that it would be possible to find deductions there.

II

It is not at all difficult to find deductions in Hume's reasonings on human understanding. Just look at his "chief argument" (as he calls it in his *Abstract*) concerning the way we make experimental inferences. In the first *Enquiry* it is presented as an investigation into "the nature of that evidence, which assures us of any real existence and matter of fact, beyond the present testimony of our senses, or the records of our memory" (EHU 4. 3; SBN 26). He claims that we believe in such facts due to past experiences which we extrapolate to the future⁹ by custom – powerful instinct implanted in us by Nature. His argument goes as follows. At first he gives an outline of two kinds of reasonings: "demonstrative reasoning, or that concerning relations of ideas, and moral reasoning, or that concerning matter of fact and existence". Then, after a few intermediate steps, in the course of which he shows that our reasonings concerning matter of fact are based on causal inferences, and that causal inferences depend on a presupposition that "the future will be conformable to the past", he shows that no demonstrative reasoning can prove that future experience will be like past experience: (1) we can clearly and distinctly conceive that "the course of nature may change"; (2) whatever "can be distinctly conceived, implies no contradiction"; (3) but everything that can be demonstrated is of such a kind that its denial implies a contradiction. After that he proves that factual reasonings could not be basis of our belief in the correspondence between past and future experience either as they presuppose it. And, showing this, he makes a final conclusion that our belief in the correspondence between past and future is not a product of reasonings, but of custom, and that "all inferences from experience... are effects of custom, not of reasoning"¹⁰ (EHU 4–5; SBN 26–47).

Now let us take a closer look at this well-known argument. Among other things, Hume tries to prove that it is impossible to demonstrate that the course of nature will not change. How can we describe his proof? Is it possible to demonstrate that this course will not change? No, this implies a contradiction. If so, then he proves a proposition which is impossible to deny without a contradiction. And this means in turn that his proof might well be qualified as a kind of demonstration itself. It is easy to see that the same is true about other parts of Hume's argument.

Demonstrations are obvious examples of *a priori* deductive reasonings, which may help to reveal some necessary truths. And it seems that we can safely conclude that the very core of Hume's metaphysics is forged out of deductions.

III

⁹ This extrapolation can be interpreted in such a way: we believe that if we would see the same experiences as we saw in the past, we would expect that they would be followed with the same experiences which followed former experiences. Such interpretation treats future in a counterfactual sense, and it gives us an opportunity to interpret Hume's extrapolation of the past to the future experiences as a wider extrapolation from observed to unobserved. Hume himself often treats this extrapolation in the sense just mentioned (see, for example, T 1. 3. 6. 4-5; SBNT 89).

¹⁰ See for a discussion Peter Millican, *Hume on Induction and Faculties* (TS at <http://davidhume.org>, 2009), 46–49.

Now let us look around to see where we have come. We have seen that in his *Treatise* Hume treated experimental reasonings as a way to construct a metaphysical science of human nature. In the first *Enquiry*, however, he changes his mind. It seems that in this work he treats this “true metaphysics” as a demonstrative science.

This picture, however, must be immediately corrected. The problem is that in the first *Enquiry* Hume denies the existence of any demonstrative science except mathematics. And he is quite explicit about that. So, for example, in Section 12 he says that “the only objects of the abstract sciences or of demonstration are quantity and number, and... all attempts to extend this more perfect species of knowledge beyond these bounds are mere sophistry and illusion”. A bit later he repeats that “the sciences of quantity and number... may safely... be pronounced the only proper objects of knowledge and demonstration” (EHU 12. 27; SBN 163). The similar things he says also in Section 4 (see EHU 4. 1; SBN 25).

What to do with these passages? Hume surely believes that his “true metaphysics” is not a demonstrative science. But what kind of science is it? It is interesting that in his classification of sciences in Section 12 he does not mention it at all. So we need to make some conjectures here. We know for sure that he uses a kind of deductive reasonings in this science and that he is aware of that. If these deductive reasonings are not demonstrations in the exact sense, so what are they?

I think we can solve this puzzle, and Hume himself helps us to do that. In his letter to Gilbert Elliot (February 18, 1751) he says that “in Metaphysics or Theology... Sophistry must be oppos’d by Syllogism” (HL 1. 151). This passage is crucial because he talks about syllogisms (under which he seems to understand formally performed reasonings¹¹) in Section 12 of the *Enquiry* as well. In a very condensed manner he informs us here about a few important features of syllogisms. He says that “syllogistical reasonings... may be found in every other branch of learning, except the sciences of quantity and number”, and that they are in some way related to “more imperfect definition[s]” (EHU 12. 27; SBN 163). He points out also that definitions help to clarify ideas and make them precise and determinate.

Hume calls such syllogisms “pretended” (ibid.), presumably because in some situations they do not provide us with a new information, and because their conclusions cannot be automatically qualified as true ones. But it does not follow that he believes that they are useless. Indeed, from the above considerations it is clear that Hume (1) was ready to describe metaphysical arguments as syllogisms, and that he (2) admitted that syllogisms could be a means to clarify our ideas. Hume gives an example of such clarification, and it helps to understand his position: “To convince us of this proposition, *that where there is no property, there can be no injustice*, it is only necessary to define the terms, and explain injustice to be a violation of property” (ibid.). Here he seems to mean that if we look at that proposition, “where there is no property, there can be no injustice”, we might not be sure at first whether it is true or not, or, in other words, what exactly is a relation be-

¹¹ Cf. his *Dialogues concerning Natural Religion* 2. 4.

tween absence of property and absence of injustice. Then we define the terms and say that injustice is a violation of property. And from this definition it follows that that proposition is true: if injustice is a violation of property, then if there is an injustice, there is a property, and if there is no property, there is no injustice. So we see here (1) definition of terms, (2) a kind of syllogism and (3) a resulting clarification of a relation between some items, not evident from the start. And we could not clarify this relation between them without a syllogism. Indeed, if our statement is a kind of imperfect definition, then to prove that it is a true one we have to give a more perfect definition and formally derive from it the statement in question.

Hume suggests that the difference between syllogisms and demonstrations is that demonstrations require more than definitions of the terms: “*That the square of the hypotenuse is equal to the squares of the other two sides*, cannot be known, let the terms be ever so exactly defined, without a train of reasoning and enquiry” (ibid.). He does not clearly explain the nature of such “enquiry”, but let us assume that this is so.¹² It is more important here to evaluate – in the light of this distinction between syllogisms and demonstrations – his above-mentioned argument (which is the very core of his theory of experimental cognition) leading to the conclusion that our factual inferences are not based on reasonings. If Hume is consistent in his methodological reflections, then it can be interpreted in syllogistic terms. For simplicity, let us restrict ourselves to a few parts of this argument. Hume considers our inferences concerning matter of fact. Are they equal to demonstrations? To answer this question we have to (1) define “demonstrative inferences” as inferences of such a kind that what is contrary to their conclusions contains a contradiction; (2) define “factual inferences” as forming beliefs in unobservable facts which are accompanied with a comprehension of the possibility to conceive – clearly and distinctly – that these facts are absent, which means that such an absence involves no contradiction¹³, – and we can *conclude* that our factual inferences differ from demonstrations. And when later Hume makes his definition of factual inferences much more detailed and says that they are based on a principle of conformity between the past and the future experiences, he can derive from

¹² I think we can interpret such phrases, with David Owen, as indicating that “Hume’s account of demonstration does not rely on any formal notion of deduction” – David Owen, *Hume’s Reason* (New York: Oxford University Press, 2000), 99. Owen, however, believes that Hume does “drop talk of syllogisms” (ibid., 107), which is not quite true. Helen Beebe recently has criticized Owen’s views and came to the conclusion that “Hume at least implicitly recognizes a distinction between demonstration on the one hand and deduction on the other” – Helen Beebe, *Hume on Causation* (London: Routledge, 2006), 30. But, like Owen, she does not discuss Hume’s remarks on syllogisms. Peter Millican attacks the very attempts to find in Hume even implicit distinction between demonstration and deduction and tries to prove that under “demonstration” Hume understood something “broadly equivalent to deduction (in the familiar informal sense)”. Therefore, a “successful demonstration is... a deductively valid argument, either from some hypothetical premiss(es) to a conclusion, or for a conclusion *tout court* (in which case any premisses must themselves be already certain) – Peter Millican, “Hume’s Old and New: Four Fashionable Falsehoods, and One Unfashionable Truth”, *Proceedings of the Aristotelian Society*, Supplementary Volume LXXXI (2007): 180. This interpretation leads to the conclusion that any proposition can be demonstrated, and Hume would surely deny this. Millican could reply that from the fact that any proposition can be demonstrated it does not follow that any proposition can be proved by demonstration (cf. his “Hume’s Sceptical Doubts concerning Induction”, in Peter Millican (ed.), *Reading Hume on Human Understanding*, 135), but the point is that Hume does not make such a distinction.

¹³ Note that all these definitions are not redundant. It is important because if we use redundant definitions we would be able to treat as syllogisms in Humean sense any demonstrative reasoning we can conceive.

this definition a conclusion that this principle cannot be founded on factual inferences, as it is clear that a non-self-evident principle cannot ground itself.

So, Hume's argument really involves syllogistical reasonings, as he understands them. And, among other things, his syllogisms clarify a relation between our inferences concerning matter of fact, based on extrapolation of past experiences to the future, and demonstrative inferences. Hume proves that they are totally different.

Now, Hume believes that "the obscurity of the ideas, and ambiguity of the terms" are the very "chief obstacle... to our improvement in the moral or metaphysical sciences" (EHU 7.2; SBN 61). He explains his position concerning this obstacle as follows. In Section 1 of the first *Enquiry* he says that "it cannot be doubted, that the mind is endowed with several powers and faculties, that these powers are distinct from each other... and consequently, that there is a truth and falsehood in all propositions on this subject, and a truth and falsehood, which lie not beyond the compass of human understanding". Some of these distinctions – between "the will and understanding, the imagination and passions" etc. – are quite obvious. But other, "finer" distinctions, "are no less real and certain, though more difficult to be comprehended" (EHU 1. 14; SBN 13–14). This difficulty in comprehending, as he explains in Section 7, is just what leads us to confusions: "But the finer sentiments of the mind, the operations of the understanding, the various agitations of the passions, though really in themselves distinct, easily escape us, when surveyed by reflection... Ambiguity, by this means, is gradually introduced into our reasonings: Similar objects are readily taken to be the same: And the conclusion becomes at last very wide of the premises" (EHU 7.1; SBN 60).

So, according to Hume, the ambiguity of metaphysical sciences is at least partly rooted in our identification of some operations of mind which in reality are distinct from each other. In other words, this ambiguity results from misinterpreting of actual relations of some mental entities. But we have seen that such a misinterpreting can be avoided with the help of syllogisms. Moreover, we have already seen that Hume actually avoids by syllogisms a possible misinterpretation of a relation between demonstrative inferences and inferences concerning matter of fact.

Let us call the syllogisms which help us to clarify relations (e. g. that of difference and sameness) between our ideas of some operations of the mind *clarifying syllogisms*. Their epistemic place in Hume's system can be further specified if we take into account what he says in Section 12 of his first *Enquiry*. Here Hume notes that as all our ideas except that of "quantity and number" are "different from each other" (that is, they are not made out of similar parts like numbers), "we can never advance farther, by our outmost scrutiny, than to observe this diversity, and, by an obvious reflection, pronounce one thing not to be another" (EHU 12. 27; SBN 163). The context makes obvious that he considers such statements as a kind of intuitive truths. Indeed, they may be classified as expressing the resemblance relations (T 1.1.5.10; SBNT 15) that he treats as intuitive. Such truths, like demonstrative ones, presuppose inconceivability of the states of affairs, which are contrary to them (if one thing is different from another we cannot conceive they are the same). And he adds: "Or if there be any difficulty in these decisions, it proceeds entirely

from the undeterminate meaning of words, which is corrected by juster definitions” (EHU 12. 27; SBN 163). In other words, in some situations we may need definitions in order to see real relations between our ideas, opposite to which cannot be clearly conceived. We know, however, that such definitions must be supplemented with syllogisms. This is so simply because these definitions are propositions, and we formally *conclude* from them to the relations in question; usually such syllogisms are short, but let us recall that Hume notes about metaphysical reasonings that they are short. And such syllogisms lead us to conclusions which are similar to conclusions of demonstrations in that the states of affairs which are opposite to them are inconceivable. These very syllogisms I have just called “clarifying syllogisms”. Hume seems to believe that there is no much need in *such* syllogisms in most our sciences, because distinctions between our ideas used in them are quite obvious¹⁴ (so we have no need in definitions to compare entities, and our reasonings about their differences are immediate judgments, not syllogisms¹⁵). But from his remarks it follows that in metaphysics the situation is different. Indeed, we have seen that he claims that in that science we must do the hard work to clarify ideas, and so to find relevant syllogisms. Therefore, they must be of a great value here. Moreover, it means that they can be interpreted as its methodological basis. And at the time of writing his first *Enquiry* Hume seemed to feel this. His new insight is reflected in his methodological observations in that work.

We have to stress, however, that from those methodological observations it surely does not follow that Hume believes now that his science of human nature is based only on syllogisms. First of all, in his other later works dealing with the science of human nature, that is, in the *Enquiry concerning the Principles of Morals* (1751) and in the *Dissertation on the Passions* (1757) he clearly approves the experimental method (see EPM 1. 10; SBN 174; DP 6. 19) and follows it. In the *Enquiry concerning Human Understanding* he makes more room for deductions. But we have just shown that his deductions or syllogisms presuppose some definitions, and if they should lead us to true (and not only formally valid) conclusions, these definitions must be real definitions, that is, have some relation to what is given, as Hume says, in “immediate perception”. Indeed, Hume tries to establish such relations by tracing origins of ideas in impressions. Such a method could not be performed without a reference to experience. We should remember also that even in his first *Enquiry* Hume occasionally describes some of his investigations as “experiments” (see EHU 5. 15–16; SBN 51–52; cf. also EHU 8. 9; SBN 85). So, his new position is rather a following one: metaphysical science of human nature is based on experience, experiments, and on syllogisms as well.

But what exactly do we mean claiming that Hume’s metaphysics is *based* on syllogisms? Let us suppose that we mean simply that Hume would admit that in his metaphysical reasonings he needs syllogisms, and that they are quite important in this field. The problem is that he is sure that some syllogisms are widely used in

¹⁴ Cf. EHU 1. 13; SBN 13.

¹⁵ It is true that Hume tries to show that there are no vast or essential differences between conception, judgment and (syllogistic) reasoning, and that all of them could be in a sense reduced to the first (T 1. 3. 7. 5, note; SBNT 95–96). But his analysis should not be understood as a claim that there could not be differences between them at all.

other sciences as well. For example, we can use them to derive consequences from empirical hypotheses. Such pretended *a priori* reasonings he is inclined to characterize as not separated from experimental ones, as they are based on experience anyway.¹⁶ In other words, the usage of syllogisms in a science does not indicate as such that this science is not based only on experience and is based on syllogisms also.

The point is, however, that syllogisms in Hume's metaphysics reveal some truths, which have the same features as demonstrative or intuitive truths. And in this respect they differ from those reasonings which are no more than logical derivations from some factual truths and, correspondently, which lead to factual truths. The truths of a latter kind may be provided by experimental method (EPM 1. 10; SBN 174), and for this reason we can claim that experience is the only basis of the sciences which make use out of those syllogisms. In fact, *such* syllogisms seem to be an essential part of Hume's experimental method, which can be summarized as consisting of (1) inductive generalizations, (2) deductive inferences from resulting general principles and (3) experimental confirmations of these inferences (Cf. EHU 1. 15; SBN 14–15; EHU 4. 12; SBN 30; EPM 1. 10; SBN 174; T 2. 2. 2. 12; SBNT 337–38). The truths, obtained in Hume's metaphysics, however, cannot be characterized as entirely factual, and as such truths of a non-experimental kind are obtained in this science by means of syllogisms, we can say that "true metaphysics", that is, the metaphysical part of the science of human nature, is based in a methodological sense not only on experience, but also on syllogisms. In other words, we can say that a science is based on syllogisms if there are a lot of important truths obtained – and have to be obtained – in it by means of syllogisms (not of demonstrations), and if these truths differ in their nature from the truths obtained by experimental method. Hume believes that experimental truths concern matters of fact and are of such a kind that what is opposite to them can be conceived. The truths the denial of which cannot be conceived concern relations of ideas. So if any science is at least partly based on syllogisms it has to contain a lot of important truths discovered as a result of clarification by syllogisms of some relations of ideas.

And we already know that Hume's syllogisms in his first *Enquiry* can be interpreted as a means by which we attain clarification of such a kind. To sum up: Hume seems to believe that (1) in common life we have some vague ideas of operations of our mind, (2) which it is impossible to clarify without much work to be done by philosophy, whose decisions, in fact, "are nothing but the reflections of common life, methodized and corrected" (EHU 12. 25; SBN 162). We must (3) trace the origins of those ideas in experience by reducing them to impressions (including not only those impressions which are their immediate causes but also those impressions that accompany them), (4) apprehended "in an instant"¹⁷, by a superior penetration" (EHU 1. 13; SBN 13). Such a reduction (5) provides us with definitions of the terms we use to refer to those operations, and, after that, even if (6) a

¹⁶ He suggests this in a lengthy note in his first *Enquiry* (EHU 5. 5, note; SBN 43–45).

¹⁷ This is Hume's reply to his own doubts about the reliability of introspection (cf. T Introduction 10; SBNT XIX).

few of those ideas remained somewhat vague as immediate copies of relevant impressions¹⁸ (if we set aside the different circumstances at which we employ those ideas, which help us in “description” of correspondent impressions and in definitions – cf. T 2. 1. 2. 1; SBNT 277), produced in us by the mental operations in question, (7) we would be able to conclude *a priori* that they are different or the same. This is the way his “true metaphysics” must proceed, at least in that aspect of this science which Hume calls “mental geography”, which is to “separate [different operations of the mind] from each other, to class them under their proper heads, and to correct all that seeming disorder, in which they lie involved, when made the object of reflection and enquiry” (EHU 1. 14; SBN 13).¹⁹ The problem is that Hume had no explicit theory concerning this methodological position.²⁰ I have tried to dig it out but this just reveals its lack on the surface of the first *Enquiry*. And I think this helps us to give a final explanation of Hume’s unwillingness to use the words “experience” and “experiments” in his introductory section of that work. If he had used them he would make an impression as if he believed that his “true metaphysics” was entirely experimental. He could correct it by specifying the role of syllogisms in metaphysics, but he had not developed an explicit theory of syllogisms at the time of writing his first *Enquiry*, while feeling importance of such deductions. So he decided to present his position by keeping silence about experience.

Anyway, we can be sure that at the time of writing his first *Enquiry* Hume believed that his metaphysical science of human nature was partly a deductive, syllogistical one. But why did not he realize this fact in his earlier times? It is easy to guess that he used many syllogisms in his *Treatise*. Just look at his discussion of causation, for example. He tries to show that it is impossible to demonstrate that “*whatever begins to exist, must have a cause of existence*” by means of “an argument, which proves at once, that the foregoing proposition is neither intuitively nor demonstrably certain” (T 1.3.3.3; SBNT 79). And this argument can be formulated in a following way (I simplify and augment Hume’s reasonings a bit): we have to (1) define demonstration as a reasoning presupposing inconceivability of a state of affairs opposite to that which is proved by it; (2) define cause and effect as distinct events, connected in a regular manner; (3) define distinct events as events which can be clearly and distinctly conceived separately, and we can make a conclusion that it is impossible to demonstrate that every event must have a cause.

As I have just mentioned, Hume used such syllogisms in many other places of the *Treatise* as well (see, for example, T 1. 3. 6. 5; SBNT 89; T 1. 3. 9. 10; SBNT 111; T 1. 3. 11. 7; SBNT 126; T 1. 3. 14. 12; SBNT 161–162; T 1. 3. 14. 34; SBNT 172). Note, however, that in the *Treatise* his account of syllogisms and of their difference from demonstrations was in even less developed state than in the first *Enquiry*. So we should not be surprised to see that in the *Treatise* Hume occa-

¹⁸ Note that such a situation is possible in some different cases (and as an effect of different causes) also – when, for example, we know definitions of two complex geometrical figures, but cannot discriminate their mental images.

¹⁹ But the mental geography is not the only aspect of the science of human nature. Another aspect of it is a search for general principles (EHU 1. 15; SBN 14–15), and it can be interpreted as, at least partly, an experimental one.

²⁰ So it is no surprise, for example, that he does not mention syllogisms in his list of arguments (EHU 6. 1 note; SBN 56).

sionally but clearly characterizes his philosophical arguments as demonstrations (T 1. 2. 2. 6; SBNT 31). Anyway, why did not he suggest in his Introduction to the *Treatise* that the science of human nature is based not only on experience?

This question might remain unanswered, of course. But I think that it is possible to give an answer to it, that is, to explain what prevented Hume from claiming that the science of human nature is based not only on experimental reasonings. To give such an explanation we need at first to answer another question: what did he mean by the “science of human nature” at the time of composing his *Treatise*?

IV

In the very first sentence of his *Enquiry concerning Human Understanding* Hume identifies the science of human nature with moral philosophy (EHU 1.1; SBN 5), i.e. with a system of sciences treating human beings in a direct way. And it seems quite natural to suppose that he did the same in his *Treatise*. Indeed, in the *Abstract* he notes that “This treatise... of human nature seems intended for a system of the sciences” (T Abstract. 3; SBNT 646), and in the *Treatise* itself he says that “In pretending... to explain the principles of human nature, we in effect propose a compleat system of the sciences” (T Introduction 6; SBNT XVI).

In the previous sections I myself have presupposed that the science of human nature, the science of man, and moral philosophy are, essentially, the very same thing. A careful reading of Hume’s early texts, however, shows that at the time of writing the Introduction to his *Treatise* he did not identify his science of human nature with moral philosophy in general. Let us look at these texts to see why it is so. Hume begins his observations concerning the science of human nature in his Introduction to the *Treatise* with a remark that “all the sciences have a relation, greater or less, to human nature” (T Introduction 4; SBNT XV). He points out that “Even *Mathematics*, *Natural Philosophy*, and *Natural Religion*, are in some measure dependent on the science of man” and, if even they “have such a dependence on the knowledge of man, what may be expected in the other sciences, whose connexion with human nature is more close and intimate?” (T Introduction 4–5; SBNT XV). Talking about sciences of this latter kind, he mentions “four sciences of *Logic*, *Morals*, *Criticism*, and *Politics*” (T Introduction 5; SBNT XV–XVI). We know that he had planned to publish five books of his *Treatise*: *Of the Understanding* (Logic), *Of the Passions*, *Of Morals*, *Of Politics*, and *Of Criticism* (T Advertisement; SBNT XII), so in that list he has omitted one of them. Why does not he mention his account of passions? Soon we will see why.

A bit later Hume says that we must “march up directly to the capital or center of these sciences, to human nature itself” and become “masters of [it]”; and continues: “From this station we may extend our conquests over all those sciences, which more intimately concern human life” (T Introduction 6; SBNT XVI). These phrases leave no doubt that he does not identify his investigation into human nature with above-mentioned “four sciences of *Logic*, *Morals*, *Criticism*, and *Politics*”. He clearly says that “we... extend” our investigations over them from that station.

If we take into account that Hume (1) does not identify his science of human nature with sciences “whose connexion with human nature is more close and intimate”, and (2) does not mention his “account of the passions” among those sciences, we can directly identify this account with the science of human nature.²¹ To make this conclusion almost inevitable we need to show that Hume was ready to admit that his theory of passions could be considered as a basis of other sciences. Indeed, Hume suggests that his science of human nature is a “capital or center of these sciences”²², so they should depend on it. So, if his theory of passions were this science, then he should treat it as a foundation of other parts of his *Treatise*. And this is what he does: in his *Abstract* he says that he “has laid the foundations of the other parts [i.e. planned Books on morals, criticism, and politics] in his account of the passions” (T Abstract 3; SBNT 646).

So we have good reason to suppose that at least at the time of writing the Introduction to his *Treatise* Hume identified the science of human nature with his account of passions. Now it is time to remind that he says that the science of man is to be based only on experience in this very Introduction. Note that by “science of man” he surely *does not* mean here the science of human nature: he treats former rather as a system of sciences, comprising the science of human nature and those four sciences of logic, morals, etc.²³ Still, talking about experience as the “only solid foundation” of the science of man he might well *primarily* refer to his account of passions as a basis of the science of man and might think about other sciences contained in the science of man on the model of this account. This, of course, presupposes (1) that at the time of writing the Introduction he had not clear idea about what his theory of understanding would look like.²⁴ Now, if we have reason to suppose that (2) at that time he had a clear idea of central features of his theory of passions, and if (3) such a theory is an experimental one, then we would be able to explain why in his Introduction to the *Treatise* Hume claimed that experience and observation are the only solid basis of the science of man.

V

Let us discuss these points, starting with the third one. First of all, it is worth to note that in Book 2 of the *Treatise, Of the Passions*, Hume describes his investigations as “experiments” about forty times. Comparing this Book with Book 1, *Of the Understanding*, we see that he does this about a dozen times (he writes about “experiments” much more here, but in a different context).²⁵ This fact clearly re-

²¹ James Noxon in his *Hume's Philosophical Development* (Oxford: Clarendon, 1973), 4, on a similar textual basis, enigmatically concludes that under the “science of human nature” Hume meant his theory of understanding.

²² He says that this capital is not a “science of human nature” but “human nature itself”; we will see, however, that the term “human nature” refers for him not only to this nature but also to a science of it. The term “science of human nature” is used only once in the Introduction when Hume discusses explaining the “ultimate principles of the soul”.

²³ That is, he treats the “science of man” as moral philosophy.

²⁴ If he had, then his actual position would be in discrepancy with those methodological prescriptions, because, as we have seen, in the *Treatise* he did not base his reasonings in that field on experience and experiments only.

²⁵ As for the Book 3, *Of Morals*, so he uses this term only once.

veals that he considers his account of passions as a kind of experimental theory. And it is no surprise that Parts 1 and 2 of Book 2 culminate in a series of “experiments”, which are designed to confirm Hume’s hypotheses. Looking at these eight experiments we can notice that they are experiments not only by name. Hume varies qualities of objects and their relations to ourselves or to others and looks at the effects of such modifications for our feelings. These effects confirm his hypotheses concerning the origin of our passions, such as pride and humility, love and hatred. And it is important to emphasize that these effects are not of such a kind that their contraries are inconceivable. For example, it is quite possible to imagine that, being a sibling of a gifted person, I feel no pride at all. Indeed, this passion is distinct from those entities and relations which in fact generate it, and for this very reason we can conceive them without conceiving this passion.

In other words, Hume’s reasonings in Book 2 of the *Treatise* can be qualified as true experimental reasonings, not as clarifying syllogisms or demonstrations. Of course, there are a lot of abstract considerations here, but almost all of them are of such a kind that they need to be confirmed by experiments. The only exception is his proof in T 2. 3. 3. 2–4 that reason is not the only cause of any of our actions. This proof is similar to Hume’s clarifying syllogisms in Book 1. But while this proof is very important to Hume, he might consider it rather as a foundation of his theory of action than as integral part of his theory of passions as such.²⁶ And even if this proof is to be included in the theory of passions, it is little doubt that Hume would consider it as an exception, which does not prevent it to be a theory that can be justly said to be based solely on experience²⁷. And if we look at this theory in general we see that while his syllogisms in Book 1 of the *Treatise* help to clarify resemblance relations between our ideas of some operations of the mind, his deductions in Book 2 are logical derivations from his experimental hypotheses. Recall that in his first *Enquiry* Hume makes a note, in which he argues that such reasonings do not differ in their essence from experimental ones. And this helps to explain why Hume says in his Introduction to the *Treatise* that his science of man is based on experience and experiments only: he thinks of this science on the model of his account of passions which is, indeed, essentially an experimental one.

Of course, this helps to explain that only if at the time of Hume’s writing the Introduction to the *Treatise* he had not any clear idea about what his theory of understanding would look like. And now we can discuss this hypothesis. We have seen that in the Introduction to the *Treatise* Hume suggests that the science of human nature does not coincide with the science of man. It is interesting, however, that Book 1 of the *Treatise* gives us quite a different picture. It is evident from Section 6 of Part 4 of this Book. At the very end of this section Hume says: “’Tis now time to return to a more close examination of our subject, and to proceed in the accurate anatomy of human nature, having fully explain’d the nature of our judgment and understanding” (T 1. 4. 6. 23; SBNT 263). The section in question might be intended by Hume to be the last section of Book 1 of his *Treatise*, followed by his

²⁶ This proof clarifies the mechanisms of the will, but, according to Hume, the will is not a passion in a strict sense.

²⁷ In the same way as Hume’s own counterexample with the missing shade of blue does not prevent him to talk as if his principle that our simple ideas are copies of preceding impressions had a kind of universal application.

account of passions²⁸. Anyway it is clear that he proceeds “in the accurate anatomy of human nature” in Book 2 *Of the Passions*. But this passage shows that his investigations into passions is just a continuation of such an anatomy, started with his theory of understanding. I think we have all reasons to identify the “anatomy of human nature” with the “science of human nature”. This means that now Hume considers his theory of understanding as an integral part of his science of human nature. And we know that syllogisms are an important part of his methodology in that field. So it seems that Hume might feel some tensions concerning his methodological descriptions as early as in 1739 when he published Books 1 and 2 of his *Treatise*. And indeed, we can find evidence of this. In his *Abstract* (1740), he says that in the first two books of the *Treatise* “Almost all reasoning is... reduced to experience” (T Abstract 27; SBNT 657). He says “almost all reasoning”, not “all reasoning”. We can interpret this passage as a sign of dissatisfaction with his methodological descriptions, given in the Introduction to the *Treatise*.

We know that this dissatisfaction grew and found its final expression in the first section of his first *Enquiry*. But why did Hume change his view on the science of human nature after writing the Introduction to the *Treatise*? One of the possible explanations is this. The book *Of the Passions* (where the science of human nature is contained in its original form) deals with “secondary impressions”, or “impressions of reflection”, i.e. with internal feelings which may be treated as reactions of human beings to “original impressions” and, even more often, ideas. But while developing his theory of understanding he had discovered that operations of this faculty concerning matter of fact also depend on internal reactions (to original impressions and ideas), or on instinct, generating belief in the objects, which are not given in the immediate experience, and also the impression of necessity, which he directly calls “impression of reflection” (T 1. 3. 14. 22; SBNT 165). And he had probably realized that in such a case he had no need to separate his theory of understanding from his theory of passions. As a result, he extended his science of human nature to the theory of understanding²⁹ and, later, to moral philosophy in general. So it is no chance that already in the final Section of Book 1 he says that “Human Nature is the only science of man” (T 1. 4. 7. 14; SBNT 273³⁰); and in his *Abstract* he also claims “that almost all the sciences are comprehended in the science of human nature, and are dependent on it” (T Abstract 3; SBNT 646).

So, the picture we have just seen suggests that when writing the Introduction to the *Treatise* Hume really had not quite a clear idea about his theory of understanding: his real findings in this field differ from what he expected in the Introduction to find there. Of course, authors often write introductions after completing the main text. But our interpretation suggests that Hume’s case was a different one. It is natural to think that he had written his Introduction to the *Treatise* after finishing the book on passions, but before composing the book on understanding. Or he

²⁸ Then he inserted between them the “melancholical” section *Conclusion of this Book*, final section of Book 1.

²⁹ So it is quite natural that at the very end of Part 3 of Book 1 Hume claims that habit (or custom) as one of the basic mechanisms of human understanding “is nothing but one of the principles of nature” (T 1. 3. 16. 9; SBNT 179).

³⁰ Note, that here he uses the words “human nature” as a name of science. The same usage was presupposed in T Introduction 6. This fact suggests that the final part of Section 7 is a revised version of some early text; initially the phrase “the only science of man” might be formulated differently – “the chief science of man”, for example.

could write it before composing both books, but, having in advance quite a clear idea about his planned account of passions, proceeded with the book on passions, as indicated in that Introduction, where he suggests that we should start with investigation into the human nature, that is, into passions. Of course, even in such a case he could actually start elsewhere, but it is very unlikely: his project requires clarifying basic features of the human nature before going further. So we need to give primacy to Book 2 *Of the Passions* to make our reconstructions plausible.

The idea of such a primacy is not a new one in the literature on Hume. N. Kemp Smith has famously claimed that “Books II and III of the *Treatise* are in date of first composition prior to the working out of the doctrines dealt with in Book I”.³¹ He was sure that Hume “entered into his philosophy” “through the gateway of morals”³² and that he “thought out the teaching of the *Treatise* in the reverse order from that in which he expounds it”³³. As an evidence for that, he cited Hume’s *Letter to a Physician* (1734). In this letter, rightly characterized as Hume’s first autobiography, we see, among other remarkable things, the following confession of the young philosopher: “I found that the moral Philosophy transmitted to us by Antiquity, labor’d under the same Inconvenience that has been found in their natural Philosophy, of being entirely Hypothetical, & depending more upon Invention than Experience. Every one consulted his Fancy in erecting Schemes of Virtue & of Happiness, without regarding human Nature, upon which every moral Conclusion must depend. This therefore I resolved to make my principal Study, & the Source from which I wou’d derive every Truth in Criticism as well as Morality” (HL 1. 16).

Now, this text is obviously similar to some passages of the Introduction to the *Treatise*. But it does not follow from it that Hume “thought out the teaching of the *Treatise* in the reverse order from that in which he expounds it”, that is, starting with Book 3 (*On Morals*) and proceeding to Book 2 and Book 1.³⁴ Rather it follows that he had started not with his moral theories but with investigations into the human nature. And we know that by such investigations he meant the theory of

³¹ Norman Kemp Smith, *The Philosophy of David Hume: With a New Introduction by Don Garrett* (New York: Palgrave Macmillan, 2005), VI.

³² Ibid.

³³ Ibid., 538

³⁴ And it does not follow from it that Hume “entered into his philosophy” “through the gateway of morals”. Indeed, it can be argued (see M. A. Stewart, “Hume’s Intellectual Development, 1711–152” in: M. Frasca-Spada and P. J. E. Kail (eds.), *Impressions on Hume* (Oxford: Clarendon, 2005), 32) that in the cited passage of the *Letter to a Physician* Hume talks about origins of his future *Treatise*, not about the origins of his philosophy in general, i. e. about “new Scene of Thought”, which had been opened before him when he was “about 18 Years of Age” (HL 1. 13). We know that before planning the *Treatise* he already had written a book on natural religion, which, as Hume indicated, had a complex structure and a kind of dialectical form with arguments and objections. He could begin writing this book about 1729, finish it about 1731 (HL 1, 154), and then he started working at his *Treatise*. So, Hume’s “gateway” to philosophy was rather natural religion, not morals (even if he already had some vague ideas about his future *Treatise* “before he left College” (EHU Advertisement; SBN 2; cf. E. C. Mossner, *The Life of David Hume*, 2nd ed. (Oxford: Oxford University Press, 1980), 73), and even if he had “read many Books of Morality” (HL 1. 14) at the time of his writing that book). The recently discovered “An Early Fragment on Evil”, titled as “Sect 7. Forth Objection” might be the only extant part of the book in question (destroyed by Hume about 1751). M. A. Stuart has shown by analysis of Hume’s handwriting and paper that this fragment could not be written before late 1730s (see M. A. Stewart, “An Early Fragment on Evil”, in M. A. Stewart and John P. Wright (eds.), *Hume and Hume’s Connexions*, 163–164), but, I think, it is possible that that fragment is a survived copy of a section of that book, made by Hume in late 1730s for some purposes of his own – for further work, for example, or to show it to somebody for evaluation.

passions. It should be noted, however, that in Part III of *The Philosophy of David Hume* titled “Detailed consideration of the central doctrines, taken in what may be presumed to have been the order of their first discovery” Kemp Smith begins just with Hume’s theory of passions. But it is not my purpose here to explain this seeming inconsistency. Instead I will try to show that we have good reason to believe that Hume indeed had written Book 2 of his *Treatise* before Book 1.³⁵

VI

This task is, of course, very difficult, because we know that Hume had heavily edited all the books of his *Treatise* after their initial composition. And he certainly tried hard to make the parts of his discourse to be tightly connected to each other, and to justify the places they occupy in the whole.³⁶ Moreover, in Book 2 there are numerous references to Book 1. Nevertheless, it seems that he had left some traces of the different sequence of their composition. I do not mean, however, those passages in Book 2 where he rather uncritically talks about “identical person” (T 2.2.1.2; SBNT 329)³⁷ and necessity of causes for events³⁸, which sound strange after his demolishing these doctrines in Book 1. The problem here is that such anomalies might be explained away. Someone might claim that he meant fictional identical person, for example³⁹, and, as for causality, Hume, as we know, stressed that he had not intended at all to show that some events had no cause.⁴⁰ Anyway, in this part of my paper I will try to find other evidence in favor of the

³⁵ Ben Mijuskovic, who criticized this Kemp Smith’s hypothesis, wrote in 1971 that it was “either an accepted interpretation or an open question”. See Ben Mijuskovic, “Hume and Shaftesbury on the Self,” *Philosophical Quarterly*, 21 (1971): 324. Nowadays, however, Kemp Smith’s approach is not anything like a common view. Of course, some authors approve this hypothesis, but without much discussion; see, for example, Jonathan Bennett, *Learning from Six Philosophers*, vol. 2 (Oxford: Oxford University Press, 2001), 197–198, or Lothar Kreimendahl, *Humes verborgener Rationalismus* (Berlin: Walter de Gruyter, 1982), 145–146. John P. Wright also sympathizes with this approach in one passage of his *Hume’s Treatise of Human Nature: An Introduction* (Cambridge: Cambridge University Press, 2009), 27, but later tends to dismiss it. In general, however, this hypothesis seems to be widely ignored or tacitly rejected. For such a rejection, see Annette C. Baier, *A Progress of Sentiment: Reflections on Hume’s Treatise* (Cambridge MA, Harvard University Press, 1991); Terence Penelhum, *Themes in Hume: The Self, The Will, Religion* (Oxford: Oxford University Press, 2000); and Claudia M. Schmidt, *David Hume: Reason in History* (University Park PA: The Pennsylvania State University Press, 2003). As for ignoring, so, for example, there is no discussion of this topic in David Fate Norton’s excellent “Historical Account of A Treatise of Human Nature from Its Beginnings to the Time of Hume’s Death”, *David Hume. Treatise of Human Nature: Critical Edition*. Vol. 2. *Editorial Material* (Oxford: Oxford University Press, 2007); as well as in David Fate Norton, and Jacqueline Taylor (eds.), *The Cambridge Companion to Hume*. 2nd ed. (Cambridge: Cambridge University Press, 2009), or in Saul Traiger (ed.), *The Blackwell Guide to Hume’s Treatise* (Oxford: Blackwell, 2006). And there is a deficit of special works on this issue.

³⁶ See for a discussion James A. Harris, “A Compleat Chain of Reasoning: Hume’s Project in A *Treatise of Human Nature*, Books One and Two” in *Proceedings of the Aristotelian Society*, cix, 2 (2009): 129–148.

³⁷ Cf. Norman Kemp Smith, *The Philosophy of David Hume*, V–VII. Kemp Smith uses the difference between Book 1 and Book 2 as regards Hume’s treating of the self as an evidence that Book 2 had been written before Book 1. For a critique of this idea of the essential difference between Book 1 and Book 2 in this respect see, for example, Don Garrett, *Cognition and Commitment in Hume’s Philosophy* (New York: Oxford University Press, 1997), 167–169. Note, however, that Hume himself points at the difference between his treating of the self in Book 1 and Book 2 (T 1. 4. 6. 5; SBNT 253). He tries to interpret this difference as a difference of the levels of our consideration of our self. This solution, among other things, might give him an opportunity to keep intact the text of Book 2.

³⁸ “...it is evident”, he suggests, by “reason alone”, that passions “must have some cause” (T 2.2.2.9; SBNT 336).

³⁹ Or “connected succession of perceptions”; see A. E. Pitson, *Hume’s Philosophy of the Self* (London: Routledge, 2002), 124. See also John Wright, *Hume’s Treatise of Human Nature*, 194. I will return to this issue later.

⁴⁰ T A Letter From a Gentleman 26; HL 1. 187.

primacy of the second book of the *Treatise* over the first – evidence of a statistical kind.

But before this let us pay attention to the fact that two first books of the *Treatise* have two sections under the same title – *Division of the subject*. Moreover, they have rather similar content. In both sections Hume makes a difference between two kinds of impressions. In the first of these sections (in Book 1) he calls them “impressions of sensation” and “impressions of reflexion”, in the second one (in Book 2) – “original” and “secondary” impressions. And in the second section he adds that this distinction is identical with that which he had made in Book 1.

Now the question is: which of these sections was composed first? He mentions Book 1 section in Book 2 section, of course, but it is clear that it does not solve our problem, because he could add this remark⁴¹ while editing the text. To find an answer to our question, we must look at these sections much more closely. Doing this we see other similarities here. For example, in both sections Hume says that he is not going to deal in Books in question with the impressions of the first kind (original impressions or impressions of sensation), because “The examination of [them] belongs more to anatomists and natural philosophers than to moral” (Book 1) and “the examination of them wou’d lead me too far from my present subject, into the sciences of anatomy and natural philosophy” (Book 2).

But there is a difference also. In Book 1 Hume claims that impressions of reflection are “deriv’d in a great measure from our ideas” or “arise mostly from ideas” (T 1.1.2.1; SBNT 7–8; my emphasis). He explains his position as follows: our mind saves in memory copies of our impressions of sensation, such as impressions of pleasure and pain, and “This idea of pleasure or pain, when it returns upon the soul, produces the new impressions of desire and aversion, hope and fear, which may properly be call’d impressions of reflection, because deriv’d from it” (T 1.1.2.1; SBNT 8). He suggests that this is a common mechanism of production of such impressions. It is interesting, however, that he does not discuss it in those parts of Book 2, which are devoted to those impressions that he mentions in his explanation. He calls here desire, aversion, hope, fear etc. “direct passions”, and at the beginning of his investigation into the direct passions, he gives an example of “immediate impression of pain or pleasure”, which – without any ideas – gives rise to a direct passion (T 2.3.9.3; SBNT 438). It is unlikely, however, that he had changed his mind: almost all other discussions of the causes of these passions here presuppose that goods or evils which immediately produce the direct passions are given as ideas. But the fact that he begins with a suggestion that ideas are of a secondary importance in this respect (T 2.3.9.2; SBNT 438), and with an example of an impression producing direct passions, indicates that he believed that this mechanism was a common one. And this explains why in *Division of the subject* in Book 2⁴² his claims seem to be much more modest than in the corresponding section of Book 1⁴³: “Secondary, or reflective impressions are such as proceed from

⁴¹ And a couple of other remarks, which can be easily identified if we agree that this section was revised and edited. I mean, first of all, T 2. 1. 1. 3, or at least a part of that paragraph, where he mentions his previous investigations.

⁴² Hereafter *Division of the subject* 2.

⁴³ Hereafter *Division of the subject* 1.

some of these original ones, *either* immediately *or* by the interposition of its idea” (T 2.1.1.1; SBNT 275). A bit later he repeats that secondary impressions arise “*either* from the original impressions, *or* from their ideas” (T 2.1.1.2; SBNT 276; my emphasis). In this Book he makes a distinction between direct and indirect passions, and he argues that indirect passions cannot arise without intermediate ideas. So we can interpret him as claiming that direct passions commonly arise immediately from original impressions. And that is indeed what he suggests here.

Now, all this would be rather puzzling if he had written *Division of the subject 1* before *Division of the subject 2* and before Book 2. Indeed, in such a case he would surely reproduce that reasoning concerning generation of the direct passions via ideas in Book 2, provided that this reasoning fits well with the actual (even if tacit) presuppositions of his theories. So, we can suppose that Hume had begun composing his *Treatise* with the Introduction; introductory section *Of the origin of our ideas* (that is, with some early version of the first section of Book 1 where he draws a distinction between impressions and ideas and outlines a relation of ideas to impressions); another introductory section, *Division of the subject 2*, where he promises to proceed with an analysis of the secondary impressions, and the main body of Book 2 itself where he deals with them. After a while (quite possibly after writing early versions of some sections of Book 3 *Of Morals*) he proceeded to Book 1. I think we have a reason to suppose that initially Hume had not wanted to write a lengthy book on understanding at all.⁴⁴ But this topic turned to be very interesting and he sunk in it. Soon he found out that operations of our understanding depend on a kind of inner instinct, resembling in that respect our passions, which are mostly based on an instinct to “unite itself with the good, and to avoid the evil” (T 2.3.9. 2; SBNT 438). And he had realized the importance of his results in that field. Therefore he begun thinking about where to place them. He realized that he had tacitly presupposed that most of our passions arise via ideas. In fact, it seems that he came to conclusion that our memories – on which our direct passions largely depend – should be better treated not as impressions (and he tends to treat them in this way in T 1. 3. 4. 1, 4; SBNT 83, 84; T 2. 3. 1. 17; SBNT 406, etc.) but as ideas (cf. later texts T_{App} 1. 3. 5. 4; SBNT 85; T 3 Advertisement; SBNT 455). And as his theory of understanding deals mostly with ideas that gave him a reason to put it before his theory of passions. He explained his decision in a new section *Division of the subject 1*.⁴⁵

While rearranging his texts Hume had to solve a terminological problem. He had classified passions as “secondary impressions”. But in order to stress their dependence on ideas he felt a need to give them another name. And, partly inspired by Locke, he found it – “impressions of reflection” (it seems that by “reflection” he

⁴⁴ Look, for example, at the title of T 1. 3. 2, *Of probability; and of the idea of cause and effect*. It is a correlate to T 1. 3. 1, *Of knowledge*; and Part 3 of Book 1 is titled *On Knowledge and Probability*. It seems that initially Hume wanted to restrict himself to those two sections, but then had changed his mind and expanded his plans. As a result he does not discuss probability at all in the above-mentioned section with the word “probability” in its title.

⁴⁵ If this picture would happen to be correct, it would disprove Paul Russell’s hypothesis that Hume had consciously modeled his *Treatise* after Hobbes’s *The Elements of Law*. This hypothesis presupposes primacy of Book 1. See Paul Russell, *The Riddle of Hume’s Treatise* (Oxford, Oxford University Press, 2008), 64–65. This would not mean, however, that Hume ignored Hobbes’s ideas and that they could not have made an important impact on him.

usually understood simply process of thinking, i.e. process involving some ideas). But having found a new name for secondary impressions he had to find a new correlate name for original impressions. And he called them “impressions of sensation”.

Now, we should not forget that all that has just been said was just a hypothetical reconstruction. But now we have a chance to confirm it. Indeed, if Hume had really invented his term “impressions of sensation” after writing Book 2 we might expect to find some discrepancy between his usage of the word “sensation” in Book 1, where it is used in the new sense⁴⁶, as a term which referent seems to be a faculty of mind directed to the original impressions⁴⁷, and in Book 2.⁴⁸

That is what we really see. If we analyze Hume’s usage of the word “sensation” in Book 2 of his *Treatise*, we find out that quite often he talks about sensation of passions, such as pride. Pride is a secondary impression, so this usage seems to be obviously incompatible with the usage adopted in Book 1. In fact, the situation is a bit more complicated, because in Book 2 Hume is close to identify sensation with a faculty of mind by which we feel pleasure and pain, and pleasure and pain are original impressions. As every passion is pleasant or unpleasant he might talk about sensation of passions meaning original impressions of pleasure and pain which might be connected with those very passions. But this surely would be a wrong interpretation because Hume explicitly claims that sensations “constitute [the] very being and essence” of passions (T 2. 1. 5. 4; SBNT 286).

So there really is a discrepancy in Hume’s usage of the word “sensation” between Book 1 and Book 2 of his *Treatise*.⁴⁹ And his usage of this word in Book 1 is surely of a later origin, because we see this very usage in his *Abstract* (T Abstract 9; SBNT 649), Appendix (T Appendix 17; SBNT 635), and Advertisement to Book 3. It should be noted, however, that later Hume had probably realized this discrepancy and tried to avoid it. In his first *Enquiry* he drops his distinction between impressions of sensation and impressions of reflection. He simply identifies sensations with impressions and makes a distinction between “outward” and “inward” impressions (EHU 2. 9; SBN 22), modifying his earlier descriptions.

Let us, however, return to Book 2 of Hume’s *Treatise*. We have just seen that it is quite possible that Hume had written the Introduction to the *Treatise* and Book 2 before Book 1. We know that Hume composed his *Treatise* in France in 1734–36⁵⁰. If he begun this process with Book 2 we might expect to find some

⁴⁶ Presumably as a result of careful editing of the text, parts of which could be written before invention of this term.

⁴⁷ In some passages it also refer to original impressions themselves.

⁴⁸ Where, as it is clear from *Division of the subject* 2, he could decide to restrict himself to a fragmentary editing.

⁴⁹ This fact has not remained unnoticed in literature, and Claudia Schmidt, for example, has paid attention to it in her *David Hume: Reason in History*, 165.

⁵⁰ See David Hume, *My Own Life*, ed. by Peter Millican (<http://davidhume.org/texts/mol>), 5. Hume claims that he composed his *Treatise* “first at Rheims, but chiefly at La Fleche, in Anjou”. In his extant letters he says twice that he composed his *Treatise* before he was twenty five (HL 1. 158, 187), so it seems that he finished the main work in Spring 1736. We know that in May 1735 he was in La Flèche already (see Ernest C. Mossner. “Hume at La Flèche, 1735: an Unpublished Letter”, in *The University of Texas Studies in English* 37 (1958): 30–3), so it seems that he wrote most parts of his *Treatise* during one year. Of course, we should not take those Hume’s remarks literally. I think we can interpret him as saying that he finished the first draft of the *Treatise* by May 1736. This does not exclude that he could rewrite some parts later or even write new parts or sections. We can be sure, for example, that he

similarities between that Book and the *Letter to a Physician* written in Spring 1734 several months before he went to France. And we can find them. When we read Hume's *Letter to a Physician* we may be surprised how often (for a rather short text) he talks about so called "spirits", ten times. Once he uses this word in a combination "animal sprits" (tiny particles in the nerves and in the brain) and it is quite probable that at least in a few other places he means by "spirits" such animal spirits.

Now, if we look at his *Treatise*, we see that this term is used here too. But not in equal proportions. Hume mentions "spirits" 14 times⁵¹ in Book 1 (0.05 per page), 19 times in Book 2 (0.1 per page)⁵², and never in Book 3. If we take into account that Hume does not mention "spirits" in his *Abstract* and uses this term in the first *Enquiry* only twice we could see a tendency here and safely claim that Book 2 contains texts written at the early stages of Hume's working at the *Treatise*.

The fact of Hume's usage of the term "spirits" reveals some important things about his way of doing philosophy at that time. We know Hume as a philosopher of a phenomenological and skeptical type, but it seems that at early stages of his philosophical career he was a different thinker. He took for granted the material world and tried to explain mental operations by brain processes. There are several examples of such explanations in Books 1 and 2 of his *Treatise*. It looks like as if he believed in existence of some exact mechanisms of movements of animal spirits directly corresponding to mechanisms producing our mental states, such as passions. Moreover, he probably considered former as basic and latter as derived. That is why he tried to explain passions by reference to movements of animal spirits.

It is interesting, however, that in Book 1 we find signs of a hesitation on his part as regards his using such explanations. If Book 1 had been written after Book 2 this fact is easy to explain. It is likely that under influence of Berkeley he had seen that we have reasons to doubt the existence of material world (to which those animal spirits belong). And in a few passages of Book 1 he even says that we have strong arguments against such an existence. This insight was a source not only of Hume's most intensive skeptical doubts but also of his phenomenalist turn⁵³. So it is no chance that in *Division of the subject 1*, written, as I have supposed, much later than *Division of the subject 2* and hence under possible influence of some phenomenalist considerations, he says that the impressions of sensation arise "in the soul originally, from unknown causes" (T 1.1.2.1; SBNT 7), while in the correspondent section in Book 2 he is much more explicit concerning their possible causes: original impressions "without any antecedent perception arise in the soul, from the constitution of the body, from the animal spirits, or from the application of objects to the external organs" (T 2.1.1.1; SBNT 275).

did that with Book 3 of his *Treatise* (see James Moore, "Hume and Hutcheson", in M. A. Stewart and P. Wright (eds.), *Hume and Hume's Connexions*, 38–39). And we can suppose he did the same with two other Books.

⁵¹ Or 16 times if we count the usages in a passage inserted from the Appendix, that we should not do. Hume used word "spirits" in the Appendix in a fragment which had to be inserted in a text where he also used it. And he used it not in an explanatory sense. This fact and that explanation of his usage will turn to be of some importance later.

⁵² I count by SBN edition of the *Treatise*.

⁵³ I mean his coming to the idea that there is no way out of our perceptions.

Note, that in both passages Hume uses the word “soul”. This word, unlike the word “mind”, has obvious substantialistic connotations.⁵⁴ But in Book 1 of the *Treatise* Hume denies that our self is an identical substance. In Book 2, however, he seems to believe in the existence of such identical persons⁵⁵. I have said above that we can argue that it is no more than seeming. However, let us suppose it is not.⁵⁶ Then we may expect that in Book 1 Hume would tend to diminish usage of the word “soul” in comparison to the usage of the word “mind”. In other words, we can expect to see in Book 1 a much larger ratio of mind-soul usages than in Book 2. And if this tendency continues in his later texts, for example in Book 3, we will have another argument in favor of that Book 2 had been written before Book 1.

The picture we see after counting words “soul” and “mind” in the contexts where their usage had not been determined by tradition⁵⁷ and was not figurative⁵⁸ not only confirms the above supposition but helps also to reconstruct the probable sequence of composing the Parts of Book 1 and Book 2 of Hume’s *Treatise*. Indeed, in Part 1 of Book 1 the ratio is 5, 7⁵⁹ (“mind” – 34, “soul” – 6); in Part 2 > 47 (“mind” – 47, “soul” – 0); in Part 3 – 21, 6; in Part 4 – 24, 5.⁶⁰ In Book 2 the ratio in Part 1 is 6, 8⁶¹; in Part 2 – 8, 2⁶², and in Part 3 – 6, 6. Compare the *Letter to a Physician* (2, 5), *Abstract* (16, 5), *Appendix* (> 34), and Book 3 (20)⁶³.

This method is not so precise as we might desire, of course,⁶⁴ but still its results indicate⁶⁵ that Hume composed his *Treatise*⁶⁶ starting with Part 1 of Book 1

⁵⁴ Cf.: “Thus we feign the continued existence of the perceptions of our senses, to remove the interruption: and run into the notion of a *soul*, and *self*, and *substance*, to disguise the variation” (T 1. 4. 6. 6; SBNT 254). It is no chance that in the influential Chambers’s *Cyclopaedia* we also find a clear insight that “The Soul is a spiritual Substance”, see Ephraim Chambers, *Cyclopaedia: or, An Universal Dictionary of Arts and Sciences*, vol. 2 (London, 1728), 97.

⁵⁵ “...Immediate object of pride and humility is self or that identical person, of whose thoughts, actions, and sensations we are intimately conscious” (T 2. 2. 1. 2; SBNT 329); see also T. 2. 2. 2. 2; SBNT 333.

⁵⁶ In fact, the seeming is rather strong. His denial in Book 1 that we are in “every moment intimately conscious” of our self and its “continuance in existence” (T 1. 4. 6. 1; SBNT 251) is in an obvious contrast with phrases like “Tis evident, that the idea, or rather impression of ourselves is always intimately present with us” (T 2. 1. 11. 4; SBNT 317).

⁵⁷ That is, ignoring phrases like “immortality of the soul” in Parts 3 and 4 of Book 1 of the *Treatise*. The word “mind” usage will serve as a kind of index of a density of mentalistic contexts, and then we will evaluate how often Hume used the word “soul” in average contexts in such cases where he could say the same using the word “mind”.

⁵⁸ I also do not take into consideration Hume’s footnotes because as a rule they are of a later origin than the main text. And I ignore texts which had been added by Hume in the Appendix and inserted in Book 1 in modern editions.

⁵⁹ Counting this usage here with the *Division of the subject* 2 (ignoring *Division of the subject* 1) we would have the ratio 4, 9.

⁶⁰ I do not count Hume’s usages of the word “mind” (as well as “soul”) in Section *Of the immateriality of the soul*.

⁶¹ I do not take into consideration Section *Division of the subject* 2. (If to take it, the ratio would be 5).

⁶² In this Part I do not count one passage where Hume talks about soul as “animating principle”, in a figurative sense.

⁶³ Cf. also his *Essays, Moral and Political* (1741), where the mind-soul ratio is 20 (“mind” – 20, “soul” – 1).

⁶⁴ Partly because some of the texts mentioned are simply too short to be statistically totally compelling. Another problem with this method is that its results in a sense depend on how we count, that is, which parts of the text we decide to consider as units. This problem, however, can be solved along the lines I indicate in the main text.

⁶⁵ It might be objected that different contexts presuppose different frequency of the usage of the same words, and we have just compared their usage in different contexts. So from the differences we have noted it does not follow that the parts with the lesser ratio had been written earlier. This might be true in some cases (and that is why I have ignored Section *Of the immateriality of the soul* in counting the usage of the words in question), but it is not true in the present case: the contexts of Parts 1 and 3 of Book 1 are similar but we see a big difference here. And the contexts of Book 2 and Book 3 are rather similar as well, but there is a difference in that usage too (note, that it is useless to compare Book 2 with *Dissertation on the Passions* (1757), because the latter work is mostly a compilation).

⁶⁶ In what follows I will not discuss the possible time of composing by Hume those “religious” sections of the *Treatise*, which he arguably removed from the *Treatise* while editing it in 1737–8.

and Parts 3, 1, and 2 of Book 2. If, however, we take into account our recent considerations we can suppose that Part 1 of Book 1 had been written not in a single move. Its oldest part, that is, Section 1 (*Of the origin of our ideas*), initially integrated with *Division of the subject* 2, has the mind-soul ratio 2, 5; as for Sections 4–7, so they could be written after Parts 1 and 2 of Book 2 (their ratio is 11, 5).⁶⁷ I have not mentioned Part 3 of Book 2. The problem with it is that the sections of which it consists can hardly be interpreted as parts of a unified whole. It means that we should treat them as separate units. We see four clear-cut blocks here. First of them (Sections 1–2 with the ratio > 19) contains Hume’s discussion of liberty and necessity, which presupposes his theory of causation elaborated in Part 3 of Book 1. The second block (Sections 3–8 with the ratio 2, 4) is devoted to discussion of the causes of actions and violent passions. The third block is Section 9 (ratio 16) where Hume considers our direct passions, the fourth – Section 10⁶⁸ (ratio > 9).

If we agree that these texts must be considered separately, we should admit that Sections 3–8 of Part 3 of Book 2 are one of the oldest segments of Hume’s *Treatise of Human Nature*. So it seems that Hume’s composing of the *Treatise* had started with the Introduction (ratio 3), first Sections of Book 1 and 2 (*Of the origin of our ideas* and *Division of the subject* 2), and Sections 3–8 of Part 3 of Book 2. Then he wrote Parts 1⁶⁹ and 2 of Book 2, and proceeded to Sections 4–7 of Part 1 of Book 1. Sometime after that, he wrote Section 9 of Part 3 of Book 2 and Parts 3 and 4 of Book 1, making discoveries concerning the nature of our causal inferences and realizing radical skeptical challenges to his system. Part 2⁷⁰ was written even later than most or all sections of other Parts, so maybe Hume had not to edit here his phrase “the present year 1738” (T 1. 2. 2. 4; SBNT 31).⁷¹

Before going further we need to clarify one point. The method I have just proposed presupposes that Hume became using the word “soul” (as a designation – along with “mind” – of our mental life or self) less because he realized that our self

⁶⁷ Section 3 *Of the ideas of the memory and imagination* seems to be written later (along with *Division of the subject* 1) as it (as well as *Division of the subject* 1) contains Hume’s later conception of memory which presupposes that it deals not with impressions but with ideas (cf. the title of T 1. 3. 5 – *Of the impressions of the senses and memory*). Note, that the fact that Hume uses the word “soul” twice in *Division of the subject* 1 can be explained by the supposition that he unconsciously modeled this section on Section *Division of the subject* 2, where he uses it three times.

⁶⁸ About the love of truth.

⁶⁹ Except Section 10 of Part 1 which seems to be mostly written later.

⁷⁰ About ideas of space and time. The mind-soul ratio in this Part indicates that in that time Hume tends not to use the word “soul” in nonspecific contexts at all, and his Appendix confirms this. He uses this word twice in his *Abstract*, but in this work he could try to detach himself stylistically from the author of the *Treatise*. As for Book 3 (1740), in which this word is used five times, so we should remember that while it was heavily edited by Hume in 1739–40, he surely had left in the final texts some of its early parts where this ratio should not have been very high. The final ratio (20) reflects an impact of these factors. And as this ratio does not radically change after subtraction of those parts which were probably rewritten, we can suppose that at least some parts of Book 3 were written between writing Parts 1–2 of Book 2 and Parts 3–4 of Book 1 – this could explain and fill a gap between the mind-soul ratio of Part 2 of Book 2 and Part 3 of Book 1. As for Hume’s related later works, so they indicate that in a time he became more tolerant to using the word “soul”. The mind-soul ratio in the first and in the second *Enquiry* is about 20 (in the second *Enquiry* we must exclude three cases of Hume’s usage of the word “soul”; once he uses it in a figurative sense, once – in a phrase “mortality of the soul”, and once when he refers to Plato’s theory of the soul.)

⁷¹ It is more likely, however, that he wrote it mostly about 1737. This Part is one of the few obviously Bayle-inspired parts of the *Treatise* (along with T 1. 4. 5), and Hume recommended to his friend Michael Ramsay to read Bayle for understanding his future book in August of 1737; see Tadeusz Kozanecki, “Dawida Hume’a Nieznane Listy W Zbiorach Muzeum Czartoryskich (Polska),” *Archiwum Historii Filozofii Społecznej* IX (1963): 127–41.

is not a soul as a simple substance, but a bundle of perceptions.⁷² But our results demonstrate that this process began before his writing Part 4 of Book 1 of his *Treatise*, where he tries to prove that our self is not a substance. In fact, however, there is no problem here. It is quite natural to suppose that before presenting his theory in Part 4 of Book 1 he spent a lot of time thinking about the nature of the self. And he had a reason to do so parallel to writing Book 2 *Of the Passions*, because his theory of indirect passions is based on a conception of relation which some entities have to bear to our self in order to produce some of such passions. And the more he thought about the nature of the self the less he was satisfied with the traditional conception of it as a kind of a substance.⁷³ And this was reflected in his usage of the word “soul”. So it seems that the method works quite well.

As a result of applying this method we can specify Hume’s philosophical views of the period when he began composing his *Treatise*. I have mentioned above that he took for granted the material world and tried to explain our mental operations by workings of animal spirits. But his view was far from that of materialism. It seems that he accepted the existence of identical mental substances, souls. That is why it would be more appropriate to call his position “naturalistic dualism”⁷⁴. So, it is not a surprise that in the earliest part of the *Treatise*, evaluating by the concentration of the “animal spirits”, that is, in Sections 3–8 of Part 3 of Book 2, we see also one of the lowest mind-soul ratios. And we can be quite sure that at least some sections of this seemingly unified block stem from about 1734.

I have just said that Sections 3–8 of Part 3 of Book 2 seem to be a unified block. Indeed, Hume himself treats them as such, because at the end of Section 8 he summarizes his claims made in these sections (T 2. 3. 8. 13; SBNT 437–438). Still there is a puzzle here, and in order to solve it we must, I believe, to make further fragmentation of the text. The puzzle is that while he widely uses the terms “spirits” and “soul” in these sections, he uses them in a very special manner. In short Sections 7–8, devoted to the influence of a distance on our passions, he uses the word “soul” six times (more than in Book 3 of the *Treatise*, Part 2 of Book I, and the Appendix taken together) without any obvious reason.⁷⁵ At the same time he does not use the word “spirits” in these sections at all. He uses this word eleven times (more than in the rest of Book 2 and Book 3 together) in Sections 4–5, however (where, among other things, he discusses the idea of a predominant passion).⁷⁶ My guess is that while Sections 4–5 stem from a time when Hume wrote the *Letter to a Physician*, and when he was under the charm of the idea of physical explanations of the mental operations⁷⁷, Sections 7–8 contain even earlier texts, probably written about 1732 (this is a time at which, according to Hume, he begun a real

⁷² In fact, we might come to the similar conclusions without such presupposition (we could just count the words and then make some conjectures concerning the probable sequence of different Parts). But as it surely makes these results more convincing it is worth to defend it. And these defense helps to clarify a number of quite important points.

⁷³ Already in Part 2 of Book 2 we read that “Ourself, independent of the perception of every other object, is in reality nothing” (T 2. 3. 2. 17; SBNT 340), which sounds as a first version of his later theory of the self from Book 1.

⁷⁴ About this term – cf. David J. Chalmers, *The Conscious Mind* (Oxford: Oxford University Press, 1996), 122–129.

⁷⁵ The mind-soul ratio here is 0, 8.

⁷⁶ The mind-soul ratio in these sections is 5.

⁷⁷ John Wright identifies a possible source of Hume’s interest in such explanations some ideas of Bernard Mandeville. See his *Hume’s Treatise of Human Nature*, 8–9.

work on the *Treatise*⁷⁸). Hume's position at that time seems more dualistic and mentalistic, and he confined himself to showing that, as he said later, "in the production and conduct of the passions, there is a certain regular mechanism" (DP 6. 19) as regards causes of their arising or turning from the calm to violent, for example.

So, we can reconstruct the phases of evolution of Hume's general views about 1732–34 in the following way: (1) soul-body dualism in ontology and the idea of an experimental account of mechanisms of the passions as phenomena of the soul (about 1732)⁷⁹; (2) search for physical explanations of such phenomena and beginning of the process of dissatisfaction with the idea of soul as a substance (about 1734); (3) limiting the area of physical explanations due to their imperfection and elaboration of the experimental account of the machinery of the passions, centered around the mechanisms of associations of impressions and ideas.

The latter stage (at which he probably wrote Section 3 of Part 3 of Book 2 and so made a first attempt to use a kind of clarifying syllogism while talking about causes of our actions, but had not yet realized a specificity of this new methodological device) corresponds mostly to the introductory sections of the *Treatise* and to the basic sections of Parts 1 and 2 of Book 2, where he deals with indirect passions, such as pride, humility, love, and hatred, and gives a complex associative account of their generation and transformation. It does not mean, however, that this order of composition necessarily implies that Hume had planned it as an order of final presentation of his ideas. We have some reasons to suppose that the section on the direct passions, written after the sections on the indirect ones, was initially intended to precede them in the final text. Indeed, it is natural to suppose that he could begin exposing his theory of passions with an analysis of the direct passions. This way is natural because they are more elementary than indirect ones. And in his *Dissertation on the Passions* he did just that.⁸⁰ We have already seen that when writing of arising of secondary impressions either from the impressions or from ideas in the *Division of the subject* 2 Hume probably meant arising of direct and indirect passions. And as it was an introductory section we can suppose that such a sequence indicates that he intended to begin with direct passions. Then he saw that they mostly presuppose ideas as well and rearranged the text: his theory of indirect passions justly seemed to him much more original. In fact, we can suppose that he did not start writing the book on passions with his account of direct passions simply because initially he had not anything interesting to say about them. As he proceeded in developing his theory of understanding, however, this situation had changed. The fact that he probably wrote the section on the direct passions (the structure of which suggests that he favors that version of his *Treatise* which opens with Book 2) approximately at the time of writing Part 3 of Book 1 confirms that

⁷⁸ HL 1. 158.

⁷⁹ At this stage Hume probably wrote his *Historical Essay on Chivalry and modern Honour* – see M. A. Stewart, "The Dating of Hume's Manuscripts," in Paul Wood (ed.), *The Scottish Enlightenment: Essays in Reinterpretation* (University of Rochester Press, 2000), 267–314; and John Wright, *Hume's Treatise of Human Nature*, 11–13.

⁸⁰ Of course, he could have additional reasons to rearrange the material in his *Dissertation* as well. See John Immerwahr, "Hume's Dissertation on the Passions", *Journal of the History of Philosophy*, 32, 2 (1994): 225–240.

he decided to rearrange the *Treatise* and begin it with his account of understanding at a comparatively late stage.

Now, let us finally look at Part 2 of Book 2 of the *Treatise*, which, as our results suggest, had been written after completion of the most or all other sections of Books 1 and 2. If this is true, we can expect to find in this Part some independent confirmation of this. Let us see if this is possible. If we agree that while writing the *Treatise* Hume had drifted from a kind of positive philosophy of a realistic type to a radical skepticism, and if we agree that this skepticism was closely connected with a kind of phenomenalist turn, we can expect to find in Book 2 some signs of his phenomenalist moods. And, indeed, in this very Part we find the most phenomenalist passage of the whole *Treatise*: “Let us fix our attention out of ourselves as much as possible: Let us chace our imagination to the heavens, or to the utmost limits of the universe; *we never really advance a step beyond ourselves*, nor can conceive any kind of existence, but those perceptions, which have appear’d in that narrow compass. This is the universe of the imagination, nor have we any idea but what is there produc’d” (T 1. 2. 6. 8; SBNT 67–68; my emphasis).

In this Part we can also find a few general considerations about the type of author’s philosophy, which resemble those we see in *Section Conclusion of this Book*, which obviously is of a very late origin. In Part 2, as well as in *Conclusion of this Book* (see T 1. 4. 7. 2, 12; SBNT 264, 271), Hume characterizes his way of philosophizing as “my philosophy”. He says also that he “pretends only to explain the nature and causes of our perceptions, or impressions and ideas” (T 1. 2. 5. 26; SBNT 64). The context of this passage is that when we try to comprehend the properties of bodies themselves, we face insurmountable difficulties, but when we look at the appearances of these bodies to our senses, we can easily resolve all of them. This position is quite specific for this Part, and it has clear parallels only in Section 5 of Part 4 (where, it is interesting, there is yet another example of “my philosophy”), near the end of Book 1⁸¹, and in the Appendix, where Hume says that he “had entertain’d some hopes, that however deficient our theory of the intellectual world might be, it wou’d be free from those contradictions, and absurdities, which seem to attend every explication, that human reason can give of the material world” (T Appendix 10; SBNT 633). If so, it is natural to suppose that in the Appendix he continues the line of thought presented in Part 2 and in Section 5 of Part 4.⁸² And this means that Part 2 might be completed after most sections of other Parts.

These considerations could be important because they may help to clarify the evolution of Hume’s skeptical views in 1734–48. I have already mentioned that his radical skeptical moods were probably initiated by Berkeley’s ideas.⁸³ Hume was convinced by Berkeley’s argument concerning subjectivity of the primary qualities, like extension or solidity. If such qualities are subjective then bodies exist in their capacity as our perceptions only. Hume also believed that “the plainest experience” can show that, while our perceptions might exist separately (this di-

⁸¹ See also T 2. 2. 6. 1–2; SBNT 366. These paragraphs of Book 2 are probably also of a very late origin.

⁸² T 1. 4. 5. 1; SBNT 232.

⁸³ Hume himself suggests this (EHU 12. 15, note; SBN 155), and many commentators tend to agree.

rectly follows from his Separability Principle), in fact they are tied in bundles, that is, depend on us (cf. T 1. 4. 2. 44; SBNT 210). On the other hand, we firmly believe that bodies exist independently of ourselves. So, here is an obvious contradiction, and contradictions feed skeptical moods. Hume mentions this “direct and total opposition betwixt our reason and our senses” in T 1. 4. 4. 15. It is interesting, however, that in the concluding skeptical section of Book 1, when talking about this contradiction he refers not only to T 1. 4. 4, but also to T 1. 3. 14, where he considers an inclination of our reason to externalize our inner data. It seems as if he wants to strengthen his claim by reference to this inclination. But why? The probable answer is that at the time of writing the concluding section of Book 1 Hume had realized that *in his exposition* the thesis and antithesis of that contradiction had no equal weight: while our belief that bodies do not exist independently of us was based on what he considered as very powerful arguments, our belief in an independent existence of bodies (as he proved in Part 4 of Book 1) seemed to be a result of an elementary misidentification of some similar entities (see especially T 1. 4. 3. 56; SBNT 217).⁸⁴ But if thesis and antithesis have no equal weight they cannot form a basis of a skeptical position. Nevertheless it looks like he still believes that they *have* the equal weight⁸⁵ and therefore tries to find some way to show this. That is why he refers in T 1. 4. 7 to that inclination or tendency. But he could feel the need for additional reinforcement of that thesis, because his account of that inclination was rather unelaborated and it was not at all clear that it could be applied to the present case. And it seems that in Part 2 of Book 1 he makes his final attempt to solve this task, using his time-tested method of explanation of mental operations by workings of the animal spirits (it is quite probable that such explanation was a hidden source of that thesis all the time). We should not forget that his position in this Part is very phenomenalistic, so it is no surprise that he says here that, in general, he tends to avoid such explanations because they go beyond the limits of experience (T 1. 2. 5. 20; SBNT 60). But he suggests that time has come when he really needs them. And he tries to show that our misidentifications of similar items could be explained by some natural movements of animal spirits (*ibid.*).⁸⁶ If this is so, our belief in the independent existence of bodies might be no less necessary than our belief that they do not exist in such a way. But soon after providing this explanation Hume had probably realized that this way leads to nothing. Indeed, to reinforce our belief in an independent existence of bodies by assuming their existence (as animal spirits) is simply to beg the question.

After realizing the uselessness of such explanations in that case (which might happen when he was finishing Part 2 of Book 1) Hume seemed to change his

⁸⁴ We consider consequent similar spatial impressions as identical and then, noticing their distinction, assume that they are still identical, which means that they should have had a continued existence when we did not perceive them.

⁸⁵ He says that operations producing these conflicting beliefs are “equally natural and necessary” (T 1. 4. 7. 4; SBNT 266).

⁸⁶ While he uses this explanation for a different purpose here he indicates that it will be useful in other sections too. And when he discusses the origins of our belief in existence of bodies in T 1. 4. 2 he refers to that place in T 2. 5 (T 1. 4. 2. 32, note; SBNT 202). He says here that in T 2. 5 he has “already prov’d and explain’d” that principle. It is quite plausible that he inserted this passage (as well as the relevant discussion in Part 2) somewhat later, because otherwise he writes here in such a manner as if he has just introduced this important principle (T 1. 4. 2. 32; SBNT 203), that is, a principle, according to which we have a natural tendency to identify resembling or connected entities.

view on this topic, and decided to go even further in the direction of phenomenalism. He suggested that while we could not talk consistently about bodies as something independent of us, there is no contradiction between our belief in bodies as independent entities and as something that exists only in the actual perception – simply because the former is illegitimate. In other words, he was ready to admit that there is nothing at all except our perceptions, but in the world of perceptions there are no contradictions. This is the very position he outlines at the end of Part 2 of Book 1, and this is the view to which he refers in Section 5 of Part 4 (I think it is quite possible that this section had been written later than other sections of Part 4⁸⁷, and even after Part 2) and in the Appendix. But, as Hume suggests in the Appendix, this view has led him to a “labyrinth” when he saw that now he could not make consistent his theory of the self.⁸⁸ He famously claims here that he cannot “render consistent” two principles: that “our distinct perceptions are distinct existences, and that the mind never perceives any real connexion among distinct existences”. He means that as (1) our perceptions do not belong to “something simple and individual” and as (2) we see no inner connection between them, we have no way to explain how they can be connected “in our thought or consciousness” (T Appendix 20; SBNT 636). This claim, however, makes sense only if he rejects a possibility of explaining such connections by workings of the animal spirits. Indeed, in Part 2 of Book 1 he clearly states that such workings *can* explain the unity of perceptions (T 1. 2. 5. 20; SBNT 60).⁸⁹ But if I am right, and if soon after saying this he had realized that such explanations ceased to be an option for him, he really should face the problem which he stresses in the Appendix.⁹⁰

⁸⁷ Section 5 looks like a continuation of Section 4, but, in fact, it presupposes a reinterpretation of its conclusions in the way indicated above. And Section 5 is in a conflict with Hume’s conclusions, summarized in Section 7, which also can be treated as a continuation and development of his results obtained in Section 4 (and Section 1) of Part 4.

⁸⁸ I mean his second thoughts on his theory of the self. In another fragment of the Appendix (T_{App} 1. 2. 5. 26, note), presumably written earlier for insertion in Part 2 of Book 1, he seems to approve his former line of thought. Indeed, this fragment is a part of Hume’s addendum to the Appendix. The main body of the Appendix, including the second thoughts on the self, has a kind of introduction and the end, and from the start is presented as an appendix. As for that additional part, so it was probably written early in 1740, when Hume hoped to arrange the second edition of the *Treatise* (cf. HL 1. 38-39). In other words, it seems to be written when Hume had not planned his Appendix yet. If this is so, then it was written before his text on the self. See for a discussion Udo Thiel, *The Early Modern Subject: Self-Consciousness and Personal Identity from Descartes to Hume* (Oxford: Oxford University Press, 2011), 398.

⁸⁹ He claims here that “relations of *resemblance*, *contiguity* and *causation*” are “principles of union among ideas”; that he could succeed in “explaining the relations of ideas” by appealing to animal spirits; and *gives* a general schema of such explanations and a detailed explanation of mistakes (discussed above) connected with those principles. In Section on personal identity, criticized in the Appendix, he explains our belief in identity of our self by applying the very “three relations above-mention’d”. He claims that “that identity depends” on “these three relations of resemblance, contiguity and causations” as “uniting principles in the ideal world” (T 1. 4. 6. 16; SBNT 260). So, the whole puzzle can be reformulated in such a way: why in the Appendix Hume rejects those methods of explanations of his principles of association which he has used earlier? And in such a form this puzzle can be definitely solved.

⁹⁰ This interpretation is close to that given by Robert Fogelin. He also claims that Hume could not make his principles consistent because of his “phenomenalist stance”. See Robert J. Fogelin, *Hume’s Skeptical Crisis: A Textual Study*, 121-124; cf. his “Hume’s Worries about Personal Identity” in his *Philosophical Interpretations* (New York: Oxford University Press, 1992), 92. Fogelin, however, is not sensitive to the relevant changes of Hume’s position during his writing of the *Treatise*, and therefore his interpretation is not so convincing as it might be. His views on this issue were criticized by Abraham Roth, who tried to show that if Hume really was dissatisfied with his theory because of the lack of explanation of the connection of perceptions, his worries would be extrapolated to the other parts of his account of understanding (Abraham Sesshu Roth, “What was Hume’s Problem with Personal Identity” in *Philosophy and Phenomenological Research*, 61, 1 (2000): 93). We can partly agree with this, but it does not contradict Hume’s position as he says that he has found inconsistencies in his account of intellectual world, and this can

In other words, Hume had realized that if we accept that our belief in the independent existence of the bodies is an illegitimate one, and, consequently, that we must accept the thesis that there is nothing except our actual perceptions, then we should admit (assuming that there is no soul as a uniting substance, and no inner connections between our perceptions – these assumptions Hume still holds as true) that our perceptions could not be connected. But, as a matter of fact, they are connected. And this is an inconsistency. In the Appendix Hume says that this is a new evidence in favor of skepticism, and for this reason we can consider the Appendix as the highest point of his skepticism. But this was not his final position. In the first *Enquiry* he makes a step which seems to be a natural conclusion derived from the above premises. Indeed, if the thesis that the belief in the independent existence of bodies is an illegitimate one leads us to inconsistencies we can deny it. To deny it is to say that this belief results not from cognitive mistakes, but from some strong and a kind of basic instinct. And Hume says something like this in his *Enquiry* (EHU 12. 7–9; SBN 151–2).⁹¹ This again could lead to a contradiction in our beliefs, but in the first *Enquiry* Hume suggests that it can be ignored or even solved.⁹²

This account of a possible evolution of Hume's radical skeptical views⁹³ was intended to show that the idea that Part 2 of Book 1 was written after most other parts of Books 1 and 2 may be plausible on independent grounds. And this, in turn, shows that the method of comparative dating of Parts of the *Treatise*, used in this paper, seems to work. But the very method shows that Book 2 of the *Treatise* had been written before Book 1. And that is what I wanted to prove in this section.

VII

So I hope that the above reasons give a sufficient confirmation of the hypothesis that Hume had written the Introduction to his *Treatise* and Book 2 *Of the Passions* before writing Book 1. And this means that by writing of the science of human nature in that Introduction he really meant his theory of passions. But this theory is an experimental one. And that is why Hume suggested that the only basis

be interpreted in a broad sense. On the other hand, elsewhere Hume could simply *use* the principles of association, but he might *explain* their possibility by appealing to the self as a substance. As his own theory of the self cannot provide such an explanation, this theory is in a real trouble and can hardly be treated as a consistent one. For recent reviews of other explanations of Hume's second thoughts see Pitson, *Hume's Philosophy of the Self*, 66–80; Jonathan Ellis, The Contents of Hume's Appendix and the Source of His Despair, in *Hume Studies* 32, 2 (2006): 195–232.

⁹¹ Annemarie Butler tries to show that there is no big difference in this respect between the *Treatise* and the *Enquiry*, because Hume mentions the relevant “instinct” in his *Treatise* – see her “Natural Instinct, Perceptual Relativity, and Belief in the External World in Hume's *Enquiry*”, in *Hume Studies*, 34, 1 (2008): 115–158. It is clear, however, that he describes this instinct differently in the *Enquiry*, and never says that it is based on any obvious mistakes.

⁹² Presumably by notion of “unknown, inexplicable *something*, as the cause of our perceptions; a notion so imperfect, that no sceptic will think it worth while to contend against it” (EHU 12. 16; SBN 155). We should remember, however, that this passage was added by Hume only in 1777 ed. In the previous editions it seemed as if he considered this contradiction as a kind of not very important anomaly that could be ignored without much consequence.

⁹³ I have discussed here only Hume's skepticism “with regard to the senses”. As for his radical “skepticism with regard to reason”, it could be at least partly induced by his discovery of the impotence of our reason, connected, on the one hand, with the fact that reason cannot initiate desires, on the other – that it does not direct our empirical inferences. From these discoveries it was easy to make a final step and claim that reason cannot even support itself. This sort of skepticism is notoriously absent in the first *Enquiry*, and this emphasizes the positive spirit of that work.

of the science of human nature is experience. But in his first *Enquiry* Hume identifies the science of human nature with the moral philosophy as such, including the account of understanding. As in his theory of understanding he widely uses syllogisms he cannot claim here that the science of human nature is based only on experience. But it does not mean that he had changed his methodology in the *Enquiry*. What he had actually changed was his treating of the science of human nature. Still this is not the whole story. We have seen that the main text of the *Treatise* indicates that Hume had already changed his conception of the science of human nature before publication of Books 1 and 2 of this work.⁹⁴ So, if we set the Introduction aside and look only at Books 1 and 2, we see that under the “science of human nature” he understood not only his theory of passions but also his theory of understanding, where he widely used syllogisms to obtain truths about human nature. But something like an inertia had prevented him to realize that he must make corrections in his descriptions of his methods in the Introduction to the *Treatise*. In a few years, however, he fully realized that and made some corrections in the *Enquiry*.

Realizing the importance of syllogistical, or deductive methods in the science of human nature did not lead Hume, however, to developing an explicit theory of such deductions. And it seems to me that the lack of such a theory was one of the main reasons why such deductions of a Humean type were obviously underestimated in the numerous discussions concerning philosophical methodology. And in this final part of my paper I will try to show that they have a potential even now.

We have seen that Hume’s deductions in Book 1 of the *Treatise* and in his first *Enquiry* help us to clarify the real relations between our ideas of some operations of the mind, for example, between demonstrative and factual reasonings. Such reasonings produce beliefs of a different kind. Beliefs are mental states which can be expressed in some statements. Note, that there is no need to restrict an analysis of the relations of statements to a simple resemblance between them. It is possible, for example, to interpret equivalence (as a high degree of resemblance) of statements or propositions as their mutually implying each other, and implication gives us an idea of dependence between propositions or statements. Now, it is a common place in analytic philosophy that our beliefs depend on other beliefs. If, for example, I look at a clock and form a belief that it is 1 p. m., this belief presupposes another one – that the clock works properly, etc. It is interesting, however, that analytic philosophers have not paid much attention to analysis of dependence among our natural beliefs, that is, our deepest beliefs of an ontological kind. One notable exception is “descriptive metaphysics” of Peter Strawson and his followers.⁹⁵ Strawson, however, came to his views from Kant’s philosophy. But Humean way to something similar is equally possible. And Hume himself had tried to show the dependence of our belief in identity on our causal beliefs (T 1. 3. 2. 2; SBNT

⁹⁴ So in this respect the difference between the *Treatise* and the first *Enquiry* is not an important one (cf. HL 1. 158).

⁹⁵ See for example, P. F. Strawson, *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen, 1959).

74).⁹⁶ To see how this project could be further developed let us take as an example two statements expressing our natural beliefs and show how their relation can be clarified by a deduction: (1) “I believe that the course of nature will not change”; (2) “I believe that every event has a cause”. I choose them because Hume had discussed both at length. He, however, did not fully explain their relation. This relation is not quite clear from the start. And even if we clarified it we might not grasp it directly. But we can prove that the second statement depends on the first. Let us suppose that we believe in the existence of causeless events. The cause is an event *A*, which always immediately precede another event, *B*, the effect. If we believe in the existence of causeless events we believe that there are no such events immediately preceding (causeless) event *B* reproducing of which would be always accompanied with *B* in the next moment of time. But this belief is in an obvious conflict with our belief that the course of nature will not change. So, if we believe that the course of nature will not change we also believe that every event has a cause.

Now, the belief that the course of nature will not change is an expression of our belief in the exact correspondence between the past and the future experiences. That is, if we believe that the course of nature will not change we believe that if any event given in experience in an environment would be reproduced in the same environment it would be followed by the event, experienced after it at its first occurrence. As the belief in causality depends on this belief, and as we surely believe in such correspondence in ordinary life, we also believe that every event given in experience has a cause which might be given in our experience (if not, we would not believe that reproduction of an experiential state of affairs preceding an event which has a non-experiential cause would be followed by the event in question).

Hume was sure that causal inferences are involved in mechanisms of sympathy, which helps us to be a kind of mirrors of other people (T 2. 1. 11. 2–8; SBNT 316–20). He believed that we know passions of other people only by their external effects. And if we agree with Hume and with most contemporary philosophers that such qualitative mental states of other people cannot be given in our experience, we must believe that their behavior has publically experienced, that is, physical causes. But sympathy presupposes our belief in the existence of other minds. We cannot observe qualitative states of other people, but still we naturally believe that such states exist. Now let us suppose that we believe that mental states of other people have no causal impact on their behavior. In such a case those mental states would be unnecessary for them. And the world with such mental states would be overloaded with superfluous entities and would be much less elegant and simple than the world without them.⁹⁷ Our belief in conformity between the past and the future, however, arguably implies another belief – in the simplest possible picture of the world we live in. Then, if we believed in the causal impotency of possible mental states of other people, we would not believe in the existence of other minds.

⁹⁶ Later he modified his view on this subject (cf. T 1. 4. 2. 21; SBNT 197). Note, that here he discusses our belief in “continu’d existence of objects”, but belief in identity is connected with such belief in a continued existence.

⁹⁷ Of course, in this world would exist at least one non-elegant entity – the epiphenomenalist. But it would be better to suppose for her that she is the only oddity in the world, than take the whole world as a one huge oddity.

And if we actually believe in their existence, we also believe that their qualitative states have a causal impact on their behavior.⁹⁸

So we see a conflict between our natural beliefs. And to solve it we might say that mental states of other people (and ourselves) are naturally supervenient⁹⁹ on physical aspects of reality in such a way that in a possible world, differing from the actual one only by the lack of actual qualitative mental states, human beings would not behave themselves in the same way as they behave in the actual world. Such a solution gives its due to both conflicting beliefs and avoids contradictions.

Indeed, we can say that human behavior is determined by physical factors, because in the actual world these factors are sufficient for producing human behavior. At the same time, we can claim that mental states have a causal relevancy, because without them those physical factors would not produce the same behavior. The physical factors are sufficient nevertheless because the mental states are supervenient on them. Of course, we must explain why some physical systems need mental states for realization of their causal powers while others, as we believe, do not need anything like this. And such explanations might give us a clue to solving the mind-body problem. But this is not my purpose here to make attempts of such explanation. I have just wanted to show that we can use deductions of a Humean type to clarify relations between our natural beliefs and that such clarifying can open a way to plausible solutions of some deep metaphysical problems.

Note also that these brief remarks have a relation to the much discussed topic of the future of conceptual analysis. Some philosophers claim that conceptual analysis is not so powerful as once believed.¹⁰⁰ If we look at the history of analytic philosophy, we see, however, that such analysis was usually focused on word usage. But this level of the ordinary word usage is no more than a surface of our conceptual schemes. Their deep structures lay at the level of our natural beliefs. Indeed, such beliefs can be interpreted as a kind of conceptual schemes which we impose on our experience. And I have just tried to show that it is possible to clarify the relations between such conceptual schemes with quite rigorous arguments. If this is correct, and if such clarifications can lead to solving some metaphysical problems, then we can hope that conceptual analysis has not said its last word yet.

Conclusion

In this paper I have argued that an analysis of one terminological anomaly in Hume's first *Enquiry* not only helps us to clarify his understanding of his own philosophical methodology but also gives us some clues to comprehending the evolution of his interpretation of the science of human nature and the order of his writing of the different parts of the first two books of his *Treatise*. Moreover, these considerations may throw some light on the nature of conceptual analysis and make us a bit more optimistic about prospects of conceptual philosophy.

⁹⁸ This argument can be easily supported by a few other common arguments contra epiphenomenalism.

⁹⁹ On natural supervenience, see David Chalmers, *The Conscious Mind*, 34–38.

¹⁰⁰ See, for a recent example, Timothy Williamson, *The Philosophy of Philosophy* (Oxford: Blackwell, 2008).

Vadim V. Vasilyev
Department of the History of World Philosophy
Moscow State University
vadim-v-vasilyev@rambler.ru

April 2012 version, revised on September 1, 2013