

philosophy, time, and consciousness: an opinionated introduction

Lajos Brons (mail@lajosbrons.net)

Department of Philosophy, Nihon University, Tokyo.

presented at

philosophy of mental time IV: time, experience, and consciousness

January 31st – February 1st, 2016, Osaka, Japan.

preface & abstract — This text was prepared for the introductory talk in a workshop intended to promote communication and cooperation between philosophers and experimental scientists on topics related to the experience and awareness of time. The workshop's two themes were: (1) the “specious present” and related ideas and research findings, and (2) mental time travel in animals (and humans). This talk/paper consists of four parts that are only loosely connected. The first (section 2) discusses the nature of philosophy and its relation to science and proposes that philosophy is a part of science characterized by a primarily deductive approach, and that philosophy's roles within science are critical and exploratory. The next three parts are more directly related to the workshop's themes but are also intended to illustrate those two roles. The second part (section 3) presents a brief grammatical analysis of the verb “to be conscious (of)” and applies that analysis to the “problem of the rock” (or the “generality problem”) for Higher Order theories of consciousness to conclude that that “problem” makes no sense. The third part (section 4) summarizes theories of temporal consciousness and points out a number of fundamental problems that most (if not all) of these theories share. The last part (section 5) is intended more to illustrate the exploratory role of philosophy in science (than its critical role) and draws together several lines of thought and research to show why a certain aspect of animal consciousness matters.

1.

According to Donald Davidson, communication does not require a shared language (or any other kind of conventions). This may very well be true in principle, but Davidson admitted later that, in practice, conventions make communication a lot easier and effective.¹ Probably something like this is true for interdisciplinary communication and cooperation as well: in principle cooperation may be possible without a shared language, but in practice, interdisciplinary communication and cooperation will be much more efficient and effective if the participants understand each other (at least to some degree) from the start.

While developing mutual understanding (at a sufficiently early stage in the cooperation) may already be a challenge, in case of cooperation between philosophy and science (or the other sciences, if philosophy is considered a science) there are some additional complications.

¹ Davidson argued against the necessity of (prior) conventions for communication in (1984), (1986), and (1992), but admitted in (1999) that his argument was “defective” because “it does not sufficiently stress the practical advantage that comes from a large degree of commonality among groups of speakers” (p. 73).

Firstly, it is rather unclear what philosophy is and what philosophers do – even to philosophers themselves. Secondly, the academic discipline of philosophy is highly fragmented, and many of its schools and traditions don't even understand each other. And thirdly, philosophical terminology routinely uses common words in (sometimes subtly) different ways, inviting misunderstanding.

This workshop aims to promote communication and cooperation between philosophers and experimental scientists on two specific topics related to the experience and awareness of time, but before we start discussing those topics, it may be useful to say something about the nature of philosophy and its relation with empirical science. That will, therefore, be the first topic I want to address in this talk (in section 2). After that, I will turn to the topics of this workshop, and will make a few remarks about consciousness (in section 3), temporal consciousness (in section 4), and animal consciousness (in section 5).

2. the nature of philosophy, and its relation with empirical science

In his *Confessions*, Saint Augustine (5th ct.) wrote: “What is time? If no one asks me, I know what it is. If someone asks me to explain it, I do not know.”² The quote is appropriate considering our workshop theme, of course, but that is not the reason why I quote Augustine. Rather, my reason is that the quote is true for “philosophy” as well as for “time”. “What is philosophy? If no one asks me, I know what it is. If someone asks me to explain it, I do not know.” Or perhaps I should refine my claim: the application of Augustine's quote to “philosophy” is true at least *for me*. I don't know what philosophy is. And that is a source of embarrassment, of course, given that I try to make a living teaching philosophy.

There are others, however, who are (or believe to be) less burdened by ignorance about the nature of philosophy and who have proposed various definitions. Unfortunately, no one has succeeded thus far in giving a definition that is not too broad, too narrow, both, and/or nonsensical. For any proposed definition there are – usually obvious – counterexamples (*i.e.* “things” that are excluded or included by the definition while they shouldn't be), and this is one of the reasons why I say that I don't know what philosophy is.

According to Plato, philosophy studies some part of reality that is beyond the reach of the empirical sciences – the world of Ideas – but that suggestion depends on the dubious metaphysical claim that there are such Ideas (and such a world). Moreover, it raises more questions than it answers. For example, if the world of Ideas lies beyond the reach of the empirical sciences, then that implies that it has no empirically detectable causal relations, but if that is the case, how can we get to know anything about it at all?

One of the most common views on the nature of philosophy is that it is some kind of source and residue of the sciences. Prominent defenders of this view include J.L. Austin and John Searle, who wrote that “as soon as we can revise and formulate a philosophical question to the point that we can find a systematic way to answer it, it ceases to be philosophical and becomes scientific” (1999: 2069). The problem with this view is that it doesn't really tell us

2 “Quid est ergo tempus? Si nemo ex me quærat, scio; si quærenti explicare velim nescio: ...” (Book 11, chapter XIV, 17)

what philosophy is, but only what it is not (*i.e.* the branches of science that split off from it), and that if we try to characterize the residue by means of a positive definition, that definition either includes religion and mythology, or excludes ethics and social and political philosophy.

Ludwig Wittgenstein famously claimed that “the philosopher’s treatment of a question is like the treatment of an illness” (1953: 255). Philosophy is therapy for our tendency to ask the wrong kinds of questions. However, while it certainly is the case that something like this is and has been a task of philosophy, it doesn’t seem to characterize the whole of it, and rather than as a definition of the actual and historical discipline of philosophy, Wittgenstein’s claim is better understood as a *proposal* for what it should be.

According to Richard Rorty, philosophy is not “a discipline which confronts permanent issues”, but rather, “is a cultural genre, a ‘voice in the conversation of mankind’ ..., which centers on one topic rather than another at some given time not by dialectical necessity but as a result of various things happening elsewhere in the conversation” (1979: 264). However, even if Western philosophy was *one* cultural genre in the past, it now consists of two major schools – analytic and continental philosophy – that hardly communicate with, or even understand each other, supplemented by a number of smaller schools such as Marxist, pragmatist, feminist, and Africana philosophy, and in addition to these, there are the various schools and traditions of non-Western philosophy: Buddhist philosophy, African philosophy, Confucianism, Mohism, Daoism, and so forth. Hence, philosophy is not *one* “cultural genre” or a “voice in the conversation of mankind”, but a cacophony of voices (or genres).

Many, but not all, attempts to define philosophy are attempts to draw a clear line between philosophy and the sciences, but typically these attempts result in either logic or ethics (and/or social and political philosophy, aesthetics, *etc.*) falling on the wrong side of the boundary line. Perhaps, the reason that all such attempts have failed is because they are based on the false assumption that there is such a line. Something like that, at least, is W.V.O. Quine’s position. According to Quine, philosophy and the sciences are continuous, and in the same way that there is no strict boundary between physics and chemistry, for example, there is no strict boundary between philosophy and the other sciences.

But even if that is the case, we distinguish physics from chemistry and sociology from economics. Even if these disciplines have vague boundaries, they appear to be identified by core topics and/or methods. What then, are the identifying topics and/or methods of philosophy?

Wilfrid Sellars famously claimed that philosophy aims “to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term” (1991: 1). This is – I think – a fairly accurate sketch of the topic of philosophy: everything, and how everything hangs together. But, obviously, this isn’t very helpful in identifying the discipline – in some sense, physics also studies everything and how everything hangs together.

Perhaps, what distinguishes (in a loose sense) philosophy from the other sciences (that study everything and how everything hangs together) is a difference in methodological focus. The other sciences are primarily inductive – their ultimate arbiter is (at least ideally) experiment and observation. Deduction plays an important role, as well, of course, but that role is always subordinate to the broader inductive project. What is deduced from previous inductions, is to be tested empirically next. In philosophy, on the other hand, the

methodological focus is deductive, rather than inductive. In philosophy deduction is not subordinate to induction, but it is the other way around: logic is the ultimate arbiter.

I have little doubt that this attempt to identify philosophy will also fail when put under sufficient pressure, but I believe that this is a useful way to think of philosophy and its relation to science because it has at least two important attractions. Firstly, it explains what philosophers do. And secondly, the ways in which philosophy can (and perhaps should) contribute to science as a whole or to (other) sciences individually almost “naturally” follow from the supposed deductive focus of philosophy.

What philosophers *do* is argue. Put two philosophers in a room together and they will almost instantly start to disagree and argue with each other (except, perhaps, if they talk about their holiday plans or last night’s dinner). That philosophers argue with each other all the time makes perfect sense, however, from the perspective of its deductive focus. If logic rather than experiment and observation is the ultimate arbiter of philosophical theories, then the only way to “test” them is to let others try to find logical flaws (after doing that oneself, of course). Argument is how philosophical theories are tested. Argument is what leads to refinement, amendment, and/or replacement of philosophical theories. (Hence, it is no coincidence that philosophical progress stalled where and whenever disagreement and argument were covered up and/or suppressed.)

The deductive focus of philosophy suggests two ways in which philosophy can contribute to science. The first of those is primarily critical; the second is more exploratory. Philosophy’s critical role is closely related to the previous point about argument. Analyzing arguments to detect flaws is as important in empirical science as it is for philosophy, and – at least in theory – philosophers excel at that. A philosophical argument fails if a term is used with (even subtly) different meanings in different parts of the argument, for example, but similar problems occur in empirical science. There may be subtle differences between what is measured or tested and what was hypothesized, or some empirical construct may deviate in slight but important ways from earlier research and/or theories based thereon. Empirical scientists rarely stop to ponder questions like “What *exactly* are we talking about?” and “Are and were we all talking about the same thing?”, but philosophers do that all the time, and they can do so in service of empirical science. Hence, the first kind of possible contribution of philosophy to the other sciences is the analysis of concepts and inferences.

A deductive argument takes its premises for granted to focus attention on what follows from them. The second role of philosophy in interdisciplinary cooperation – or perhaps even within science in general – departs from this principle to explore “to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term” (as Sellars phrased it – see above). In its exploratory role, philosophy takes the results of empirical science – more or less – for granted, and combines them with premises derived from other sources to see what follows from them. These “other sources” include other sciences, but also normative ideas and/or “self-evident truths” suggested by philosophers, as well as more ordinary (individual and social) human concerns. Consequently, in its exploratory role, philosophy can build bridges – not just between the sciences, but also (and much more importantly) to the rest of society by trying to answer the questions how and why certain scientific findings (broadly understood) matter.

What I am proposing – in summary of the foregoing – is that philosophy as a part of science is primarily concerned with two kinds of questions, roughly: “What exactly are we talking about?” and “What does it matter?”.

I said before about Wittgenstein’s conception of philosophy as therapy that it is better understood as a proposal for what philosophy should be (according to Wittgenstein) than as a definition or description of the actual and historical discipline of philosophy, and probably the same applies to my sketch of the discipline and its relation with science. Actual (rather than ideal) academic philosophy tends to be much narrower in scope than the second role requires. For various historical and institutional reasons philosophy – like all other academic disciplines – has fragmented into a collection of narrow specializations. However, the difference in methodological focus makes specialization much more problematic in philosophy than in the other sciences. To see why this would be the case, consider the following two claims (which may be true or false – that’s besides the point):

[AF] Great apes can imagine the future.

[OA] Orangutans are great apes.

The inductive focus of the empirical sciences means that their primary concern is with the truth or falsity of these claims. The deductive focus of philosophy, on the other hand, means that its primary concern is what follows from them, in case of this example:

[OF] Orangutans can imagine the future.

As suggested before, this doesn’t mean that empirical scientists are not interested in [OF] or that philosophers are *only* interested in [OF]. What it does mean is, firstly, that for empirical scientists [OF] is relevant mostly to judge the plausibility of [AF] or the classification of orangutans according to [OA]; and secondly, that the assessment of the truth or falsehood of the two premises [AF] and [OA] is part of a philosopher’s “job” only if they lie within the scope of her branch of philosophy. (Of course, the difference between the two approaches is not as sharp as I’m sketching it here.)

By implication of the latter, the narrower the scope of some deductive research field, the more it excludes, and therefore, the more it takes for granted. Consequently, in a deductive science (*i.e.* philosophy) there is a risk that specialization deteriorates critical reflection on basic assumptions or even leads to dogmatism. Certainly it seems to have had this kind of effects in academic philosophy, as there is a large body of beliefs that are more or less uncritically accepted by the majority of philosophers. (This is true for all schools of philosophy; they just differ in the content of that body of beliefs.)

Within this body of largely shared beliefs there are several that are important in the present context. At the top of that list would be a number of beliefs related to the typically high confidence of philosophers in their own cognitive abilities. Philosophers typically assume that their own minds are largely transparent to themselves, and thus that they are authorities with regards to their own mental contents and thought processes. (This is a peculiar assumption as it seems to imply that these philosophers believe that they can think and observe and think about that thinking at the same time, which is implausible, to say the least.) And most philosophers trust their own intuitions (as sources of knowledge), but not the intuitions of others. Another widespread belief is that mental states like beliefs, thoughts, and

desires are representational; that is, that such mental states are involve representations of things out there in the world. Awareness of assumptions like these (and quite a few others) matters in cooperation between philosophers and scientists because they may not be shared by (all of) the latter (in all cases), and they may lead to misunderstanding. (The problem, of course, is that it is in the nature of such background assumptions that the person making those assumptions is rarely aware of them, and even less of the need to make them explicit.)

Before moving on to topics related to consciousness and time, let me briefly recapitulate the three main points I tried to make in the first part of this talk. Firstly, philosophy differs from the other sciences in its methodological focus, which is primarily deductive, while the other sciences are primarily inductive. Secondly, from this deductive focus two roles for philosophy within science as a whole or within interdisciplinary cooperation follow: a critical role and an exploratory role. In the first role, philosophy's "job" is to check terminological and logical consistency; in the second role, it combines research findings as premises to see what follows from them and why they matter. Thirdly, in practice both of these roles may be hampered by narrow specialization and by uncritically accepted assumptions. Consequently, whether philosophers can actually play a useful role depends very much on their open-mindedness.

Of the following three parts of this talk, the first two (on conscious in general and temporal conscious in particular) illustrate the critical role of philosophy, while the last (on mental time travel in animals) illustrates its exploratory role. I'm not sure whether they are *good* illustrations, but I hope that at least they're useful *as illustrations*.

3. concepts of consciousness

"Consciousness" is a notoriously ambiguous concept. In the literature various specific notions of consciousness are distinguished: creature consciousness, state consciousness, phenomenal consciousness, and so forth. Within creature consciousness, usually a distinction is made between transitive and intransitive variants. In linguistics, transitivity is normally a property of verbs, and for that reason, the latter distinction raises the question whether it would be helpful to think of consciousness as a verb. I think the answer to that question is "yes". In the following I hope to somewhat disentangle the conceptual mess by analyzing the verb "to be conscious (of)".

In English, there is no standard relation between the two variants in transitive/intransitive verb pairs, and even the identification of such pairs can sometimes be disputed. While "to eat" and "to fly" clearly have both transitive and intransitive uses, and thus the two variants in the pair are homonyms, pairing "to fall" and "to drop" may be less obvious. In Japanese, on the other hand, the latter are as obviously related as the two forms of "to eat" in English, as both are written with the same Chinese character 落 - 落ちる *ochiru* means "to fall" and 落とす *otosu* means "to drop". I'll ignore this complication here, however, and will just assume that "to fall" and "to drop" constitute a transitive/intransitive verb pair similar to the two variants (or uses) of "to eat" and "to fly".

An obvious difference between *fall/drop* and *eat/eat* or *fly/fly* is that in cases when alternative forms are used to describe the same situation, then the latter two have the same

gramatical subject in either variant while this is not the case for the former. That is, if x eats y then x eats, but if x drops y then y falls. Confusingly, in case of “to be conscious (of)” intransitive variants can be both about x and y : if x is conscious of y , then it can be said that x is conscious, but also – in case of beliefs and decisions, for example – that y is conscious. Something like the first is often called “creature consciousness” and the second “state consciousness”. Hence, if x is conscious of y , then x is creature conscious, and y is a conscious state (or belief, or decision, *etc.*) (of x) and thus state conscious. However, both creature consciousness and state consciousness may have further requirements, so I propose two more neutral terms: *agent consciousness* and *patient consciousness*, defined as follows:

[D1] If x is conscious of y , then
 x is **agent conscious** (of y), and
 y is **patient conscious** (to x).

As mentioned, creature consciousness may have further requirements, but there is substantial disagreement about what these requirements could be. To say that a creature is conscious may mean that it is awake, that it is sentient, that it is self-conscious, that it is in a state such that there is something it is like to be in that state, and so forth.

Eating implies eating something and the other way around, but the same isn't true for flying. One cannot eat without eating something, but a bird flies without flying something (and the same is true for an airplane passenger if it is the pilot who flies the plane and the passengers fly *in* the plane). Perhaps, there are other examples where the conditionality works the other way around.

It seems that there is at least a minimal sense of creature consciousness that is analogous to the case of *eat/eat*: to be (creature) conscious means to be conscious of something (which can be something external, but which can also be a thought or desire). If a creature is conscious of something it is conscious, and the other way around, it makes no sense to say that some creature is (or even could be) conscious without being conscious of anything. This minimal sense of creature consciousness coincides with conceptions of consciousness as being awake and being sentient, but not with more specific kinds of agent consciousness such as being self-conscious or being in a state such that there is something it is like to be in that state. (The first of these is an odd requirement for general intransitive consciousness, however, as it mentions a specific object (or patient) of consciousness: the self, whatever that may be.) In either conception of creature consciousness (and also in further conceptions I know of), creature consciousness is agent consciousness plus some further requirement. Hence:

[D2] x is **creature conscious** if and only if x is agent conscious, and x satisfies Ψ

in which Ψ is the further condition (which may be empty). From [D1] then follows that:

[D2'] x is **creature conscious** if and only if there is some y such that x is conscious of y , and x satisfies Ψ

It should be noted, however, that x 's satisfaction of Ψ may also involve y , which is a reason to prefer [D2'] over [D2].

Similarly, state consciousness *appears* to be a kind of patient consciousness. Some mental state is a conscious state – and thus state conscious – if there is some creature in that state, and that creature is aware of the state, or if that creature’s being in that state involves something like qualia or phenomenal properties, or if there is something it is like to be in that state, or if the information carried by that state is available to the creature, and so forth. However, while a mental state cannot exist without there being a creature in (or having, or experiencing) that state, there is considerable disagreement about what exactly is required for a mental state to be conscious, and particularly whether a conscious mental state requires that creature to be (in some sense) conscious of that state. According to some Higher Order theories of consciousness (something like) this is a requirement indeed – that is, a state is a conscious state if and only if the creature in that state is conscious of that state – but there are other theories with a different view on the matter, and different Higher Order theories differ from each other in several, important respects, including what it means to be conscious of something.

However, if some things (namely mental states) can be said to be conscious while being neither agent conscious nor patient conscious, then there must be a third way for a thing to be conscious, and it is doubtful that that suggestion even makes sense. (At least, I can’t make sense of it.) If something is conscious then either it is conscious of something else, or something else is conscious of it. And consequently, (because it isn’t creature consciousness) state consciousness must be a kind of patient consciousness:

[D3] x is **state conscious** if and only if x is patient conscious, and x satisfies Φ

in which Φ is the further condition (which, like Ψ , may be empty). And as above, from [D1] then follows that:

[D3'] x is **state conscious** if and only if there is some y such that y is conscious of x , and x satisfies Φ

in which x ’s satisfaction of Φ may also involve y (analogous to [D2'] above).

I find the notion of state consciousness rather confusing, however. The clarity of debates about consciousness would be greatly enhanced if we could get rid of “state consciousness”; that is, if instead of saying that some state is conscious we would say that some being is conscious of that state. For that reason, I would prefer the following re-phrasal:

[D3''] x is a **conscious state** (and thus conscious *in that sense*) if and only if there is some y such that y is conscious of x , and x satisfies Φ .

If the above is right, then the most basic notion of consciousness is the transitive verb “to be conscious of”. The other notions of consciousness can be defined in terms of that notion plus the additional Ψ and Φ . This raises questions about the nature of those, of course: What could Ψ and Φ be? and what does it mean to be conscious of something? I think that Ψ and Φ are empty – that is, that creature conscious is the same thing as agent consciousness and that state consciousness is the same thing as patient consciousness – and that “to be conscious of x ” means something like perceiving x and (simultaneously or subsequently) believing that one is perceiving x (which would be a variant of Higher Order theory), but I will not defend these

views here. Instead, I want to illustrate my attempt to disentangle the conceptual mess of “consciousness” by applying it to the so-called “problem of the rock”.

The problem of the rock (or the generality problem) is an objection against Higher Order theories of conscious first suggested by Alvin Goldman (1993). Goldman wrote:

A rock does not become conscious when someone has a belief about it. Why should a first-order psychological state become conscious simply by having a belief about it? (p. 366)

If I understand this objection correctly, then either it is nonsensical or it depends on an obvious equivocation fallacy. Since it is unlikely that Goldman’s argument is so obviously flawed, I probably don’t understand it correctly, but fortunately that doesn’t make it less suitable for the illustrative purpose I have in mind.

I’m not sure what exactly Goldman means with “become conscious”, but the most charitable interpretation of the second sentence of the above quote seems to be an interrogative version of the claim that

[HO] a mental state x is conscious if and only if someone has a belief about x

which summarizes Goldman’s interpretation of some version of Higher Order theory. (And I’ll assume that interpretation to be correct.) He then argues that, lacking a good argument to the contrary,³ we should be able to generalize [HO] into:

[GP] x is conscious if and only if someone has a belief about x

from which it follows that:

[RP] a rock x is conscious if and only if someone has a belief about x .

However, [RP] is false because rocks aren’t conscious, and therefore, [HO] is false. Or at least that is what the objection seems to suggest that the enthymematic conclusion should be.

To say that something is conscious is to say that it is either creature conscious or state conscious. Mental states cannot be creature conscious (because they aren’t creatures), so “conscious” in [HO] must mean “state conscious”. If we ignore Φ (see [D3] above) and identify state consciousness with patient consciousness then the left-hand side of [HO] translates into “there is some y that is conscious of mental state x ”. Since “someone” in the right-hand side refers to that same y , the quantifier that bounds y should have the whole sentence in its scope, and thus [HO] translates into:

[HO'] for any creature y and a mental state x , y is conscious of x if and only if y has a belief about x .

which can be generalized into an analogous version of [GP] and applied to an analogous version of [RP] without causing serious problems.

[RP'] for any creature y and a rock x , y is conscious of x if and only if y has a belief about x .

³ Attempts to refute the “rock objection” generally focus on providing such an argument. That is, they try to show why the generalization from [HO] to [GP] is invalid. If my analysis is correct, this generalization is harmless, however.

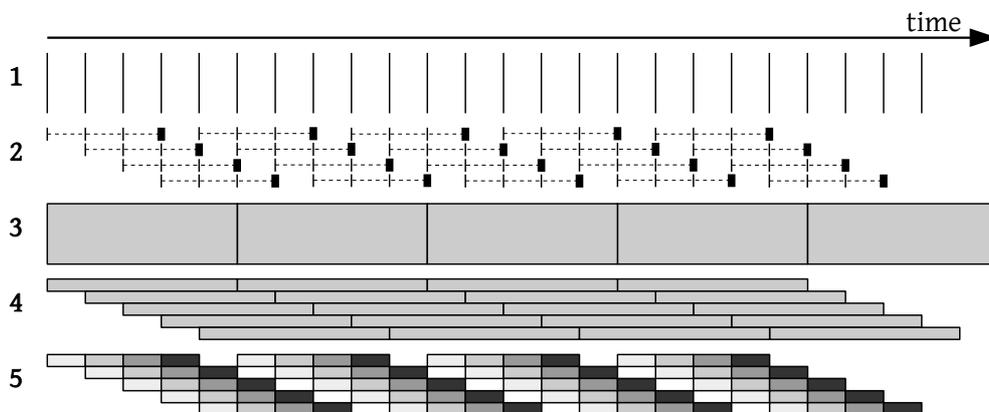
The problem with the so called “problem of the rock” as I understand it, is that it needs consciousness in [RP] to be creature consciousness for the problem to arise, but consciousness in [HO] to be state consciousness to (more or less) accurately characterize a Higher Order theory. [HO] cannot be about creature consciousness (as mental states aren’t creatures), and [RP] is not problematic if is about state consciousness – all it claims then is that someone is conscious of a rock if one has a belief about that rock. Consequently, there is no “problem of the rock”.

4. temporal consciousness: experiencing the present

Among the things and phenomena we are supposed to be conscious of, time is one of the most puzzling. When or if we are conscious of time, what exactly are we conscious of? And what do we perceive pre-consciously that feeds into that conscious experience of time?

Philosophical theories of temporal consciousness give very different answers to these questions. Figure 1 graphically summarizes the main kinds of theories, which will be briefly introduced in the following. It should be noted, however, that there is no standard terminology, and thus that terms used here deviate from those employed in some other overviews (and by different adherents to variants of the kinds of theories mentioned).

figure 1: theories of temporal consciousness



The most basic division between theories of temporal consciousness is that between extensional and non-extensional theories. The first assume that we are always (perhaps necessarily) conscious of a short, but extended period of time; the second assume that we are always (perhaps necessarily) conscious of moments without duration. Moments without duration are drawn as vertical lines in the figure (but very short ones in case of 2), and periods as blocks.

Non-extensional theories come in two kinds (1 and 2 in figure 1). According to cinematic theories (1), we are conscious of a series of moments in rapid succession (as the rapidly succeeding images on a movie screen), and the connection between them is provided by our minds. Hence, we do not really see things move, but we see things at different locations, and our minds create the perception of movement.

In case of momentary retentional theories (2), we are not just conscious of the present moment, but – in some sense – also of some of the preceding moments. That is, at any time, we are directly conscious of the present moment (short thick line in the figure) and indirectly – through retention – of the preceding moments (short thin lines, connected by a dotted line). Our conscious temporal experience is constructed out of these momentary experiences, and consequently, what we are conscious of at any time overlaps with what we are or were conscious of at adjacent points in time.

According to extensional theories, we are conscious of short periods of time, often called the “specious present” after William James (1890). The positing of a specious present is usually defended by the claim that we are conscious of movement and of multiple notes of a melody in a single act of consciousness. Extensional theories come in two main kinds (3 and 4 in figure 1). Sequential theories (3) assume that our temporal consciousness consists of a series of discrete periods. First we are conscious of one block in the figure, then of the next, and so forth. Dependent on how quickly the musical fragment shown in figure 2 is played, we would be conscious of the first three notes first, for example, then of the next three, and so forth. Consequently, we would be conscious of c-a-c (the first three notes) and e-e-g (the next three notes), but never of any note sequence including c-e (the third and fourth notes, which are the last and first notes of temporally adjacent conscious experiences).

figure 2: the first two measures of Гарын Арван Хуруу (Mongolian traditional)



Overlap theories (4) find this latter consequence implausible and suggest that we aren't conscious of discrete, sequential “blocks” of time, but of overlapping “blocks”. Hence, our consciousness of c-a-c is not followed by the next three notes (*i.e.* e-e-g), but by the overlapping sequence a-c-e, and so forth. (See figure 1, number 4, for a graphical representation.) Such overlap theories have further variants: the overlapping sequences may be continuously replaced like the (apparent!) view from the window of a moving train, or they may be punctuated (as in figure 1). In the first case there would be sequences in between c-a-c and a-c-e, but not in the second case.

In addition to these four kinds of theories of temporal consciousness, a fifth kind could be added, although I'm not aware of anyone defending it. This fifth kind would combine aspects of momentary retentional theories (2) with extensional overlap theories (4), and could be called extensional retentional theories (5 in figure 1). As in momentary retentional theories we combine the present with the recent past in our consciousness, but the atomic experiences that make up temporal consciousness have (short) duration as in extensional theories. According to such theories, our temporal consciousness would be like a sequence of blocks to which new blocks (the now) are added at the one end, while older blocks disappear at the other. In terms of the example of figure 2, we would be directly conscious of the present note – for example, the fourth note (e) – and indirectly, through retention, of the notes directly preceding it (*e.g.* a-c).

All of these theories have their problems and attractions. I will not summarize all of those here, but will instead focus on a few points about theories (and theorizing) of temporal consciousness in general.

Probably the most fundamental problem was already mentioned before, albeit only in passing: most theories of temporal consciousness assume transparency of our own minds to ourselves. Extensional theories and the notion of the “specious present” depend on introspective reports, but it is by no means clear that this is a reliable source of information. It is at least in principle possible that much of our temporal consciousness is constructed or even illusory. Introspection may be able to tell us what we think that we are conscious of, but that is not necessarily the same as what we really are (or were!) conscious of. The mind in general, and temporal conscious in particular, may be much more opaque than commonly assumed (*e.g.* Carruthers 2011).

A closely related problem is the relative lack of attention for the *processing* of sensory signals. There may be many steps and/or phases of processing by our sensory systems and brains before something reaches conscious experience. And it is not *prima facie* impossible that different steps or phases in that process resemble different theories of temporal consciousness. Of course, the notion of temporal *consciousness* implies that these theories refer to the last, conscious stage in that process only, but earlier stages may restrict that last stage in important ways, and it is, moreover, not completely clear whether all of the theories mentioned are just about *conscious* experience indeed. For example, if the cinematic theory is right, we are *not* conscious of the rapid succession of atomic experiences, but only of the stories our mind concoct to connect those in a coherent fashion. Hence, the cinematic theory is as much a theory of at least one pre-conscious stage as of the conscious experience.

The third – and final – point I want to make is – more or less – an extension of the previous point. Arguments for the “specious present” and against durationless conscious experience as assumed by non-extensional theories typically appeal to our conscious experience of movement as movement (rather than as a series of different locations) and musical sequences, but there is a much more fundamental objection against durationless experience, namely that it is physically impossible. Sound (we hear), light (we see), and temperature (we feel) all require time. There is no sound at a point of time, nor light (and certainly not light of a particular color), or temperature. Sound, light, temperature, and many of the other “things” we sense require duration. A sensation of any of these *must* have some duration. Hence, our bare sensing cannot be durationless, and if that is the case, then neither can there be durationless atoms of our conscious temporal experience.

These bare sensings (I hesitate to call them experiences, as that term seems to imply consciousness, and bare sensings are not conscious) can have a very short duration, and their lengths may very well differ between modalities. Furthermore, these temporally extended bare sensings must be sequential rather than overlapping because our neural infrastructure just doesn’t allow for the redundancy in signal transmission implied by overlap models. This means that *on the level of bare sensing*, something like the sequential theories (3 in figure 1) must be true. Recent research by neuroscientists suggests that our visual sensory system indeed produces sequential, discrete samples with a length of 50 to 100 ms (*i.e.* at a sampling

rate between 10 to 20 Hz; see for example Dubois & VanRullen 2011; Kitazawa 2013).⁴ 50 to 100 ms is much shorter than the supposed length of conscious blocks in extensional theories of temporal consciousness. The supposed duration of the “specious present” is supposed to be around 1 s; that is, 10 to 20 times longer. That doesn’t mean, however, that this neurological evidence refutes extensional theories – we are (probably) not conscious of the discrete samples themselves (in normal circumstances, at least), which means that these are somehow further processed and/or combined. How exactly? would be the next question.

Answering that question is no easy task, however. And there is (at least) one other puzzling fact that somehow needs to be taken into account: patients suffering from motion blindness have conscious visual experience that is sequential and discrete at a significantly lower rate (around 6 Hz; see Hess, Baker, & Zihl 1989). Could this mean that the sensory samples (of 50 to 100 ms) are first grouped into longer, semi-conscious units of 150 to 200 ms, which then are collected at a next level in longer specious-present-like “blocks”? Perhaps. Perhaps not. Obviously, I’m not qualified to answer the question, but my point in asking it is to illustrate a complication: the process between bare sensing (or raw sensory input) and conscious experience may involve several steps and may be far less straightforward than most theories of temporal consciousness seem to assume.

5. mental time travel and self-consciousness in animals

While the previous two parts of this talk were illustrations of the critical role of philosophy more than its exploratory role, in this last part I want to illustrate that exploratory role by contemplating why it would matter whether (some) animals are capable of *mental time travel*, the ability to mentally project oneself backward in time to relive some past experience, or forward in time to “prelive” some expected or desired future experience (e.g. Suddendorf & Corballis 1997).

In his *Introduction to Principles of Morals and Legislation* Jeremy Bentham wrote:

The day *may* come, when the rest of the animal creation may acquire those rights which never could have been withholden from them but by the hand of tyranny. The French have already discovered that the blackness of the skin is no reason why a human being should be abandoned without redress to the caprice of a tormentor. It may come one day to be recognized, that the number of the legs, the villosity of the skin, or the termination of the *os sacrum*, are reasons equally insufficient for abandoning a sensitive being to the same fate. What else is it that should trace the insuperable line? Is it the faculty of reason, or, perhaps, the faculty of discourse? But a full-grown horse or dog is beyond comparison a more rational, as well as a more conversable animal, than an infant of a day, or a week, or even a month, old. But suppose the case were otherwise, what would it avail? the question is not, Can they *reason*? nor, Can they *talk*? but, Can they *suffer*? (1789:17.1.IV.n.)

⁴ A quick probe of some of the literature on the topic teaches that there are various other closely related sampling rates and that sampling rates may differ from person to person and from occasion to occasion. Hence, the sampling rate of visual sensory input is probably not a settled matter yet, and may be (even) more complicated than suggested here.

The question whether animals can suffer is really two questions, however, and only one of those can be answered by empirical research. The other one is the question what it means to suffer. Bentham founded Utilitarianism, the moral theory that identifies the good with the greatest happiness of the greatest number. Happiness and pleasure is bad, according to that theory, and unhappiness and pain are bad. In its further development the good was (and is) often identified with “utility” and the bad with “dis-utility”, but the focus tends to be on utility (often reduced in practice to material well-being). Conceptualization of the negative side – that which Utilitarianism aims to avoid – is less well-developed, but if it is the ability to suffer that makes some creature morally considerable, then we need to know what that means.

Outside Buddhism,⁵ the concept of suffering is most central in medicine and nursing, which have the alleviation of individual suffering as one of their traditional goals. Hence, those fields would appear to be a good place to look for informed contemplation on the nature of suffering.

The dominant understanding of suffering in the medical and nursing literature distinguishes suffering from pain and is based on Eric Castell’s definition of suffering “as the state of severe distress associated with events that threaten the intactness of a person” (1982:640; 1991:33),⁶ wherein the term “person” refers – more or less – to someone’s self-concept. Most important critical responses to Castell’s definition are those by Steven Edwards and (2003) and Franco Carnevale (2009). The first stresses that suffering is something felt, must have a significant duration, and must have “a fairly central place in the mental life of the subject” (p. 65). The second argues convincingly that suffering is an emotion (while pain is a localized sensation). Although Edwards and Carnevale succeed in showing some flaws in Castell’s definition, neither rejects it completely, and for that reason, I’m inclined to see their critique as proposals for amendment more than rejections. What stands is that suffering is subjective (*i.e.* can be judged by the sufferer only), and that suffering is associated with threats to the self and/or the self-concept.

In the case study Castell uses to introduce and illustrate his argument he writes that the patient “feared the future” (1982:639). Indeed, suffering (either before or after Edwards’ and Carnevale’s amendments) is a kind of fear more than a kind of pain (and fear indeed is an emotion). “Suffering occurs when an impending destruction of the person is perceived” (1982:640; 1991:33). Castell’s conception of a “person” is a self-concept. It is the sum-total of the “things” (in the broadest possible sense) that matter to someone: life, body, relations, roles, self-images, and so forth. The perception of the *impending* destruction (of a part) thereof is a projection of the self into the future. Suffering, therefore, is a kind of fear associated with self-projection into an undesirable future. And since projection of oneself into the future is a kind of mental time travel, this means that the ability to suffer depends on the ability of mental time travel. By implication, the ability of (some) animals to suffer (in addition to being in pain) depends on their ability of mental time travel.

5 It is controversial whether “suffering” is a good translation of the Buddhist core concept of *dukkha*, but it remains the most common translation.

6 Although this definition is widely accepted in theoretical work, actual medical practice is more often informed by an identification of suffering with observable indicators of pain, or by medical workers’ preconceptions of what circumstances would count as resulting in suffering.

If the foregoing is right, then the ability of mental time travel is a *necessary* condition for the ability to suffer, but it may not be a *sufficient* condition. Suffering is conceptualized as a kind of fear (or a fear-like emotion), and thus also requires the ability to feel such an emotion. It seems plausible, however, that these two abilities – mental time travel and fear – are jointly sufficient for the ability to suffer. There appear to be many kinds of animals that are capable of fear, and consequently, if we take the moral significance of suffering as conceptualized here for granted, then, if (some of) those animals have the ability of mental time travel, they can suffer and are, therefore, morally considerable.

By implication, answering the question whether (some) animals have the ability of mental time travel matters a great deal to how we relate to, and how we should treat (those) animals. Answering that question is not easy, however, not just because of experimental difficulties, but mostly because there is a further conceptual issue: it is not exactly clear what it means to project oneself into the future (or the past), and especially how sophisticated and/or articulate the projected self-concept needs to be.

Cameron Buckner (2013) has pointed out that philosophers and scientists have a tendency to assume a lot about human cognitive abilities (often based on introspection) while demanding ever stronger evidence for the capabilities of animals. Our minds may not be as transparent (to ourselves) as many of us believe, however (*e.g.* Carruthers 2011), and our beliefs and concepts including our self-concepts may not be as sophisticated as we'd like to think either. Demanding the conceptual sophistication *we think we have* from animals before considering to attribute them the ability of mental time travel is an example of the kind of “anthropofabulation” Buckner argues against.

Perhaps, the most minimal version of having a self-concept – and this minimal version may very well be sufficient – would be a (non-propositional) belief (-like state) in the existence of what Mark Johnston calls “the one at the center of this arena of presence and action” (2010: 139). The question is: What would count as evidence for such a belief (-like state)?

I have no answer to that question. My aim here was merely to draw together several lines of thought and research to show how and why some specific aspect of animal consciousness might matter.⁷

references

- Bentham, Jeremy (1789), *Introduction to Principles of Morals and Legislation*.
- Buckner, Cameron (2013) “Morgan’s Canon, Meet Hume’s Dictum: Avoiding Anthropofabulation in Cross-Species Comparisons”, *Biology & Philosophy* 28.5: 853-71.
- Carnevale, Franco (2009), “A Conceptual and Moral Analysis of Suffering”, *Nursing Ethics* 16.2: 173-83.
- Carruthers, Peter (2011), *The Opacity of Mind: An Integrative Theory of Self-Knowledge* (Oxford: OUP).

7 “Might”, because there are various gaps in the argument(s) presented here that (would) need mending as well.

- Castell, Eric (1982), “The Nature of Suffering and the Goals of Medicine”, *The New England Journal of Medicine* 306.11: 639-45.
- Castell, Eric (1991), *The Nature of Suffering and the Goals of Medicine* (Oxford: OUP).
- Davidson, Donald (1984), “Communication and Convention”, in (1984), *Inquiries into Truth and Interpretation* (Oxford: Clarendon), pp. 265-80.
- Davidson, Donald (1986), “A Nice Derangement of Epitaphs”, in (2005), *Truth, Language, and History* (Oxford: OUP), pp. 89-107.
- Davidson, Donald (1992), “The Second Person”, in (2001), *Subjective, Intersubjective, Objective* (Oxford: OUP), pp. 107-22.
- Davidson, Donald (1999), “Reply to Peter Pagin”, in: U.M. Żegleń (ed.), *Donald Davidson: Truth, Meaning and Knowledge* (Milton Park: Routledge), pp. 72-74.
- Dubois, Julien & Rufin VanRullen (2011), “Visual Trails: Do the Doors of Perception Open Periodically?”, *PLoS Biology* 9.5.
- Edwards, Steven (2003), “Three Concepts of Suffering”, *Medicine, Health Care and Philosophy* 6: 59-66.
- Goldman, Alvin (1993), “Consciousness, Folk Psychology, and Cognitive Science”, *Consciousness and Cognition* 2: 364-82.
- Hess, R.H., C.L. Baker, & J. Zihl (1989) “The ‘Motion-Blind’ Patient: Low-Level Spatial and Temporal Filters”, *The Journal of Neuroscience* 9.5: 1628-40.
- James, William (1890), *The Principles of Psychology*.
- Johnston, Mark (2010), *Surviving Death* (Princeton: Princeton University Press).
- Kitazawa Shigeru (北澤 茂) (2013), 「こころの「現在」の科学—時間の錯覚からわかること」, *BRAIN and NERVE—神経研究の進歩* 65.8: 911-21.
- Rorty, Richard (1979), *Philosophy and the Mirror of Nature* (Princeton: Princeton University Press).
- Searle, John (1999), “The Future of Philosophy”, *Philosophical Transactions of the Royal Society of London B* 354: 2069–80.
- Sellars, Wilfrid (1991), *Science, Perception and Reality* (Atascadero: Ridgeview).
- Suddendorf, Thomas & Michael Corballis (1997), “Mental Time Travel and the Evolution of the Human Mind”, *Genetic, Social & General Psychology Monographs* 123.2: 133-167.
- Wittgenstein, Ludwig (1953), *Philosophische Untersuchungen* [*Philosophical Investigations*].