

## TRANSCENDENCE OF THE EGO (THE NON-EXISTENT KNIGHT)

*Bas C. van Fraassen*

### *Abstract*

I exist, but I am not a thing among things; X exists if and only if there is something such that it = X. This is consistent, and it is a view that can be supported. Calvino's novel *The Non-Existent Knight* can be read so as to illustrate this view. But what is my relation to the things there are if I am not identical with any of them – things such as my arms, my garden, the city I live in? I name this the Gurduloo problem, after the Knight's page. This relation must be one that admits of degrees; I suggest that we say that I *manifest* myself through the things thus associated with me. Several pseudo-problems, pertaining to volitional action, supervenience, observability, and the emergence of consciousness, dissolve upon inspection.<sup>1</sup>

*I am as inseparable from the world as light, and yet exiled as light is, gliding over the surface of stones and water, never gripped nor held.*<sup>2</sup>

I exist, but I am not a thing among things. I am neither a physical object nor a mental substance or abstract entity, nor a compound thereof. I am not animal, mineral, or vegetable. Nor am I a thing constituted by or composed of things of that sort taken together. I am not some piece of furniture of the universe.

Certainly I have a body and I have thoughts as well as feelings. I have a spatial location as well as a place, however modest, in history. But I am not to be identified with any of this.<sup>3</sup> I am in this world, but not of this world. I am not a thing among things.

<sup>1</sup> I want to thank Galen Strawson and Béatrice Longuenesse for encouraging me to continue despite my obvious ignorance of the subject.

<sup>2</sup> 'Aussi inséparable du monde que la lumière et pourtant exilé, comme la lumière, glissant à la surface des pierres et de l'eau, sans que rien, jamais, ne m'accroche ou ne m'ensable.' Jean-Paul Sartre, *Le Sursis*. Paris: Gallimard 1945. From the Chapter 'Mardi 27 Septembre'.

<sup>3</sup> Ingmar Persson's 'Self-Doubt: Why We Are Not Identical To Things Of Any Kind' signals agreement in its title, though in ways that I still need to explore.

### Logic: I exist

The view just expressed is logically coherent. Despite my leading quotation from Sartre, my own usage of 'exists' is not that of the Existentialists. Accepting the ground rules of Quine's 'On what there is', I equate '. . . exists' with 'There is such a thing as . . .'. Pegasus exists if and only if there is such a thing as Pegasus.

How can I say then that I exist, hence that there is such a thing as me, but also I am not a thing? There is an ambiguity in the English language. In such 'quantifier' locutions as 'something' or 'there is such a thing as' or 'everything', the word 'thing' does not occur with any substantive meaning, but is a sort of pronominal device. In elementary logic we paraphrase 'Something is . . .' as 'There is  $x$  such that  $x$  . . .' For example, 'Everything beautiful is good' we render as '(All  $x$ ) (if  $x$  is beautiful then  $x$  is good)'. The word 'thing' has disappeared. The example's logical structure is displayed in the paraphrase 'Everything is such that if it is beautiful then it is good'. Two of the three occurrences of ' $x$ ' correspond therefore to the relative pronoun 'it'. But the first occurrence of ' $x$ ', corresponding to the 'thing' part of 'Everything', does not play a different role from the others.<sup>4</sup> In venerable terminology, its use is not categorematic but syncategorematic.<sup>5</sup>

### 'Thing' and 'object'

When I say that I am not a thing, not an object, I mean this in the more substantive senses of 'thing' and 'object' in which they are genuine common nouns.

*I am not an object.*

*The self is not an object.*

*The word 'I' does not, on any proper occasion of use, refer to an object.*

<sup>4</sup> Clearer in combinatory logic: a universally quantified sentence says that a certain predicate has universal application.

<sup>5</sup> In model theory the domain of a quantifier is a set; but nothing implies that it is a set of objects. Note moreover that since each model has a set as its domain, and there is no set that contains everything there is (on pain of contradiction), it follows that the meaning of 'for all  $x$ , . . .' as provided by a model is never 'for everything there is'. If in fact there is a set of all *objects*, or of all *things*, or of all *substances*, it also follows that 'everything' does not mean 'all objects', 'all things', or 'all substances'.

Physical events and processes as well as persistent, enduring materials count as objects. So do mental substances and mathematical objects, witches and demons, quantum fields and quarks, space-time *pace* the substantialists, other possible worlds on at least one conception thereof – always assuming that there are any. Is there anything at all besides objects? Yes: Selves, among other things (once again using ‘things’ syncategorematically).

My topic is the Self; but I equate ‘What is the Self?’ with the question ‘What am I?’. I take the former to be but a quasi-impersonal way of posing a mystery that can in the last recourse only be expressed in first-person, indexical language.<sup>6</sup> So here is my first assumption: I refer to my self when I say ‘I’, and you refer to your self when you say ‘I’. For I take it that I refer to my self when and only when I refer to myself.

That my self is the referent of my word ‘I’ is not a necessary assumption. Galen Strawson does not accept this, and says from the outset that the self is not the whole person.<sup>7</sup> In this he is in accord with much of modern philosophy, and I am not. But I take it that on all accounts, the question ‘What am I?’ – to which I limit myself – is crucially pertinent to this subject of the Self.

My second assumption: most things we ordinarily say about ourselves are true. For example, I was in Princeton in March and in Oxford in May. Therefore I am a continuant, in the simple and ordinary sense that I existed throughout at least some stretch of time.

### The no-self view

On one view the putatively referential use of ‘I’ creates a grammatical illusion. We recall Lichtenberg’s critique of Descartes’ *Cogito*: there is thinking going on, but it does not follow that there is something that thinks. The view that there isn’t anything at all that thinks I’ll call the No-Self view. I do not subscribe to it, for I insist that although *the self is not some thing*, nevertheless *it is not nothing*.

<sup>6</sup> I must here leave aside how first-person, indexical expression differs from ‘objective’ description.

<sup>7</sup> On Galen Strawson’s view the question ‘what is the self?’ is distinct from ‘what am I?’. I rely here on an unpublished lecture at Princeton University as well as his ‘The Self’, *J. Consciousness Studies* 4 (1997) 405–428.

Hume was certainly right that looking within oneself, one does not find any such *object* as the self or as his mind. This thinking *thing* (whatever it was that Descartes maintained we know better and more clearly than anything else) is not findable by introspection. Good point. Such phenomenological reflection however, if carried further, refutes identification of the self *with any recognizable object at all*. Introspecting I notice only my thoughts, my feelings, . . . , not the putative entity whose thoughts or feelings they are.<sup>8</sup> But this applies to my hands and feet as well. Here is Sartre's protagonist Matthieu in *Le Sursis*:

He extended his hands and moved them slowly along the stone balustrade. . . . But just because he could see them they were not his own, they were someone else's hands, they were outside, like the trees, like the reflections on the Seine, hands that were cut off.<sup>9</sup>

This is a universally recognizable experience: when we focus on the hands as objects – as when we focus on our thoughts – they become foreign objects.<sup>10</sup>

The underlying point here, about the form that experience can take, is disputed by Quassim Cassam in his *Self and World*.<sup>11</sup> He argues for a 'materialist conception of self-consciousness' in the following form:

our awareness of ourselves is a matter of one's being sensibly or intuitively aware of oneself qua subject as a physical object

Cassam's argument is ostensibly phenomenological; in fact, it accepts a menu of traditional options, which he eliminates one by one until only that conclusion remains. His argument actually pays scant attention to the phenomenology of experience. That we experience ourselves as incarnate, as involved

<sup>8</sup> Resist the temptation to postulate an object truly there but inaccessible to such reflection, as 'substratum' for what does appear. The self *transcends* all such appearances but that does not imply that the self is a *transcendent object*.

<sup>9</sup> 'Il étendit les mains et les promena lentement sur la pierre de la balustrade. . . . Mais, justement parce qu'il pouvait les regarder, elles n'étaient plus à lui, c'étaient les mains d'un autre, dehors, comme les arbres, comme les reflets qui tremblaient dans la Seine, des mains coupées.' Sartre, *Le Sursis*, *ibid.*; p. 296.

<sup>10</sup> If they are mine, *where* is the *me* whose they are? However, this experience does not push us inexorably or logically toward the view that there is no me after all, as if there were only my thoughts, hands, feelings, and so forth, travelling together like a small circus troupe. Experience is neither the warrant for, nor the fruit of, metaphysical beliefs.

<sup>11</sup> Quassim Cassam, *Self and World*. Oxford: Oxford University Press 1997.

in nature, is a far cry from that materialist conception of self-consciousness.<sup>12</sup>

### Incoherence of the no-self view

Whatever meta-linguistic dress we put on the No-Self view, it appears to imply that I do not exist. But I do exist! Can this ostensible contradiction be finessed? The options are just two: to reinterpret the thesis about 'I', or to reconstrue 'I exist' as deceptive in its grammatical form.

Call the latter option *Version One* of the No-Self view: there is something drastically wrong with our way of speaking and thinking. When we say 'It rains', 'It thunders' or 'It snows' the word 'it', though grammatically the subject term, does not have a referent. By analogy, the surface subject-predicate form of 'I am thinking' may be deceptive.

This version is actually just a promissory note that we have no way of cashing in.

*(a) When thinking is going on, is it also the case that there is nothing that is thinking?*

Judging by the surface grammar, 'Nothing is thinking' appears to contradict 'I am thinking'. The difference is that in the case of 'It rains' we can add an explanation whose surface grammar is not deceptive – about humid air and condensation. Therefore to press the analogy is to invite a similar completion, and what are the options there? Materialist or dualist stories about how there is thinking going on. But such a completion entirely defeats the charm and appeal of the 'grammatical illusion' move.

*(b) How do we distinguish 'I am thinking' from 'You are thinking', since in both cases thinking is going on?*

In the case of rain, we can say of course that it is raining in Princeton and not in Oxford. Again, the analogy invites some completion, to locate the thinking in various places, bodies, minds, or the like; and we may enter the same comment as before.

*(c) How do we construe 'I exist now, but did not exist before 1900'?*

<sup>12</sup> I doubt the very intelligibility of this 'as . . . qua . . .' locution; but let this pass.

The example of thinking, natural enough when confronting Descartes, is rather special. If the No-Self view is to be taken seriously it must extend to other uses of 'I'. In this example, I convey the information that the person BvF has existed only after 1900. So the most natural construal is that 'I' does have a referent, namely that person. Why here and not everywhere else?

*Version Two* reconstrues the thesis that 'I' is not a referring term to: its uses do not ever refer to an object. I am not an object of any sort. But that allows us to add that my use of 'I' does refer, namely to me – who am not an object.

### **Identification: I am not a thing**

Why am I not the most obvious thing that people see when they look at me – namely my body?

My reason for this denial is the unforgiving, uncompromising Principle of the Excluded Middle. I do not insist that this principle holds with logical universality. But it holds for discourse about bodies and physical objects in general. If I am my body then there must be a fact of the matter as to which object in the world is my body. Once we test this presupposition it fades into thin air. There is an insightful dramatization of the options open to an advocate of body-self identity in a novel by Italo Calvino.<sup>13</sup>

### **The non-existent knight**

Charlemagne is inspecting his arrayed knights before a military campaign against the Infidels. One knight on a white charger, in immaculately gleaming armor, has not raised his visor. Commanded to do so, he demurs; when he does raise it, there is nothing inside the armor. 'I do not exist', the knight explains. If he does not exist, how does he do his duty? 'By willpower and faith in our holy cause!' 'Oh yes, yes, well said, yes, that is how one does one's duty', the Emperor responds, a little thoughtfully.

<sup>13</sup> Italo Calvino, *Il Cavaliere Inesistente*. Torino: Giulio Einaudi Editore S.p.A., 1959.; tr. Archibald Colquhoun in Italo Calvino, *Our Ancestors: The Cloven Viscount. The Baron in the Trees. The Non-Existent Knight*. New York: Vintage/Ebury, 1980.

This non-existent knight, Agilulf, has armor, a horse, a sword, habits of movement as well as of thought, though he has no body. What does he mean then, that he does not exist? Without accepting a general materialist position, one may hold that human existence is necessarily or essentially embodied. So if with the existentialists one reserves the term for human existence – then there is such a person as Agilulf but he does not *exist*.

The knight Agilulf (who has no body) and his squire Gurduloo (who cannot distinguish himself from his material surroundings) are the illuminating extremes:

The only person who can be said to be definitely on the move is Agilulf, by which I do not mean his horse or armour, but that lonely self-preoccupied impatient something jogging along on horseback inside the armour. Around him pine cones fall from branches, streams gurgle over pebbles, fish swim in streams, . . . ; but all this is mere illusion of movement, perpetual revolving to and fro like waves. And in this wave Gurduloo is revolving to and fro, prisoner of the material world, he too is smeared like the pine cones, fish, maggots, stones and leaves, a mere excrescence on the earth' crust.<sup>14</sup>

We, in contrast, are 'creatures of the middle', to echo Pascal. But the extremes illustrate aspects of the sort of existence I (and you reader, *mon semblable, mon frère, ma soeur*) can claim.

### The Gurduloo problem

What in the world is my body? Where is the demarcation between that body and the rest of nature? We can consider ways of drawing the line by purely physical criteria, by phenomenology, by psychological factors, or by social construction.

When I think about the crucial importance of my body in the determination of what I am, I think first of all about how I express myself bodily, through posture and motion as well as speech, writing, and other manual labor. But I express myself also through my clothes, car, house, garden, . . . Should I not count these part of my body, in an extended sense, if I am to have a plausible sense for 'I am my body'? True, those are all detachable parts. But so

<sup>14</sup> Calvino, *Non-Existent Knight*; in Ch. 9.

are my hair and nails. True, I can change my clothes without endangering my identity; but so can I change my hair color and even skin color (at least from pinko-grey to bronze). True, I can exchange my house for another house and car for another car. But I can also have a cornea, kidney or heart transplant. Transplants of both sorts are only contingently easy or hard: under certain conceivable social and economic conditions, a heart transplant may be easier to achieve than a change of jeans.

The difficulty of distinguishing our bodily self from our physical environment is graphically portrayed by Gurduloo. We first see him paddling among a flock of ducks, quacking, 'the rags he wore, of earthen color . . . had big greenish-gray areas the same color as feathers'. (p. 24) Does he think he is a duck too? The girl whose ducks they are says No, he thinks the ducks are him. Eating soup, he ends up in the kettle, exclaiming 'All is soup'. (p. 53) An old peasant says 'He's just a person who exists and doesn't realize he exists'. (p. 28. Well, it was a French peasant.)

Gurduloo is an extreme case, of the kind Harry Frankfurt calls 'wanton'.<sup>15</sup> He has no second order preferences that could sort out his immediate desires. Hence we can't read a sense of 'self-preservation' into his behavior. If we could, we would be seeing a privileged line of demarcation between his body and environment. That we cannot shows how such a demarcation is not, ever, independent of the importance of what matters to us. This makes for a problem for any attempt to identify oneself with one's body. If I am a physical object, then there is a fact of the matter as to which I am. What fact is that? How could it, without circularity, fail to be independent of what matters to me?

The difficulties in the demarcation of my body on a 'folk' level of discussion pale by comparison when the question is transposed to a scientific realist world picture. How are ordinary movable middle-sized objects to be identified in a world of quantum fields, for example? The quandaries we are in here are tellingly explored in Brian Smith's *On the Origin of Objects*.<sup>16</sup> Despairing of a purely 'objective' (physicalist) solution, one may try for a pragmatic solution, such as that *being an object* is a 'response-dependent' concept. One might elaborate the idea by saying, for instance, that the objects in this world are precisely our selective clumpings of bits

<sup>15</sup> See his *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988.

<sup>16</sup> Brian Cantwell Smith, *The Origin of Objects*. Cambridge, MA: MIT Press, 1998.

of stuff as a function of our needs and interests. (This seems to be what Brian Smith favors.) But this cannot be an option if I am asserted to be identical with one of those objects! The pragmatic turn loses us any remaining grip on the factual question 'Which object am I?'

### **Minimalist body-mind identifications**

If the Gurduloo problem, of determining one's body among all the parts of nature, is insoluble in purely physical terms then we might look to other criteria: biological, medical, legal, psychological, phenomenological, social. I submit that these are either linked indissolubly to specific purposes (with the constraints dictated by special interests) or else intensify the Gurduloo problem.

#### *The life sciences*

My body: which object precisely is that? In the life sciences there has been need for a precise answer to this question.<sup>17</sup>

Such an animal as the human is topologically speaking a torus, since the gastro-intestinal tract is a pathway going all the way through. The surface of a torus divides a three-dimensional space precisely into two parts, one finite (interior) and one infinite (exterior). My body is precisely the interior of this torus, plus its boundary – the skin, to be precise. Of course the precision is only of a degree sufficient for the life sciences' purposes. Some conventions also need to be introduced: hair is largely outside the skin, but we can re-classify it as skin; nails too of course. Such conventional choices are also constrained by the purpose at hand.

Alternatively one may define the interior/exterior by sterility: 'outside the body' is any non-sterile area. This removes the outside surface of the skin, but also certain regions we tend to think of as 'inside'. A more serious ambiguity derives from the functioning body's active resistance to things that do not respond in certain ways, through its immune system. Daily lifelong administration of immunosuppressive drugs is part of the regimen of all transplant recipients. If the body does not reject the transplant or implant during this drug regimen, is the tolerated artifact at that

<sup>17</sup> Thanks to Carmen dell' Aversano for helpful discussion.

time part of the body? It may be entirely in the interior of the torus, and entirely sterile – but the body originally recognized it as alien, and drugs were needed to depress the immune response.

...

Imagine that the philosophical answer *Yes: I am my body* is met with the response: *In that case you are the interior plus boundary of the above described torus, qualified as indicated in the amendments listed.* We cannot read this response as scientific support for the assertion as a factual claim. We can only read it as specifying what the assertion means if read as a sentence *in an idiolect created to serve certain professional purposes.*

Suppose however we took ‘The body is [thus and so, such and such an object]’, presented as a *factual claim* coming from the life sciences. We would then still have to ask ‘Fine; but are they right?’ There have been many false hypotheses that were once accepted in science. So this would not be an idle question. But it would be puzzling, for what it asks for does not seem to be a matter to subject to experiment. An experiment could test whether the interior of this torus is everywhere sterile, but how could it test the hypothesis that the entire interior and/or the surface are part of the body?

Legal arguments (e.g. about whether preserved sperm should be buried with the donor when deceased, and limits on lap dancing) do not simply draw on the life sciences for demarcation of the body.<sup>18</sup> This underlines the purpose-relative character of the question.

#### *The absolute minimum?*

If I am wearing clothes, perhaps I can say: my body is everything that is inside my clothes. That leaves a little uncertainty about the head and hands sticking out. If I am wearing a space-suit, I can try ‘everything inside the space suit’. But that includes air and moisture outside my skin. The presence of this air and moisture are crucial to my existence there, as much as my blood and lymph. By what criterion can we exclude the absolutely necessary life support system for the bodily functions from the body? Alternatively we can ask: if we allow for any and all forms of life support as equally legitimate, how much of me has to survive for me to

<sup>18</sup> I have only anecdotal evidence, and would appreciate references; see further Andrews, Lori B. and Dorothy Nelkin, *Body Bazaar: The Market For Human Tissue In The Biotechnology Age*. New York: Crown, 2001.

survive? Eric Olson's *The Human Animal* argues that only the brain stem's existence is necessary (and as a matter of fact, also sufficient) for the survival of the same animal.<sup>19</sup> Does that give us the principled criterion for demarcating the self?

Consider the principle that I am whatever thing is such that its existence is both necessary and sufficient for my survival as the same animal. If Olson is right, this principle would then imply that I am my brainstem. (I am not attributing this conclusion to Olson.) That I am my brain stem is at least a possible extreme of the 'I am my body' variants. Here after all we have the engine of reason and the seat of the soul, according to some – is that not precisely enough? Leave aside how curiously and arbitrarily selective it is to ask what is necessary and/or sufficient for survival after first excluding conditions of life support from consideration. We cannot accept this principle while also taking the self to be the person that I am, for the candidates this principle sanctions just do not seem to qualify as persons.

### **The unforgiving excluded middle**

Attempts to identify myself with my body are thus thoroughly stumped by the Gurduloo problem.

The problem at issue is not the vagueness of our concepts, whether ordinary or crafted for scientific and practical purposes. There are bodies in the sense of the life sciences. There are well-defined organisms to be studied in those sciences, from cells to entire ecosystems. The level of precision in definition suits the purpose at hand; no more is required. Nor is the problem Gurduloo's case raises a problem about vagueness in general. The contours of most bodies to which we freely refer are indeed vague: clouds, roads, seas, prairies, mountains, . . . Nevertheless, K2 is a specific mountain, although it is not a specific heap of earth and stones; and it is distinct from other mountains such as Mt. Etna or Mt. Everest. Even vague objects can be distinguished from each other. So if there is a general problem of vagueness, it does not remove the question before us. We can sensibly ask which of those

<sup>19</sup> *The Human Animal: Personal Identity Without Psychology*. Oxford: Oxford University Press, 1997. Thanks to Eleonore Stump; see further her challenging 'Non-Cartesian substance dualism and materialism without reductionism', *Faith and Philosophy* 12 (1995), 505–531.

vague objects I am. The Gurduloo problem is not which body is which, although that is pertinent, but

How can I possibly *identify* myself with any part of nature, however defined, given that I can discern only differences of degree and delineations that shift and deform from one perceptual context to another?<sup>20</sup>

*The Excluded Middle is very unforgiving.* If there is a fact of the matter about whether I am a physical object or system, then I am at most one of these so demarcated objects, and not another, and the facts must settle the matter. *But what facts could settle it?*

A philosopher may be tempted to replace that question with 'Is there any non-arbitrary, principled way to make the selection?' These two questions are not at all the same! A non-arbitrary, principled distinction may just be one that appeals to certain indisputable theoretical virtues. Then an ostensibly factual question is being settled on the basis of preference – however non-arbitrary – for certain kinds of theory. *That is not settling by facts alone.* The question is not 'Which object *is it best* to postulate to be my body, in the development of a *satisfactory* scientific or metaphysical theory?' That question may be answerable, but does not respond to our concern.<sup>21</sup>

### Manifestation, not identity

... Agilulf, by which I do not mean his horse or armor, but that lonely self-preoccupied, impatient something jogging along on horseback inside the armor.<sup>22</sup>

While identity is a black-and-white sort of relationship (I am either identical with x or distinct from x), my relationship to any of the things I call mine is not. We find only graduated and nuanced relations between ourselves and the things that we call ours, whether physical, psychological, or social. Questions such as

<sup>20</sup> There may be no way to express this problem in any but first-person language. The essential indexical may well be at the heart of this entire discussion.

<sup>21</sup> Some disagree here, because they accept explanatory power, unification, and simplicity as a guide to warranted belief. For them it must follow that theories are likelier to be true if they have these humanly desirable characteristics. What, I ask myself, could account for that? What does nature care?

<sup>22</sup> Calvino, *Non-Existent Knight*; from Ch. 9.

'which particular object is my body?' is eventually unanswerable precisely because of its questionable – arguably untenable – presupposition of uniqueness. Therefore we need a distinct relational concept, different from identity yet non-trivial, to characterize our incarnate existence.

### **How do I relate to my body?**

I express and reveal myself through bodily features and movements: through my posture, my physical comportment, my way of walking, my 'body language', my interaction with my natural and social environment, and thereby with the other selves who express themselves in similarly incarnate form. I express and reveal myself not only through the flesh but also through my clothes, my style of dressing, whether I ride my bicycle sedately or hands-off, how I arrange my living space, by where I walk as much as by how I walk.

Some ways of expression are more direct and less mediated. While none suffices to identify me, there is also no possibility of existence devoid of all such self-expression.

There is no bare homunculus at the heart of this living, moving presence. So how shall we picture our own existence to ourselves?

Expression is possible only in a language, in a broad sense of 'language'. Body language has a movement vocabulary, and means of composition of movements that have meaning. But there cannot be such a thing as a private language: you have a language only if you are a member of a language-endowed community. That is therefore a pre-condition of the very possibility of self-expression. To this social aspect I will return below.<sup>23</sup>

### **'Ye shall be as gods'**

Satan said to Eve 'Ye shall be as gods', deceiving her into wanting what she already had. In the continuum-relation between her self and her bodily environment she already transcends all that is given in experience. This goes for her as well as for the as-serpent-appearing demon.

<sup>23</sup> Thanks to Isabelle Peschard for discussion of this point. See also Anthony Synnott, *The Body Social: Symbolism, Self, And Society*. London: Routledge, 1993.

Mythical depiction of gods, angels, and demons dramatizes the distance between being in the world and being an object, as well as their self-expression through diverse incarnate and natural appearances. Needing a nuanced relational term also for us, let us therefore adapt one from myth. I *manifest* myself in nature, through my body, my movements, my words, my decorating and clothing activities, my artistic endeavors and literary output, my passions and tantrums, . . .

Think about us in the way the myths depict the Homeric and Hindu Gods, as well as the lesser supernaturals in the religions of the Book. (Remember Leda and the swan, Krishna as charioteer in the Bhagavad Gita, Shiva who appears in creation and destruction, Satan's minions who at their choice appear to us as incubi or succubi.) The gradation that replaces both self-body separation and self-body identification is then evident. I manifest myself through temporarily enduring parts of nature such as my torso and limbs, as well as through my habitual though inconstant appearance as clothed in certain materials or none at all, as walking rather than driving, and so forth.

### My physical attributes

I have many physical attributes, such as mass, position, and velocity. I have these precisely because I have a body; more accurately: because I am embodied. Our ordinary way of speaking about this fact underlines two salient points. First of all I speak of my body in the same way I speak of my clothes, house, car, friends, as something that is *not me but mine*. Secondly, the fuzzy boundaries that characterize those objects characterize equally the transfer-attribution of their physical characteristics. What is my body mass when I have just eaten and not yet digested a large pizza? Does it include the small parts of pizza still in my mouth, if I am still chewing? And what is my precise position when I am moving, and my hair, scarf, and billowing sleeves are blowing in the wind? Thirdly, this verbal distancing slides very easily into metaphor. For that reason we can't make it the arbiter of truth. Literally, the eye does not see. The eye sees only in the sense that the eye reads, which is the sense in which the hand writes and the mouth speaks. These are derivative properties of the parts, deriving from the whole, with the relationship metaphorically raised into an identity. Turn the metaphor back on itself. The hand writes only in

the sense that I write with my hand, but I am extended in the sense that I have limbs and the like which are extended.

### Spatial location

If the relations between ourselves and the objects, events, and processes in the world admit of degrees then many questions about us will be subject to irremediable ambiguity.

*Where am I?* If I were identical with a specific physical object, that question would have a definite answer. But does it? Trying to identify my location, I naturally turn to the center of my sensory perspectives.<sup>24</sup> A famous psychological experiment in which the subject aligns sticks to point at himself confirms our intuitive response. That places me somewhere in my head, behind my eyes and between my ears. Immediately I follow Descartes in picturing myself as a small homunculus sucking its sensory life out of the pineal gland. But now I close my eyes, it is deep in the night with not a sound anywhere, I move around, I feel around me. Now my perceptual center is somewhere behind my hands, it is not inside my head, I am phenomenologically located somewhere close behind the touch of those hands.

Psychological experiments severely disturb the pictures that bewitch philosophy. Michael Kubovy's is simple and striking:

If I trace a 'b' on the back of your head, you'll report that I wrote a b; if I trace the same character on your forehead, you'll probably perceive it as a d. It's as if you had a 'disembodied eye' behind your head reading the pattern traced. We are engaged in research . . . to discover the body's natural systems of coordinates.<sup>25</sup>

### The 'causal order'

It may well be objected that I am on this view rather insubstantial. Precisely so. I am no substance, nor was meant to be. I am

<sup>24</sup> See Michael Kubovy, Ch. 9 'The psychology of egocenters' in his *The Psychology Of Perspective And Renaissance Art*. Cambridge: Cambridge University Press, 1986.

<sup>25</sup> Kubovy, *ibid.* and <http://www.virginia.edu/~mklab/skin.html>. But 'natural coordinates'? The origin and orientation change in this case while the body's spatial configuration remains the same.

present, and I act, but am not an object entering into the interactions studied by the natural sciences. I may be said to do so derivatively; for I have a body – however circumscribed in a given context – which does enter into those interactions. The body and its physical interactions are the subject of physics and physiology; not so my actions.

Princess Elizabeth of Bohemia asked Descartes: *how does the mind interact with the body?* Descartes had no satisfactory answer, nor could he have. What should we answer her if she raises her question again, for our current account of the Self? Briefly: *there is no need for mechanisms of interaction* if I want to act, *precisely because I am not an object*. Neither forces and collisions between physical objects nor – if such there be – powers, potencies, or principalities to connect the abstract or occult with the physical, are pertinent. Not being an object, when I wish to act I just do it.

This may take a moment to assimilate. If I am not a physical object, how can I exert force on the ground, how can I be heavy enough to make a dent? How do I move my arm? What relation is there between my decision to move my arm and the contraction of the muscles that produces this movement?

On their most literal reading, these questions trade on the assumption that I am a body interacting in a bodily way. Bracketing that assumption, we see that such a question as ‘How do I move my arm?’ can legitimately be construed in only three ways.

- a) The straightforward request addressed to the physiologist and physicist, to explain the bodily movement.

The scientist is to provide us with a theory or model, written or constructed within a given set of parameters, pertaining to the physical processes that take place when I move my arm.

- b) The question *How do I move my arm?* in the sense evident from Agilulf’s reply to Charlemagne.

Agilulf’s answer, ‘by willpower and faith in our holy cause’ locates the question in the framework of ‘person discourse’, in the ‘space of reasons’.

- c) The request for instruction, to be shown how it is done.

After I had my leg in a cast for six weeks in 1995 I needed to relearn stepping down a stairway. Asking someone else ‘how do you step down a stair?’ might have helped. The obvious initial

answer, *I just do it*, is certainly correct. But it needs to be supplemented with something more to focus my attention in a useful way.

There is no peculiarly philosophical question beyond these three, except for someone who takes a peculiarly philosophical position on the Self. Descartes could not have answered Princess Elizabeth in this way. For he had identified selves as immaterial substances, in a context with set categories of action and causation. These he had mainly inherited from the metaphysical tradition he had been at pains to dislodge in other respects, and was now asking for its due. But that is a while ago; why keep revisiting his predicaments?

### Location in the social fabric

I have a location also in the social fabric, what Bradley called my Station and Its Duties. What I ought to do derives from what I am: a citizen, a lover, a son, a father, a teacher. So does *what I am*: despicable because bourgeois, admirable because bloodied yet unbowed, and so forth. That is my position in the world, in a somewhat different but no less important sense than that of spatial position.

I manifest myself in nature and in the social world. Of course this is possible only for an embodied being. The relationships that define my social position have a strong physical core: ancestry, place of birth, verbal acts of commitment. They are displayed in my body, however narrowly or broadly demarcated.<sup>26</sup> The physical embodiment enters every aspect of human existence.

Listen to the narrator of Ermanno Bencivenga's fable 'Io'.<sup>27</sup> Once I was just me, he says, but that made things very difficult when meeting people. The question would always come up: who was I? All I could say was 'I'm me'. These meetings never led to anything. So I chose a name for myself, quite at random: Giovanni Spadone. That made all the difference. 'Oh, Giovanni Spadone! so glad to meet you! and where are you from, what do you do, who are your parents . . . ?' Of course, if they had responded instead with 'And who is Giovanni Spadone?' I would have landed

<sup>26</sup> See Anthony Synnott, *The Body Social: Symbolism, Self, and Society*. London: Routledge, 1993.

<sup>27</sup> In his *La Filosofia In Trentadue Favole*. Milan: A. Mondadori, 1991.

back in the same predicament. . . . But you know, they never do that. . . .

This moral so evident in the necessary birth of a social identity has its twin in the conditions of a social death, witnessed in the demise of Agilulf. When other knights brag about their glorious deeds, the glaring inaccuracies don't really matter – they are who they are anyway.<sup>28</sup> But for Agilulf there is no recourse beyond his history. A doubt about the deed for which he received his knighthood calls him – who and what he is – into question. He is truly defined by his station and its duties, as constituted in detail by his history; there is nothing more to him. Sadly, when he finds this doubt confirmed he wanders off alone, takes off his armor. Then there is no longer any such thing as Agilulf; and perhaps there never was. . . .

### **Pseudo-problems pertaining to the self**

Almost everything to be said on the subject of what I am is prey to readings that turn it into metaphysics. Princess Elizabeth's question concerning volitional action invited such a reading already. Let me try to prevent some more before they can take hold.

### **The unobservable and I**

In philosophy of science I take the empiricist view that accepting a scientific theory does not require belief in the reality of any *unobservable* entities it postulates. Taking this view of what science is does not by itself imply disbelief in unobservable entities. But what an uncomfortable tension, if someone agnostic about the unobservables postulated in physics were then to profess to be an unobservable entity himself!

Am I observable? Certainly I am; did you think I was invisible or intangible? But I am visible and tangible only because I have

<sup>28</sup> Calvino, *Non-Existent Knight*, Ch. 7. I resist the reading that the person must have a reality or substance that supports counterfactuals, the 'could' and 'could have' involved in what they claim to be. It seems sufficient that *they are what they will have been*, and what they will have been is still in the future, held jointly in their hands and the hands of fate.

a body; my visibility and tangibility are derivative properties. This reply may be challenged as follows:

[1] If I am distinct from a given entity, then I am observable only if there is a way to observe me without observing that entity.

[2] But observation is by the senses.

[3] Anything observed by the senses is identical with some physical entity (thing, event, or process).

[4] Therefore I am observable only if I am identical with some physical entity.

I have more faults to find with this reasoning than I can count! I will skip the obvious ones.

It would be uncomfortable for an empiricist about science to profess to be an unobservable entity. But 'entity' is another synonym for 'object' or 'thing', and I am not one. It is indeed not true to say that I am observable, except in the sense that I have an observable body. But that is because such terms as 'observable' and its contraries classify *things* (including of course events, processes, all those object-like things). Attributes pertaining directly to certain kinds of objects can apply to me derivatively, but not in any other way. With this distinction made we can see how the argument trades on equivocation. Surely what is at issue is what is directly encountered in experience, in any reasonable sense. You do encounter me directly in experience; but that encounter has certain physical relations involving our bodies as a precondition.

The most important point, though, is one about our philosophical dialogue. There is no relevant direct parallel at all between our discussion of the self and the empiricist's controversy with the scientific realist. For this talk about the Self, remember, is just talk about myself and about yourself. Whether to be agnostic is a question that simply does not arise. You are not a postulated theoretical entity, introduced for reasons of theorizing or fitting evidence; and neither am I.

### **Reduction, no; but supervenience?**

Twentieth century analytic philosophy sometimes looks like just a sustained attempt at a consistency proof for materialism (physi-

calism, naturalism . . .).<sup>29</sup> The most direct way to prove arithmetic consistent would be to reduce it to pure logic – as the logicists tried. Similarly the most direct way to prove the consistency of physicalism about thought, emotion, charm, and other such person-attributes would be to reduce all that to what is described in physics. I won't rehash the history that made us give up on both those simplistic ideas. But there remains the recently most popular position of 'supervenience without reduction!'

Does everything true about me supervene on facts about physical objects and structures? My view, as expressed so far, may be compatible with that; but it is not what I mean. I do not contradict it, but only because the entire game looks to me like a prime example of what Carnap had in mind in his 'Pseudoproblems in Philosophy'.

Supervenience is usually explained with a heavy dose of modal metaphysics, to me of dubious intelligibility. But the concept of supervenience between two sorts of discourse is intelligible. These could be the language of physics and the part of our language in which we use 'folk' psychological terms – 'person talk'. Abstractly, consider two languages, Q-discourse and F-discourse. Let the supervenience claim be:

[**Supervenience**] for anything X stated in F-discourse, if X is true then the world could not be different in that respect unless also something formulated in Q-discourse were different.

An intuitive example: I am currently thinking about dragons. I could not be thinking differently at this moment unless something physical (e.g. the state of my cortex) were different as well. Rendered in terms of language, we construe this as: the F-sentence 'I am thinking about dragons' could not be false without some change in the truthvalues of the Q-sentences as well.

Although I have used the subjunctive, this can be explicated as follows. Both forms of discourse admit of many models

<sup>29</sup> In logic or mathematics, when a consistency result is proved the next challenge is to develop alternatives. Given the consistency of Euclidean geometry, the challenge lies in non-Euclidean models. If ZF is consistent with the axiom of choice we want at once to know if there are models of ZF that violate the axiom. This attitude leads to the creation of new methods and strange structures, hence new understanding. A consistency result for materialism should similarly be philosophically exciting only if it reveals ways to create new strange, paradoxical, and esoteric *alternative* conceptions of what we are.

(intuitively, representing 'ways the world could be') and these models can be combined into models for the two together only in certain ways. For these combined models, what the above claim asserts is this:

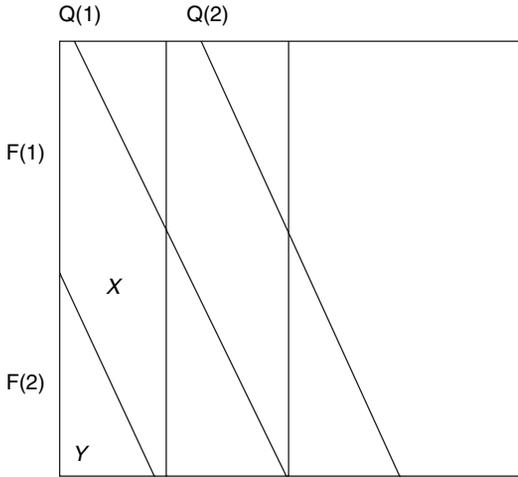
**[Supervenience]** there are no distinct combined models in which a given F sentence has different truthvalues while all Q sentences have the same truth values.

Suppose that the supervenience claim is correct.

It is easiest to think about this in terms of the maximal consistent sets of sentences within the given language (roughly, what Carnap called 'state-descriptions', ignoring his restriction to finite size). Such a set sums up all that is true in one given way that things could be, to the extent formulable in that language. Let these 'state-descriptions' of the Q language form the family  $\{Q(i): i \text{ in index set } I\}$ . Finally, let us write [S] for the set of combined models that satisfy given sentence S (of either discourse).

Then if F is an F-sentence, there can be no combined models M and M' such that F and Q(k) are true in M while F is false but Q(k) true in M'. Hence the set of models satisfying Q(k) must either lie entirely within the set of those satisfying F, or entirely outside that set. That is, [Q(k)] is either included in [F] or else disjoint from [F]. Since the state-descriptions jointly exhaust what can be true, any model that satisfies F must lie in one of the [Q(i)]. Putting these two points together, [F] is just the union of all the sets [Q(i)], with i in I, that are included in [F]. In our example that means that a 'personal discourse' sentence is just the 'disjunction' (in a suitably extended sense) of 'physical discourse' descriptions.

Let us display this argument graphically, with a diagram depicting a violation of the supervenience claim. (I am representing propositions by the sets of worlds in which they are true). Let X and Y be worlds: in our diagram they belong respectively to contrary propositions F(1) and F(2) in F-discourse. But in Q-discourse, both make Q(1) true. Hence things are different in X and in Y with respect to aspects expressible in F-sentences, but the same in any respect describable in Q-sentences. Therefore the F-discourse does not supervene on Q-discourse.



Conversely then, if one discourse does supervene on the other, we must be able to construe its propositions as such ‘disjunctions’. This will not amount to reduction if these ‘disjunctions’ – which are sets whose elements are maximally consistent sets of sentences – are not finite, not recursively enumerable, and not definable or specifiable in any other way.<sup>30</sup>

See how close this is to straightforward reduction! To have supervenience without reduction means to have no translation sentence by sentence or paragraph by paragraph or even definable set by definable set. . . . but there is still a perfect description ‘at the far edge of infinity’.<sup>31</sup> The supervenience claim then still entails only that there is, so to speak, a reduction for God or for the angels, just not for finite beings like us.

*This is obviously a position designed to be irrefutable.* What are the benefits of believing in such a relation of persons to physical objects? The mere assurance of consistency? Cold comfort! Add to this that no such ideal ‘physicalist’ language exists, or is likely ever to be had . . . Why play these games?

<sup>30</sup> Our specifications will be finite sets of sentences in our own language, of which there are only denumerably many, while even the family of ‘state-descriptions’ is already larger than that.

<sup>31</sup> I use Pascal’s phrase, so apt in a similar context.

**'How does consciousness emerge?'**

Once upon a time there were only inanimate objects; now there are living, behaving organisms, and some are conscious. How did consciousness emerge?

As a question for science, this is well-posed to the extent that the term 'consciousness' is well defined. But taken in that way, there is no problem in principle at all, regardless of one's philosophical stance. Agilulf's case can illustrate this point. At first sight his existence would raise the scientific problem: how could he possibly lift his sword? If there were such a knight, science would have the task of constructing an adequate model of the phenomena thus displayed. To begin it would produce a barely adequate model with little or no predictive power. That model would be replaced by a better one, and so on, as long as we had more factual questions to be answered and nature cooperated. The models would be accompanied by testable explanatory narratives.

What argument could there possibly be that scientists could not succeed in this? There are observable phenomena as the suit of armor, sword, and shield move through the world; science models those phenomena. Whatever the phenomena are there can be a scientific story that depicts them, with unobservable things postulated and inserted if needed. That is simply what science does.

Similarly, there must necessarily be a scientific story about the transition classified as the emergence of consciousness. That is the change from when the world held as yet no living organisms to when it did, and from the states in which there was as yet no consciousness to when there was. The question 'how does consciousness emerge?' has therefore a very straightforward answer at any given time. That answer is provided by the model (perhaps still very unsatisfactory or inadequate) which science so far provides for that geological/biological stage. So we have an answer already, though it is as yet disappointingly uninformative; but it is improving all the time.

You can readily see the analogy between the relation of science to the emergence of consciousness (however defined) and Hume's confusion about miracles. Take any given miracle, such as water turning into wine. If this really happens, then physical science is adequate only if it implies the possibility of that happening. If scientists accepted it as a phenomenon to be accounted

for, they would account for it. They would produce models, accompanied by explanatory stories, that fit the phenomenon. If the problem posed by the phenomenon is severe (as it was with radioactivity, for example) it may engender pervasive foundational changes in the natural sciences. The question 'How can miracles happen?' *cannot be* 'How can there be observable exceptions to empirically adequate scientific theories?' For the latter is self-contradictory.

Philosophers who confess themselves perplexed by the question *How could consciousness possibly have emerged at all in this physical world?* must be understanding the question in a different way, a way that has nothing to do with science. Their problem is a pseudo-problem unless (a) they have a clear and distinct understanding of this as different from the scientific question, and (b) they do not confuse the question so understood with the question as posed in any scientific context. I see no evidence that these desiderata are met.

There is a mystery of consciousness. But it is not among the mysteries that the sciences confront, which they so fortunately and habitually address and solve.

### Disquisitions on substance

Does my account of the self not make it a substance?<sup>32</sup> The technical philosophical meaning departs, as usual, from our ordinary use, where cloth, water, and molasses are substances to be found in the home. One thing the technical question can mean is whether the self is a substance in the sense of 'a being which can subsist by itself, without dependence upon any other created being'. There is a rival technical construal perhaps equally familiar: 'the bearer of attributes, the subject of predication but is not itself predicable, that which receives modifications and is not itself a mode'. And there is a third, the most important one according to Leibniz: 'that which acts'.

Am I – is my self – a bearer of properties and a terminus of relations? Certainly, in the innocuous sense implied by the fact that I am thinking now, have a body subject to sickness, old age, and death, and have a brother as well as a disillusioned distant

<sup>32</sup> Question very reasonably raised by Peter van Inwagen at the Reading conference.

admirer perhaps, or something of the sort. Am I the subject of predication and not myself predicable? I should hope so. Do I act? There is no doubt that I do. Not only do I raise my arm, and respond to critical questions, but I do so deliberately and intentionally, thus clearly engaged in action and not just behavior.

Since my affirmative answers derive from very ordinary every day assertions, for which I deny any need of metaphysical underpinning, I may still be misunderstanding the question. In any case, I have as yet omitted what is perhaps the most important sense. Am I – is the self – something that can subsist by itself, without dependence upon any other created being? Of course not. Suppose even (however implausibly) that for each person and thing in the world it is the case that I could have existed without it. It still does not follow that I could have existed without anything else at all.

If substances are things, then I am not a substance. If being a substance requires only that most ordinary things to be said about me are true, then I am. But of what use is that to metaphysical theorizing?

*Princeton University*  
*Princeton, NJ USA*  
*fraassen@princeton.edu*