

Stephen Priest *Light on the Subject* (2016 Draft)

Light on the Subject: Kant and Hegel on Self-Consciousness

Stephen Priest

Stephen Priest is Senior Research Fellow in Philosophy at the University of Oxford.

Contents

1. Kant's Deduction	3
2 Hegel's Master and Slave Dialectic	80
Appendix: Sense-Certainty and the Natural Attitude	127
Bibliography	130
Index	

KANT'S DEDUCTION

by

Stephen Priest

A 103 - A 130 p. 1-36

B p. 36

Kant's transcendental solution to the problem of how experience is possible is insufficiently metaphysical. *Pace* Kant: The transcendental is the transcendent.

Without synthesis experience would not be possible according to Kant. Unless intuitions were subsumed under concepts and we could be conscious that what we think is the same as or meaningfully connected to what we thought a moment before, 'all reproduction in the series of (re)presentations would be useless'. (A103. 133) Our having concepts depends upon some of them being categories. That we have the categories we do is also a necessary condition for our experience of the physical world. (A 135. 135) But Kant thinks these conditions of experience themselves need justification, so the Transcendental Deduction involves a discussion of the self.

Why so?

He says that 'All necessity without exception is grounded in a transcendental condition'. (A 106. 135) In other words, if *p* is necessary then there is always some non-empirical explanation of *p*'s necessity. There is an equivocation on 'necessary' here. Read psychologically, 'necessary' means something like 'could not conceivably be otherwise', read logically or metaphysically 'if true, then could not be false' or, 'if the case then could not be otherwise'. Kant tends to assimilate these senses, perhaps not in the last resort without good reason, because it might turn out that the cognitive constraints on our thinking delimit precisely what counts as a logical or metaphysical possibility. Then no sense could be attached by us to a distinction between 'possible' and 'thinkable by us self-conscious beings'. But *prima*

facie the senses are quite distinct.

A condition or a ground of x is what makes x possible. So if a is a condition or ground for x a is at least a necessary condition for x. By the use of 'transcendental' here Kant is excluding the need for an empirical search for conditions.

Transcendental knowledge is knowledge of how knowledge is possible, and these conditions are not to be met with in experience because, allegedly, they are what make experience itself possible.

If there are conditions for experience then it seems coherent to suppose they are transcendental partly at least in Kant's sense. This is because, suppose we wish to explain how x is possible. Then the conditions for x's possibility cannot (except tautologically) be part of x, otherwise they would already stand in need of explanation themselves. They would be ever part of the *explanans* and never the *explanandum*.

Non-trivial grounds for experience lie outside experience. However, Kant neglects a possibility for which there is conceptual room. Some ground might be transcendental in relation to some experience and be empirical in relation to a distinct experience. I call such non-Kantian grounds transcendental. Kant neglects transcendental grounds because he tries to show how experience as a whole is possible. He assumes that experience as a whole has wholly transcendental grounds. This however is a *non-sequitur* and might be false.

If Kant's project is coherent, we need to ask next what exactly the transcendental unity of apperception is invoked as a condition of. I have isolated the following as ultimately grounded in the transcendental unity of apperception (TUA) as their condition. There is no further transcendental condition for the TUA, but each of these requires the TUA:

- a) Synthesis
- b) Concepts of Objects in general
- c) all objects of experience
- d) the forms of intuition (space and time)
- e) the laws of nature
- f) rules
- g) the objective reality of our empirical knowledge
- h) appearance in experience
- i) the categories

j) the unity of consciousness

Clearly some of these are conditions for each other. For example, the categories (i) are conditions for concepts of objects in general (b) and the laws of nature (e). Also, the unity of consciousness (j) is a condition for all the items listed under (a) to (i) inclusively. Although (a), synthesis, seems to ground it reciprocally, even the unity of consciousness is made possible only the TUA. As Kant says 'There must be a transcendental ground of the unity of consciousness'. (A 106. 135)

Because it grounds all other conditions Kant says 'This original and transcendental condition is not other than transcendental apperception' (A 107. 136). We need now to understand what this is and decide whether it exists.

Kant distinguishes transcendental from empirical apperception. Empirical apperception is the intellectual component of self-consciousness, of which inner sense is the experiential component. In addition to this alleged faculty of self awareness there exists a non-empirical or transcendental capacity to be conscious of one's numerical identity over time. It is perhaps for this reason that Kant's theory of the self is more satisfactory than Hume's. Without naming Hume, Kant at A 107. 136 presents a very Humean view of introspection. He says:

'Consciousness of self according to the determinations of our state in inner perception is merely empirical and always changing. No fixed and abiding self can present itself in this flux of inner appearances. Such consciousness is usually named inner sense or empirical apperception.' (A. 107. 136)

Compare Hume's famous remark that he never happens upon any single enduring impression of self that could give rise to the idea of himself as persisting over time. The TUA is invoked to solve the problem of the self Hume bequeathed: What makes my experiences mine? But, is it needed? Is it true that no single abiding self is presented to oneself in self-awareness? What makes one set mine and another yours? There are two component hidden in this question, one to do with individuality or uniqueness the other to do with subjectivity.

Taking individuality first, if we ask 'what makes this mental field yours, and this mine then this can be analysed partly along these lines: What makes x not y and what makes y not x? Once we possess well defined concepts of any two entities, then if they are numerically distinct then to say x is not y or y is not x is to point out a logical or modal property of x or y. I am not you and you are not me. My soul is not your soul and yours is not mine. If, as we are postulating, I am my soul and you are yours in a weakened transcendental sense, then these sentences express the same proposition. The non-identity of numerical distinct particulars is not just a modal property of each but one that is logically primitive. I mean the rule $x, y (x \neq y) = (x \neq y \neq x)$ does not admit of any non-tautological explanation, or any derivation from premises which do not themselves employ the axiom of identity. We could say it

follows from $A, x, y (x = y) = (y = x)$ than $A x, y (x = y) = (y = x)$ but this is hardly likely to convince someone who was unconvinced by the negative formulation of the principle. We should, I think, point out that someone who putatively denies either principle has not succeeded in saying anything sensible, not even in producing a denial.

The idea of uniqueness, it's being logically primitive that each x is just the x that it is and not some y , is part of the answer to what makes my mental states mine and your yours. That I am not you and you are not me is not a point just about selves. It is a particular case of the modal fact that x is not y and y is not x if they are numerically distinct.

But an important issue remains. This is what 'I am not you' says that 'this one is not that one' has not expressed already. What information does 'I am this' provide over and above 'This exists'? What information does 'You are that' provide over and above 'That exists'? To understand this we need to supplement the account of individuality or uniqueness with one of subjectivity.

A traditional and Kantian account of subjectivity would go something like this. Many items fall within my experience but one in particular never does. This the I that actually has or is the owner of these experiences. This is subjective in the sense that it is always experiencer but never itself the object of any experience. Now, there is a great deal wrong with this account of subjectivity; especially, its assumption that experience is to be understood as a relationship between subject and object. But, if there is anything which really exists which is experienced but not often experienced it is the soul. Again, though, this is a contingent fact and one that is capable of empirical verification. I mean, for example, although it is often pointed out by physiologists that the soul does not directly sense itself, has no 'feelings' (Hearth) it does not follow that it could not in principle observe itself, as Wittgenstein has pointed out (BB).

So, the soul is the transcendental subject in the sense that typically, when transforming the physiological input into the mental field, it is not thus transforming itself (except, in an imaginative way when it thinks of itself).

It is also a contingent fact that the soul is the transcendental subject. I mean by this something else might have been. As Wittgenstein points out, if my skull were to be opened it is logically possible it should be found to be empty. It seems also, that some highly intelligent people have very little soul tissue.

Suppose the physiological structure of the soul were quite altered so that it could transform itself wholly into part of the mental field: in the way at present it transforms light waves into colours with hue and saturation. This is not difficult to suppose. After all, the physiological input it transforms already is quite immense in amount and complexity. Then, I maintain, there would in just that degree be no such thing as subjectivity. The soul is a gap or void in the field of its mental states. If this gap were to be closed there would remain individuality but not subjectivity. It is the opacity, most of the time, of the soul; its not featuring amongst its own mental field, most of the time, that gives rise to the subject-object relational structure of

experience.

One final point is partially illustrative and partly constitutive of this account of subjectivity: It is true that when you see your eyes do not see themselves. We do not on that account speak of one's eyeballs as 'subjective', even though prima facie they are good candidates for the 'subject' position in some account which entails perception is a relationship between perceiver and object perceived. If we wish to talk about a subject of visual perception then the eyeballs should feature in such an account. No doubt other physiologically necessary conditions for a person seeing an object will also be mentioned: the optic nerve part of the soul etc are causally necessary, so also arguably are other aspects of the person which are subjective: contingently subjective. But the physical eyes are clearly contingent transcendental conditions for that experience we call seeing. This is most clearly metaphysically testable: damage or remove someone's eyes and they cannot see. To put it in Kantian terms: the eyes are a transcendental condition of seeing even though it is analytic that we use the eyes to see. The eye is just whatever the organ of sight is.

I should not wish to conclude from this that subjectivity is physical because that term has misleading connotations. Its use here might suggest that descriptions of the subjective could be reduced to descriptions of the objective. This translation is not I think feasible, even if one and the same set of entities is described in two vocabularies: one subjective and one objective. But, if subjectivity is not physical then it is not clearly mental either. The eyeballs as part of the subject are not mental or spiritual even though subjective.

I suggest the empirical testing of the subjectivity of eyeballs may be extended to the soul as subject of the mental field in general with advances in neurology. Of course the soul and the eyeballs are physical in what I shall call the naive-Newtonian sense of 'physical'. This concept of physical employs as its paradigm of the physical object a physical object as observed. We need to make room, if we wish to retain the word 'physical' to talk about subjectivity, for this: I am a physical object, or there is something it consists in to be a physical object, and this does not consist in observing a physical object.

We are now equipped to answer the original questions. I have knowledge my numerical identity over time for the following reasons: I am, transcendently speaking, my soul. My soul endures over time $t_1 \dots t_2$, and that soul that I am is able to formulate the judgement 'I am now, at t_2 , numerically the same being as at t_1 . This judgement is possible because the field of consciousness between t_1 and t_2 has an uninterrupted history. (I leave aside cases of falling asleep, comas etc because solving Hume's problem does not require addressing the issue of 'breaks' in experience). There is no awareness of this history, except the judgement, the propositional thought, that it exists and is of a certain nature.

This transcendental account of one's identity over time is then just the account of the identity of one's soul over time. If we wish, we may say of the history of the mental field: 'I am that history'. This is the correct analysis of what Kant calls the empirical self. I am metaphysically that transformation of my environment that I

effect as my soul, transcendently.

In a sense then the problem of personal identity does not arise because it was misconceived. It was the wrong sort of question to ask. I am (metaphysically) a set of mental states. Once we have answered the questions; What does it mean to call them mine, and what makes them a set in a way that identifies oneself with them, no residual questions remains of the form: how do I know I endure? or what is the permanent element in the set that is really me? I am the whole set, not an element in the set. This captures another part of the truth contained in the Humean account of subjectivity. It is true that I am not an item within my experience, but neither on the other hand am I an item outside it. I am it.

This is to answer the question 'What am I?' or 'What is the self?' in an empirical tone of voice. If we ask it in a - transcendental voice, the answer is: I am a soul: A specific soul with a unique spatio-temporal career.

What I hope to have done is sketch an answer to Hume's and Kant's question which is consistent with actual and foreseeable neurology, and thus which is metaphysically testable. This has required rewriting Kant's word 'transcendental' to allow for a conditions being - transcendental and contingently transcendental. Kant could not possibly allow this adaptation of his term, indeed, 'metaphysically transcendental' is contradictory for him. But I propose the concept be amended to allow as examples of transcendental objects eyeballs and souls: not other people's eyeballs and souls as they appear to me but each person's own eyeballs and souls as used by oneself. nscious'. It follows a fortiori that 'X is conscious cannot yield as a logical consequence 'X is what a person merically identical over time' falls into this second logical category:

- (1) 'I am an object of inner sense'
- (2) 'All time is merely the form of inner sense'
- (3) 'Consequently I refer each and all of my successive determinations to the numerically identical self, and do so throughout time, that is, in the form of my inner intuition of myself'
- (4) "'The personality of the soul" says nothing more than that in the whole time in which I am conscious of myse

TUA2

We have seen that some of the conditions for the possibility of experience which Kant identifies as transcendental are transcendental and to that degree empirical. This is a fact that would have appalled Kant, but one I argue for in the light of contemporary neuroscience.

To further substantiate this I examine Kant's solution to the Hume-Kant problem of the self.

Accepting Hume's point that no 'fixed and abiding self' appears in self-consciousness, Kant concludes from this that

'What has necessarily to be [re]presented as numerically identical cannot be thought as such through empirical data' (A. 107 136).

If we read 'empirical' weakly to mean 'knowable through experience' and not strongly to mean 'knowable through sense experience' then all that prevents this claim being false is its including the term 'necessarily'. Now, it may be that Kant thinks it is a necessary truth that persons are conscious of their numerical identity through time. This is perhaps criterial of being a person. But, if it is, this is partly a matter of stipulation. There could in principle be beings who were just like persons except they were not conscious of their numerical identity over time. I should not wish to deny that they would be deficient in not just that respect but in consequential respects also. But it is logically possible that such beings should be numerically identical and a fortiori numerically continuous over time without their in addition being aware of this fact. (I use words like 'aware' and 'conscious of' just to mean 'know' here). So, the insertion of 'necessarily' by Kant is just some unpacking of what is implicit in his concept of a person. If we abandon that criterion of personhood, then the claim becomes synthetic and, arguably, false. This is because what can be represented as numerically identical can be thought as such through empirical data.

Indeed, Kant has an element of the correct account when he uses 'thought' here: because 'I am my experiences' or 'I am the same being that existed yesterday' are thoughts in the propositional sense, and not experiences.

There is another construal of the claim which emphasises a non-empirical component. This is the notion 'numerically identical'. It could be that Kant conflates this with the modal concept of self-identity. The latter as employed for example in the logical axiom $x (x = x)$ is certainly not an empirical concept. This despite the fact it is obviously persistently metaphysically verified and never metaphysically falsified. But the idea of numerical identity is an empirical one. If we say A is B but A is not C then we are exercising an empirical power to discriminate between A and C and to assimilate A and B. Thus this notion of sameness has a broadly empirical use. Nor does the fact that 'A is B but not C' a thought imply that the concept of numerical identity is not empirical, only that it is indeed a concept.

By drawing attention to this we just point out that arguably no concepts are straightforward empirical items: they are perhaps rules, or dispositions or skills. But we should discriminate non-empirical from empirical concepts by the sort of use they have. Whatever the correct ontology of concepts, 'A is (numerically) B' clearly has an empirical use.

So, what is represented (thought) as numerically identical not only can be but frequently is thought through empirical data.

Nor is there any objection to extending this empirical account to oneself. Just as we may say person A at t_1 is person B at t_2 , so we may add to this 'I am person A'. In other words, substituting the first person pronoun for 'A' or 'B' does no violence to the empirical subject matter of the terms of the identity statement. In all this we have to read 'empirical' weakly not strongly.

So, an extended and amended Humean account is really enough.

Kant though thinks a non-empirical account of a) my numerical identity over time and b) my consciousness of my numerical identity over time is needed. It should be clear now why I think this further quest for conditions is redundant, but it bears examination nonetheless.

Kant thinks there must be a condition which precedes all experience, and which makes experience itself possible (A. 107. 136). The necessary conditions for some phenomenon, in this case experience, will typically turn out to comprise a large set. I have mentioned empirical/transcendental conditions: the soul as a biological condition for experience, optic nerves and operative eyeballs as necessary conditions for seeing in particular. But clearly, that there is, for example, something rather than nothing at all is also a necessary condition for experience.

It is primarily a matter for empirical investigation what the conditions for experience are. The procedure should be: select some candidate condition C, remove C and decide whether experience still exists. Kant's conditions are not testable in this way. He says:

'There can be in us no modes of knowledge no connection or unity of one mode of knowledge with another, without that unity of consciousness which precedes all data of intuitions and by relation to which representation of objects is alone possible' (A. 107. 136)

Clearly the unity of consciousness referred to here is not identical with the unity of consciousness he mentioned at A 106. 135 as requiring a transcendental ground, because he says now: 'This pure unchangeable consciousness I shall name transcendental apperception' (A. 107. 136).

In the first use of 'unity of conscious' this denoted something in need of a ground, in the second it is a ground, in fact the ground of the first mentioned unity. This equivocation on 'unity of consciousness' is unfortunate because it is inconsistent with one transcendental and plausible interpretation of Kant's position. It could be that the empirical and the transcendental unity of consciousness are not numerically distinct,

but one and the same unity described in two ways, viz metaphysically or transcendently. This though is precluded by the principle we mentioned earlier; x's being a condition of y precludes x's being y. Perhaps the difficulty could be avoided in this way: that a transcendental description of the unity of consciousness is possible is a necessary condition of an empirical description of that same unity being possible. The transition from an ontological medium to a distinction between descriptions mitigates the incoherence to some degree. Not entirely though. Kant says transcendental apperception is a kind of intellectual self consciousness: the true thought of my numerical identity over time. This, which is perhaps a disposition or a possibility cannot plausibly be identified with the set of occurrent mental states comprising my unity of consciousness.

Dispositional and occurrent states are distinct (even though occurrences realise dispositions). It is not *prima facie* clear for example how a tendency or a propensity could be numerically identical with an experience where the two sorts of mental state seem so qualitatively dissimilar.

So it is better for Kant to drop his talk of the TUA as a unity of consciousness and replace it with a dispositional account. This is not because occurrent self consciousness does not exist but because it accords better with Kant's later talk of the TUA as a possibility of self-consciousness.

The needs, and the semi-logical, semi-psychological structures Kant invokes to meet them can be accounted for within the empiricist-physiological model I have outlined. The first use of 'unity of consciousness' translates into 'mental field'. The second misleading 'unity of consciousness', which is 'transcendental' - refers to the soul (or perhaps as the account is enriched, to activity of the soul or propensity to soul activity). This translation also retains the attempts to render the two uses of 'unity of consciousness' consistent in terms of different descriptions.

To do much of the cognitive work Kant requires we may abandon 'self-consciousness' and even its 'possibility' but may employ instead a 'member of the mental field'/'physiological item' distinction instead. Thus, colours for example have hue, saturation, shade etc as components of the mental field, but are lightwaves of certain lengths under a physiological description. As items in the mental field they are straightforwardly empirical, as purely physiological they are - transcendental, that is, empirical-transcendent. They are transcendental because they are not qua light waves direct objects of visual perception. They are empirical transcendental because they are indirect objects of observation, in fact theoretical objects of empirical science. Similarly, a mental image may be metaphysically coloured, of the lake district, etc but - transcendentially, metaphysically a soul state.

Added to the ambiguity between one unity of consciousness describable in two ways and two numerically distinct unities of consciousness, one empirical and one transcendental is another, closely related one. Kant draws a distinction between the unity of consciousness (*Einheit des Bewusstseins*) and 'all data of intuition' (A 107. 136). He says the unity of consciousness precedes (...) the data of intuitions. Whatever 'precedes' means it implies not only if not A then not B but if not A then B

then A is not B. For example, if it means that A exists chronologically before B then, just so long as A exists but B does not exist then A is not B. This does not preclude the possibility that if A exists and then if B exists, A and B might be identical in one sense of that term: they might fall under the same sortal. Thus I at an earlier time am identical with (am numerically the same person as) myself at a later time.

The chronological interpretation of the unity of consciousness-data of intuitions relation can be excluded by appeal to the context of Kant's critical epistemology. He is clear that knowledge and experience are not possible without both the categorical contribution of the understanding and a sensory input. This precludes the existence (in the case of persons) of any unity of conscious at anytime chronologically prior to the having of some intuitions. This follows if the existence of the unity of consciousness entails for Kant either having some experiences, or knowing something. If it does not include this then the term is being used vacuously.

But, if this is right, a difficulty arises for the logical reading of 'precedes' too. This is because we have just established that intuitions are in a sense a condition for any unity of consciousness. Kant though at A 107 is asserting the converse: the unity of consciousness is a condition for intuitions. The problem is only insuperable if A's being a condition for B precludes B's being a condition for A (whether or not this turns out to be what Kant intends here). Clearly 'logical condition for' can be a reciprocal relation. (For example, having size might be a condition for having shape and having shape might be a condition for having size, at least in the case of physical objects.)

But the relation 'logically prior to' does not admit of this reciprocity. If A is logically prior to B then it is logically impossible for B to be logically prior to A. If this is what Kant means by 'precedes' here then some further ground needs to be found for saying the unity of consciousness makes possible the data of intuition in at least one way which is not reciprocated. There is more than one such way. Before examining these I shall briefly restate the issue in the vocabulary I have introduced.

Is there a sense in which A is logically prior to B but in which A is not prior to A, where A and B are given these values: experience, experiences; set of mental states, mental states, mental field, results of soul-transformations? In discussing logical priority, one issue must be left completely aside. This is the causal priority of the soul over experiences: the contingent truth that the existence of a soul is a causal condition for there being experiences in the case of human beings.

The priority of the unity of consciousness over all data of intuition can be understood in at least two complementary ways. One involves emphasising that Kant, in these passages, is trying to answer the question How is experience possible? The other is noticing that answering this question requires for him mentioning some transcendental facts, but not empirical facts.

Our knowledge of all data of intuition is, in one sense, empirical: I know I am having some experiences because I am able to exercise inner sense, but knowledge of that prior unity of consciousness which allegedly precedes all data of intuition is not empirical. 'Empirical' has to be read weakly here to mean not just 'knowable through

sense experience' but 'knowable through experience'. In particular Kant wants to emphasise that the transcendental ground of experience cannot be discovered introspection. I think an argument for this conclusion is embedded in the text at A 107. Here it is extracted:

- (1) 'Consciousness of self according to the determinations of our state in inner perception is merely empirical and always changing'
- (2) 'No fixed and abiding self can present itself in the flux of inner appearances'
- (3) 'What has necessarily to be represented as numerically identical cannot be thought as such through empirical data'

Therefore

- (4) 'There must be a condition which precedes all experience, and which makes experience itself possible'

The three premises are thoroughly Humean. The first amounts to: In introspection I am aware only of changing experiences, and knowledge of these is contingent, the second to: in introspection I am not aware of any permanent or continuous self, and the third to: no necessary truth, for example the fact of my self identity, can be an empirical truth.

The conclusion, that there is a non-empirical condition for experience, is facilitated by the remark: 'To render such a transcendental supposition valid [...]' where the transcendental supposition is my self-identity over time. The conclusion at (4) is the conjunction of two claims, the first of which does follow from (1) - (3) on a certain reading of them, the second of which only follows if we supplement Kant's argument with additional premises. The first conjunct of (4) is: 'There must be a condition which precedes all experience'. If we read 'condition' here to denote the condition for ability or my self-identity over time then this does go through if we allow Kant to assume there is such a condition: that my self-identity over time is the sort of truth that admits of having conditions, and secondly if we allow him to assume he has exhausted the empirical and the transcendental possibilities. I shall later deny this second assumption and say there are transcendental and empirical conditions for my self identity over time, but these are not present to introspection.

But if 'transcendental truth' and 'empirical truth' are mutually exclusive predicates, as Kant assumes they are, then the first conjunct of the conclusion follows: there is a condition of my self-identity over time which is not metaphysically discoverable.

If 'transcendental' and 'empirical' are jointly exhaustive of the possibilities, then this condition must be transcendental. We need to read the argument in this highly Kantian way to obtain half its conclusion.

The second conjunct of 4) is that this condition 'makes experience itself possible'. I think this only follows on several new assumptions. So far Kant's premises are only sufficient to yield 'there is a transcendental condition of my self-identity over time', not 'there is a transcendental condition of experience in general' still less, 'they are one and the same condition'. What does Kant need to make the premises yield this?

There are various ways this might be done but one in particular is most in keeping with Kant's intentions. Suppose A is a condition for C. Using this model we can say my self-identity over time is a condition for my experience, but there is a further condition for my self identity over time, *viz* the transcendental unity of apperception so, the transcendental unity of apperception via the transitivity of the relation 'is a condition for' is a condition for experience. This is, I think, what Kant intends. He certainly does not assimilate the TUA to one's self identity over time. These are not the same. The TUA is the non-empirical possibility of my awareness of my self-identity over time. It is a non-empirical disposition of self-conscious thought of a particular restricted sort. As he will later put it, it is the possibility of the 'I think' (*Ich denke* [...]) prefixing any of my thoughts.

It should be clear now why the relation 'precedes' between the transcendental unity of consciousness and all data of intuitions is in at least one sense non-reciprocal. All data of intuitions are experiences, and the transcendental unity of consciousness is precisely a condition for experiences.

There could in an extremely minimal and truncated sense be a sort of entity which would count as an experience if not subsumed under the TUA. Kant for example is mentioning what he takes to be a logical possibility here (and not something that entails a contradiction): 'It would be possible for experiences to crowd in on the soul and yet to be such as would never allow of experience' (A 111). Not only could those experiences not be parts of one and the same experience, they would not be experiences of anyone even considered singularly.

There is a further interpretation which retains the asymmetry between the TUA and the data of intuition. This requires making a distinction between the transcendental unity of apperception and the transcendental unity of consciousness (TUC). This distinction is not well made out by Kant, but there are grounds for drawing it.

If the TUA is a disposition to self consciousness, to have the thought 'I think', then the TUC is what is thus thought as a unity. There is a relation of mutual dependence between the TUA and TUC, but mutual dependence is not identity so they are not identical. Unless the disposition to the use of 'I think [...]' were possible, the TUC would not exist. Reciprocally, unless there were thoughts capable of being prefaced by 'I think' the TUA would not be possible either.

There are even grounds for saying that the TUC does already possess a kind of unity not immediately provided by the TUA. This is the unity of the categories, not the category of 'unity' but the completeness and a priority of the twelve.

Clearly, that we have categories is for Kant a condition of our having experience, and that we have the specific categories we do is a condition for our having the sort

of experience we do. On this second interpretation, the unity of consciousness is the unity of the categories and these 'precede' all data of intuition in a way that is a familiar theme in Kantian epistemology. The categories are prior to experience in that they are not empirical concepts derived from experience but *a priori* concepts intellectually imposed upon it. They also fall under the description 'conditions for experience'.

Finally, these contentions taken together substantiate the ordinary language fact that any experience is somebody's experience. The verb 'experience' takes a grammatical subject and an object. This though should not mislead us into unquestioningly adopting the ontological assumption that experience is necessarily a relationship.

The unity of categories - data of intuition reading is also substantiated by Kant's remarks that

'There can be in us no modes of knowledge, no connection or unity of one mode of knowledge with another' without the TUC as a condition for data of intuition. Nor could there be '(re)presentation of objects.' (A. 107)

On this reading, Kant is reiterating his critical epistemological claim that *a priori* concepts and sensory input are jointly necessary for knowledge, but with a special proviso: The unity of the categories is also a condition, and this unity is logically prior to the possibility of intuitions being experiences. But Kant thinks these conditions of experience themselves need justification.

TUA3

Disposition and Occurrence

There is a deep ambiguity in the concept of the TUA and this needs to be addressed now. Sometimes Kant offers a dispositional account of the TUA but sometimes, less frequently, an occurrent account. The dispositional account is invoked in arguments which purport to establish purely formal or analytic conditions of experience, the occurrent account in arguments which purport to prove the existence of transcendental synthetic *a priori* conditions for experience.

An example of the occurrent view is this: 'This pure original unchangeable consciousness I shall name transcendental apperception.' (A107) This could be understood as a psychological faculty that is perpetually exercised (that is, exercised by a subject at all and only those times when that subject experiences). That would be to read 'unchangeable' to imply 'perennial' or something equivalent. But that is not the most perspicuous reading. The force of this analogy must not be entirely lost:

'The numerical unity of this apperception is thus the *a priori* ground of all concepts, just as the manifoldness of space and time is the *a priori* ground of the intuitions of

sensibility.' (A107)

'Unchangeable' means 'unchangeable in principle'; not the sort of fact that could change if there is experience. This is not just the formal point that I am self identical, not just a special case of $x (x=y)$, but an ontological commitment to a unified consciousness. We could put it this way. The unity of consciousness is a condition for states of consciousness. Indeed, we could develop Kant's analogy further on this own terms. Intuitions are either temporal or else both spatial and temporal. In other words they occupy portions of space-time. But, purportedly distinct spaces and times are in fact parts of one and the same space and time. So, also purportedly distinct experiences of mine are in fact parts of one and the same experience. Here in contrast is an example of the dispositional view:

'This unity of consciousness would be impossible if the mind in knowledge of the manifold could not become conscious of the identity of function whereby it synthetically combines it in one knowledge.' (A 108)

'The unity of consciousness' refers to the TUC not the TUA. The TUA is precisely consciousness 'of the identity of function' mentioned in the same sentence. 'Could' here does not commit Kant to the view that TUA is a kind of occurrent or perpetual of subliminal self consciousness. On the contrary, it is the possibility of being self conscious in a specific sense: being capable of the thought 'I think [...]'.
Contained in this notion of self-knowledge is the possibility of this sort of thought: 'I am numerically identical, now at t_2 with a person having some experiences at an earlier time t_1 '. This is what Kant intends to capture by 'original and necessary consciousness of the identity of the self'. (A 108)

Prima facie there is a problem about the dispositional reading which can be avoided by amending Kant's theory slightly. The difficulty is deciding a referent for 'the mind' in the above passage. If 'the mind' refers to the same faculty as 'original unchangeable consciousness' then the dispositional account is in danger of regressively falling back into the occurrent account. This is because 'mind' would then conceptually imply an existing transcendental ground even for the TUA and the TUC. This is a conclusion Kant would not wish to be committed to. He is on one very plausible reading committed to just that because 'mind' is the subject of the verb 'could' in 'could not become conscious of [...]'. As we have seen, the TUA is a condition for the TUC, so on the principle if A is a condition for B, and B a condition for C, then A is a condition for C, the mind is a condition for the TUA (as subject of it), the TUA is a condition for the TUC so the mind is a condition for the TUC.

Clearly this is unsatisfactory from a Kantian point of view on at least two grounds. It could be read as a case of that very Cartesian reification of the subject he opposes in the Paralogisms chapter. Secondly, if 'mind' is not given a Cartesian analysis it has not here been given a Kantian one. It is in fact left unexplicated by Kant in this passage.

The solution is to analyse 'mind' here in a way consistent with Kant's use of the psychological concepts he has introduced so far. Two routes are open. One is to identify the mind with the totality of functions mentioned in the passage, the other is to preferentially identify it with just one or more of them. The totality analysis is substantiated *prima facie* by something Kant says. The mind becomes conscious of the identity of function whereby it synthetically combines (the manifold) in one knowledge. (A 108) If the mind is that which is conscious of that identity, and that which synthesises the contents of the manifold then the mind cannot be identical with either of those. This is true if A's being the subject of some experiences E1 En precludes A featuring amongst those experiences or being the totality of them. The mind is the subject of three activities: knowledge of the manifold, consciousness of self-identity, and synthesis. The totality analysis would be facilitated if 'being the subject of' does not entail 'being ontologically distinct from'. Then it could be argued that the mind is nothing over and above the exercise of these various functions. It is then synthesising activity, not what synthesises, it is consciousness of self-identity over time, not what is thus conscious, it is knowledge of the manifold, not the possessor of that knowledge.

If Kant does not adopt this or a comparable reductivist solution then he is in the incoherent position of assuming a subject for the TUA: a condition for the condition of all conditions (that is not itself).

There seem to be no criteria for excluding one or more of the various psychological functions from subsumption under 'mind' or preferentially choosing one to be called 'the mind'. If such criteria can be produced there are no Kantian grounds for precluding such an amendment.

There is a further and unKantian solution to the analysis of 'mind'. For this, 'analysis' must not be read as 'explanation of the meaning of the word "X"' but as 'saying what X really is'. Clearly, although these procedures sometimes coincide, this is not necessarily so. In keeping with the psycho-physical identity theory in terms of which I have tried to resolve some Kantian issues, the solution to What is the mind? here is; the mind is the soul. Indeed, it is because we have souls but each person cannot directly sense their own that Kant is misled into smuggling in the mind as the covert subject of experience.

In Kantian terms 'mind' is redundant in the A 107-8 account of the possibility of experience, but if he would consider all metaphysical and, in my sense, transcendental, conditions of experience then the soul is the obvious candidate. It is the transcendental subject of experience.

Kant then proceeds to identify as one and the same function, the consciousness of self-identity and the consciousness of the unity of consciousness:

'The original and necessary consciousness of the identity of the self is thus at the same time a consciousness of an equally necessary unity of the synthesis of all appearances according to concepts.' (A 108)

This identification is maintained despite the use of 'equally' in 'equally necessary'. 'Necessary' here can be read as 'necessary for experience'. A difficulty, however, is that if A and B are equally necessary for C then, implicitly, A is distinct from B. If the identity of the consciousness of unity and the consciousness of identity depend for their being one and the same act on the identity of their object then Kant's view is inconsistent: he would have asserted what his own analysis is precluded.

One remedy is to say that my self-identity over time consists in the unity of my consciousness. This would be either a contingent identification or a constitutional claim, but one that is useful for Kant to make. It is only contingent, if a fact at all, because my identity over time could be or be constituted by something other than the unity of my consciousness (for example by my soul). Nevertheless if we read 'is' here as the 'is' of constitution we can coherently assert that the unity of consciousness is self-identity over time without thereby subscribing to the false view that 'unity of consciousness' and 'self-identity over time' are synonymous.

On another interpretation of the relation between consciousness of identity and consciousness of unity, they are not numerically identical. How we read Kant here depends upon the criteria for the individuation of psychological functions. But 'thus at the same time' could be taken merely to imply that it is not possible for there to be consciousness of one's self identity without consciousness of the unity of consciousness. Or, on a similar reading, the consciousness of the unity of consciousness is the means of being conscious of one's self identity over time. Either of these variants is consistent with the consciousness of identity not being strictly (ie numerically) identical with the consciousness of the unity of consciousness. Rather the relation is now either or both of 'is a necessary condition for' or 'end and means'. Indeed, if A is a necessary condition for B then it follows A is not B (if we leave aside the modal consideration that everything is necessary for itself). Also if B is the means to A as an end it follows that B is not A. Kant gives us no criteria for selecting any of these options in particular. He writes at a certain level of generality.

At A 108-9 there is a suggestion that is consistent with both the necessary condition and the means-end readings. He says the unity of the synthesis of all appearances is 'according to concepts' (A108), and 'that is, according to rules'. Here he is thinking of concepts as rules for making experience intelligible.

The employment of rules which is psychologically unavoidable and a condition for experience has two consequences. One is to make appearances 'necessarily reproducible', and the other is to 'determine an object for their intuition'. (A108) 'Necessarily reproducible' must be a reference to numerically distinct appearances of the same sort, because it would be a mistake to suppose that one and the same appearance could be reproduced at distinct time intervals, unless extra reasons were provided for allowing one and the same appearance to exist intermittently. I read 'determine' here to mean 'make what it is' so, for example, a concept determines an object if and only if it (at least partly) makes it what it is.

The employment of rules, and the resulting component parts of the intelligibility of objects can be read as a necessary condition or as parts of the means by which the

consciousness of unity facilitates the consciousness of identity. The consequences of rule following, and perhaps the rule following itself, are partly constitutive of the unity of consciousness. This activity of subsuming the object of intuition under concepts generates 'the concept of something wherein they (appearances) are necessarily connected' (A108). This something, though Kant does not make this explicit, is the unity of consciousness.

So, the passage as A108 can be read as an explication of what is involved in the expression 'unity of consciousness' as well as a partial account of how we (putatively) come to find objects intelligible. This accords well with his earlier claim that the consciousness of the unity of identity of the self is (in a loose sense) a consciousness of the unity of the synthesis of appearances according to concepts.

Although the necessary condition and the means and ends analyses are mutually consistent, neither, taken singularly or jointly, is consistent with the identity analysis which entails the claim that the consciousness of unity is 'at the same time' the unity of consciousness. Kant provides no criteria for preferring one interpretation over the others.

Kant at A108 introduces a complication into his account which is, I think, a mistake. He says:

'The mind could never think its identity in the manifoldness of its (re)presentations and indeed think this identity a priori, if it did not have before its eyes the identity of its act, whereby it subordinates all synthesis of apprehension to a transcendental unity.' (A108)

There are three errors here. Firstly 'mind' is invoked again in an unanalysed way. Secondly, 'before its eyes' is a metaphor with all the misleading connotations associated with 'inner sense' and criticised in the last chapter. Thirdly, he is now not just claiming that it must be possible to be conscious of one's self-identity through the unity of consciousness but that one must in turn be aware of this very act, or more specifically of the 'identity' of that act of self-conscious thought. 'Identity' is equivocal. It might mean what it meant in 'numerical self-identity'; that it is one and the same capacity for rational self-reflection which is exercised at different times. Or, possibly, by knowing the identity of the act I just know what it is. The first sense would be a specifically Kantian one, the second an ordinary language use of the term.

But, which ever is the correct analysis of 'identity' Kant has introduced a new condition which threatens to become regressive, and said something inconsistent with the explanation of the TUA he has presented so far. The postulation is regressive, because although it enables us (putatively) to know of the existence of the transcendental unity of apperception, and not just know of the need for the TUA, the question newly arises: how was that knowledge possible? and so on. Kant introduces conditions for conditions for conditions, but the regress is halted quite arbitrarily; where he no longer pursues a chain of reasoning to its logical conclusion. The awareness of the exercise of that disposition which is the TUA is blatantly

inconsistent with the idea of the TUA as the ultimate condition for other psychological functions. This is because Kant is committed to the view that 'the mind could not think its identity in the manifoldness' unless it was aware of 'the act whereby it subordinates' synthesis to a transcendental unity'. (A108) So not just this act, but the awareness of this act, is a condition for the mind thinking its identity. But the possibility of the mind's thinking its identity is the TUA. But, allegedly, there are no further psychological conditions of the TUA.

Kant has unduly complicated his account by raising it to a meta-level that is really redundant. All he needs for the argument to go through (accepting his other assumptions for the moment) is that the act of synthesis exist, not that the awareness of the activity of synthesis exist, still less yet another sort of activity described as the act whereby synthesis is 'subordinated to a transcendental unity'. It is very difficult to imagine what such an act could amount to. The 'synthesis of apprehension I' he says is empirical (A108), but the unity it is subordinated to is transcendental (A108), so allegedly there is an act which relates the empirical and the transcendental, and the mind can have this act 'before its eyes'.

Rather than invoke psychological function upon psychological function with dubious coherence Kant should adopt this approach. Suppose, as he does, one wished to make the unity of consciousness a condition of experience, and a condition of self identity over time. Then all that is needed, in Kantian terms, for oneself, the unity of one's consciousness, and experience to be possible is that the objects of intuition are in fact synthesised. There is no additional need to invoke the consciousness of this unity or even the possibility of the consciousness of this unity to explain how that unity is possible. It might (though I should dispute this) be invoked to explain something else: namely how self knowledge is possible. It is not true that the possibility of self-consciousness is a condition for experience. There are no doubt beings who are conscious, in the sense that they have experience, but who are not self conscious. If Kant, as he sometimes suggests, is trying to give an account of how self-conscious beings can experience, of how this is possible, then it becomes analytic that self-consciousness or something very much like it will be invoked amongst the conditions of their experience.

Although Kant is discussing the transcendental conditions for experience the regressive character of the account parallels an analogous regress in the empiricist entailments of his philosophy of mind. For example, he says: 'All representations have, as representation, their object, and can in turn become objects of other representation'. (A108)

This model of experience of experience is subject to the same criticisms levelled at inner sense. Again analogously, the thesis of the intentionality of the mental could be salvaged without resort to the meta-claims.

Kant offers in addition to the possible identification of the TUA and the TUC yet another identification. He says of 'our empirical concepts in general relation to an object' that 'this relation is nothing but the necessary unity of consciousness'. Unless this identification held according to Kant, 'knowledge would be without an object'

and our empirical knowledge would have no 'objective reality'. (A109-10)

So the relation of our concepts to the objects of intuition is the unity of consciousness. There are two complementary ways of understanding this.

Firstly, it is a recapitulation of the mutual dependence of concepts and intuitions which is a tenet of his critical epistemology. That mutual dependence is thus partly constitutive of what he now refers to as the 'necessary unity of consciousness'. Secondly, he says the relation is 'also [...] the synthesis of the manifold'. (A109) This is perfectly consistent, as synthesis is the activity of the subsuming intuitions under concepts.

As well as this though he says the 'pure concept of the transcendental object' (A 109) confers objective reality on our empirical concepts, and that this in turn rest on the 'transcendental law' that all possible objects of experience conform to a priori rules of synthetical unity. I understand the latter to be the categories and the principles as a priori transcendental rules for their application. So unless we self-conscious beings were equipped with the categories, and unless the principles were true, our experience could not be of an objective reality.

He seems to subscribe to the further view that we must possess the concept of an objective reality in order for our experience to be objective. This is the force of 'concept' of the transcendental object which 'cannot contain any determinate intuition'. (A109) I think it is false that we have to believe or even possess the fairly sophisticated idea of our experience being of an objective realm in order for it to be true of us that our experience is in fact objective in that sense. On Kantian terms the categories, especially of substance and causation having a real empirical application is sufficient for our experience being objective. He need not have further recourse to a conception of objectivity. If we do have that conception, or if we believe our experience to be of an objective world then it is a necessary condition of our experience being objective that that conception be accurate, or that belief be true. This is right if there is an analytic connection between experience being objective and being of an objective world, say if no experience is objective unless it is experience of an objective world. It is worth maintaining this analytic connection because it is worth distinguishing objective experience from experience that just seems objective (for example, from the inside). At (A 110) Kant ends this discussion by saying:

'In other words appearances in experience must stand under the conditions of the necessary unity of apperception.' (A 110)

This is a rather loose way of summarising a discussion which has not invoked the necessity of the possibility of awareness of my numerical identity, but just explicated what this identity consists in. 'In other words' should really be deleted unless we subscribe wholly to the occurrent and not at all to the dispositional account of the TUA. Then on the occurrent reading the TUA can be identified with the TUC, and that in turn with the conditions of the objectivity of experience in synthesis.

There is clear sense in which this unity is transcendental. When I am having

some experiences it is true that I am subsuming intuitions under concepts, but this is not a fact that is introspectively available to me. It is not an item within my experience but a condition for and a structure of my experience.

The passage at the end of (A 110) does to an extent mitigate against the assimilation of the TUA to the TUC. There he speaks of the 'synthetic unity of appearances in accordance with concepts' and at (A111) says this is 'based on a transcendental ground of unity'. I assume if A is 'based on' B then A cannot be identical with B. He can be read consistently if we do not identify the synthetic unity of appearances (SUA) with the TUA but with the TUC. The evidence for this identification is extremely strong. He has already mentioned the essential role of synthesis in constituting the necessary unity of consciousness. Now he goes on to draw the same analogy with space and time for the SUA as he did for the TUC:

'When we speak of different experiences we can only refer to the various perceptions, all of which, as such belong to one and the same general experience.' (A 110)

He says the SUA is the 'form' of experience, as were space and time and the 'one single experience' of (A 110) is just the TUC.

Kant's next move seems to me illegitimate. He says that unless the SUA were based on a 'transcendental ground of unity' then synthesis would be 'altogether accidental'. (A 110) On my account, in two important senses synthesis is accidental: It is a contingent fact that there do exist any experiences at all. Secondly, it is a contingent fact that the conditions for experience exist also. On top of this, the best candidate for a transcendental condition of experience is the soul, and this, in a weak sense compatible with its being transcendental, is empirical. The transcendental ground he refers to is the TUA, construed dispositionally.

Kant is misled by a familiar Kantian equivocation on 'necessary'. He assimilates 'necessary for experience' to 'logically necessary' or 'formally necessary' or, here 'non-accidental'. Two tasks should be separated: firstly the quest for conditions for experience and these may pace Kant be empirical, contingent and transcendental, and secondly, the conceptual analysis of concepts like 'experience', 'self', 'consciousness'. Kant persistently assimilates these two tasks. Even though '*p* is necessary for experience' entails '*p* cannot be refuted by experience', this kind of necessity is weaker than logical necessity.

Even if the premise just quoted about the need for a transcendental ground were true, the following conclusion is a *non-sequitur*:

'Otherwise it would be possible for appearances to crowd in upon the soul, and yet be such as would never allow of experience'. (A 111)

The TUA, whether understood as a transcendental unity of consciousness or as a disposition to think the unity of that consciousness, just does not explain the transition from an input of intuitions, to the experience of objects. If anything does

explain this transformation it is the existence of the soul. Indeed, it is rather a condition of there existing a disposition to think 'I think [...]' that there should be something to think about, some propositional content of the thought 'I think [...]'. This content is either explicable partly in terms of experiences or else requires experiences. So, far from the TUA effecting the transition from appearance to experience, it presupposes that that transition is already achieved.

Two further criticisms: Arguably, what stands in need of a transformation into experience should not be called 'appearances'. This is because appearances are grammatically appearances to something or other. Also, they are arguably appearances of something or other. As we have seen, for them to count even as appearances in this sense on Kant's very own terms a great deal of empirical and transcendental psychological machinery has to be invoked.

Second, 'soul' is smuggled in as 'mind' was in the earlier passages. It is left unclear what the soul is upon which the appearances would hypothetically crowd. This much is at least clear: the soul (unlike the mind) cannot possibly be a generic term for the various mental functions in their mutual dependencies. This is because Kant is speculating what, *per impossibile* on his terms, would be the case if these functions were in fact suspended. In order to say something coherent here he needs to make an ontological assumption explicit. For example does 'soul' denote a Cartesian ego, or might it denote the soul? The Paralogisms preclude the first option. I recommend the second.

Despite these defects in its formulation the question, Kant's concept of the TUA is designed to answer is a genuine one. Indeed, at (A111) he defines 'appearances' (in this context) as 'intuition without thought', that is, hypothetically, as a passively received input not subsumed under categories. This is a sensible way of posing the problem, but I should be inclined to say that if there is anything which genuinely is a passively received input not directly experienced then this should be understood as purely physiological: lightwaves, perhaps, or soundwaves. A reformulation of Kant's question would then be: What facilitates the transition from the physiological level of description to the phenomenological level of description? In keeping with the idea of a soul transformation I hold the metaphysical hypothesis that it is the soul which effects this transition.

Thus it can remain true on Kantian lines that perceiving an object constitutes it, that is: partly at least makes it what it is, but this constitution is a transition from a physiological description being uniquely true of the object to a phenomenological description also being true of it. We could call this view 'physicist-idealism' or 'monism'.

Thus we can write Kant's 'crowd in upon the soul' in this way: there could in principle be a physiological input via the physical senses to the soul but no experiences or secondary qualities thereby be caused to exist. Indeed, it remains true in the world of which both physical and psychological descriptions are true that light and soundwaves are 'nothing to' the person who sees and hears in a sense strongly analogous to Kant's. Kant though is thinking of the logical conditions for experience

again, and it is part of the meaning of 'experience' for him that a concept is applied to an intuition in the unity of consciousness.

If, as I maintain, the existence of light waves, soundwaves and objects which resist pressure are conditions for objective experience then this is nonetheless a contingent fact about them. This for three reasons: They are empirical conditions not of any possible experience but of just the sort of experience we do have. Secondly, the sort of experience we do have could, in principle, have been produced by some other set of conditions. Finally, there could, as just noted, have been a world just like ours with only one difference: no psychological or phenomenological descriptions are true of it: only physiological descriptions. Thus, in a third sense, the items falling under a physiological description only contingently fall under the additional one: 'conditions for experience'.

Kant does not leave open the possibility that his question, What are the conditions for a sensory or inner input being experience? It might allow of an empirical answer.

[The first premiss is the assumption of the simplicity of the self. Kant does not criticise this proposition directly at CPR 372-3, presumably because he thinks it vulnerable to all the objections he has amassed so far.]

TUA 4

The conditions for experience are empirical or analytic, or transcendental in the sense used so far. Kant makes the possibility of self-consciousness a condition of experience:

'All possible appearances, as representations belong to the totality of a possible self consciousness.' (A113)

But this is either analytic: All my experiences are mine, or else a non-empirical ontological claim. It is non-empirical because there is no conclusive empirical evidence that any given experiences belong to a single self consciousness. Although 'all my experiences are mine' is analytic, Kant makes the more general claim here that all possible experiences are someone's which is also analytic. It entails the specific claim, if I exist and have experiences.

Now, the claim that all experiences are someone's does not yield the conclusion that all experiences are episodes in a self-consciousness, unless the additional assumption is added that any being which experiences has a capacity for self-consciousness. This assumption is false unless 'capacity for' is read weakly as 'could in principle be' or something equivalent. It is metaphysically possible and probably true that some beings, certain animals, are capable of experience but not self-consciousness. They have experiences and those experiences form parts of a single experience, a single 'life' experience. In Kantian language, the correct formulation should be that all experiences are parts of some consciousness, not some self-

consciousness. But this clearly is not his view.

Kant is led to suppose self-consciousness is a condition for experience partly because of his tacit, and rather Cartesian, emphasis on the first person singular case. It is especially apparent in the later discussion of the 'I think [...]' and at (A113), his canonical formulation is the first person singular claim that all my experiences are mine. If we give this a third person rendering: all her or his experiences are her or his, or it's then there is no implicit reference to self-consciousness. But the use of the token reflexive 'I' means that the true utterance of 'all my experiences are mine' does at least presuppose a capacity for self-reference, even if not self-consciousness. None of this is to deny that Kant has independent arguments for the existence of the TUA. It is a hypothesis about how he might be misled into thinking the grounds 'X experiences' provides for 'X is capable of self-consciousness' are stronger than they are.

'Totality of self consciousness' (A113) is yet a further amendment to the relation between TUA and the TUC. It is a generic name for both in their mutual dependence. He means that unified consciousness which has a disposition to think its own self-identity. Both these notions are included in the claim that 'as self-consciousness is a transcendental representation, numerical identity is inseparable from it (A 113).

Two points about 'representation': This is a case where it is correct to translate *Vorstellung* as 'representation' and not 'presentation', because clearly if the TUA could be presented it would be empirical. If on the other hand we deny this, and say I am presented with myself in self-consciousness, then we force Kant to subscribe to a view manifestly inconsistent with his Humean conclusion that the self is not available to introspection. Secondly, 'representation' is a kind of thinking in this case and not a kind of experiencing. It is the thought of one's own self-identity over time.

The claim that numerical identity is inseparable from self-consciousness is ambiguous. One interpretation can be dealt with straight away. This is: That the axiom of identity is true is a condition for there being self-consciousness. This is right but only in the modal sense that the truth of $x (x=x)$ is a necessary condition for the existence of anything whatsoever. But if we read 'numerically identical' as something like, 'same one at t_2 as at t_1 ' then we need to decide what 'inseparable' means here. In one sense the claim is clearly separable. It is possible to produce many true statements of the format 'A at t_1 is (numerically identical with) A at t_2 ' without any need to invoke the concept of self-consciousness. So, it is clear that in that sense, the doctrines are separable. That only leaves the possibility that self-consciousness requires consciousness of numerical identity over time. This is feasible because if A and B are in some sense inseparable, but we know that A is separable from B, then B must be inseparable from A. How can this be? He is using 'inseparable' as a non-reciprocal relation, or a one-way relation. If A is separable from B but B is not separable from A then if A exists then B may or may not exist but if B exists then A exists. It is thus unfortunate that Kant says 'numerical identity is inseparable from it (self-consciousness)' because what he means is self-consciousness is inseparable from numerical identity.

This is the only remotely plausible interpretation of Kant's intention here. But adopting it leads to difficulties and possibly inconsistency. To begin with, the TUA is the condition of all other conditions of experience. It is at variance with this claim that my continued self-identity over time is presupposed by the TUA. Secondly, the TUA is supposedly constitutive of the TUC, the unity of my consciousness, and this in turn is a condition of my continued existence over time. So an order of priorities has to be decided: Either the TUA makes my numerical identity possible, or my numerical identity makes the TUA possible. A third possibility is that the dependence is in this case reciprocal, but that option, or the second would require Kant's giving up the earlier assumption that the TUA is the ground of all other psychological structures.

There are arguments which could be mounted in defence of each of these options. For example, if the TUA includes my disposition to think my numerical identity, then arguably this is only possible on two conditions: firstly, I have to exist in order to think (a Cartesian assumption) and my self-identity is a modal condition of my existence; secondly, I have to be self-identical over time to be (correctly at least) thought as such. These considerations suggest Kant should adopt the first option and make the TUA's possibility dependent on my self-identity. One objection to this doctrine though is that it rules out *a priori* the possibility of a punctual or instantaneous self-consciousness. Also, in principle, a being could be self-conscious at some time and at some later time, where there further was no intervening time when such a being was not self-conscious. This would not make Kant's numerical identity condition redundant though, because it is a condition of this continuous consciousness being correctly called 'self-consciousness' that it is what it is a consciousness of at t1.

The middle interpretation, that the TUA is a condition for my continued self-identity over time seems only remotely plausible on one construal. This is that it makes possible the use of 'my' or indeed, the first person pronoun 'I'. Other than this, the TUA cannot be what my continued self-identity consists in because it is the (possibility of) consciousness of that identity. It is a principle, part of the grammar of 'conscious of', that if A is conscious of B then neither A nor B can be identical with the consciousness that relates them. (even if, as allegedly in self-consciousness, A and B are not numerically distinct).

So, it is the possibility of a unity of consciousness that can be referred to in the first person which is facilitated by the TUA. Its role in making personal identity possible is to allow us to speak of personal identity. Some other account is needed of what entitles us to speak truly of personal identity.

The final possibility, that the dependence between TUA and 'numerical identity' is two-way should now be clear. The self-identity of that which I am is a condition for the TUA. The personality of that which persists over time is bestowed by the TUA: This is the most sympathetic way of reading Kant here even though it does require weakening the claim that the TUA is the condition of the possibility of all other psychological structures.

Kant says not only that numerical identity is 'inseparable from self-consciousness' but that it (self identity) is a 'a priori certain' (A 113). There are at least three ways of taking this.

Most straightforwardly, but most trivially, it is *a priori* that $x (x=x)$. More contentiously, but more interestingly it is a priori that I am self-identical over time. This could, arguably be inferred from the minimal premise 'I exist' without the need to make any empirical observations. For example it could be maintained that 'I' could be given no sense in the absence of some ground for someone's persistence over time. But there could perhaps be punctual selves or intermittent selves. Even if there were intermittent selves these could be thinkable as discrete episodes constituting one and the same self. This is awkward, but not utterly implausible, rather as someone could attach sense to this: 'I continued the same dream last night that I started the previous night'.

Finally there is a more obscure interpretation. Straight after the claim that self-identity is *a priori* certain, Kant says 'for nothing can come to our knowledge save in terms of this original apperception'. (A 113) 'For' here suggests this sentence is intended as a reason for accepting that self identity is a priori certain. I think it can only function as such on a new interpretation of 'a priori'. This is to be roughly equivalent to the sense of 'a priori' in Kant's talk of 'a priori categories'. Obviously the categories are not *a priori* in just the sense in which judgements or propositions are a priori because categories are not truth value bearers and only truth-value bearers can be the subject of the complex predicate 'knowable to be true or false independently of (sense) experience'. The categories are *a priori* in the sense of not being abstractable from sense experience on any Lockean account, yet imposed psychologically on the contents of experience in rendering it intelligible. Now, 'this original apperception' is a priori in this sense. The TUA is not an empirical faculty, neither in its exercise, nor in its origin. But like the categories it is a condition for experience. These three points taken together are sufficient for its being a priori in the new sense.

On the last interpretation then, my numerical identity is a priori certain because in a sense just so long as I have experiences, that presentations are something 'to me' as Kant puts it, so my numerical identity is guaranteed. The assumption is that there are no un-owned experiences. An experience is an episode in a self-conscious mind, and this can be known *a priori*, given just: at least one experience exists. So when Kant says 'appearances are subject to a priori conditions' (A113) he is not just referring to the categories, but to the TUA and the TUC, including my numerical identity over time.

Kant's conclusion is worth quoting, because it involves him in an inversion of what is really the case:

'All appearances stand in thoroughgoing connection according to necessary laws, and therefore in a transcendental affinity, of which the empirical is a mere consequence.'
(A 113-114)

I call this an 'inversion' because it makes the empirical a consequence of the necessary and transcendental. Now, in principle there may be instances of this order of priorities obtaining. We have mentioned that logical axioms are conditions for anything's existing. Also, an idealist argument could be mounted to the effect that a transcendental subject of consciousness is a necessary condition of any object of experience. So, I should not wish to stipulate a priori that the empirical cannot in principle be a 'mere consequent' of the necessary or the transcendental. In this particular case though Kant needs standing on his head or setting on his feet.

If there is a condition to be singled out as necessary for experience then it is the soul. But *pace* Kant the soul is metaphysical because transcendental. Unless this empirical condition obtained: my soul exists, statements of the forms 'All my experiences are mine' would not be true. This is because the indexicals 'my' and 'mine' require referents. So, although 'All my experiences are mine' is a necessary truth, indeed it is analytic, it is only true on condition it may be formulated, and the conditions for its formulation are empirical. In this sense the necessary is a consequence of the empirical.

Indeed, if the soul is the transcendental subject then in a clear further sense the transcendental is a consequence of the empirical. Souls are empirical objects, but my soul, to me, is in addition a contingently transcendental object.

Kant says (A114) that the reader will find his order or priority - transcendental over empirical 'strange and absurd'. But we should not find it so surprising, as he says, if we consider that 'nature is not a thing in itself but is merely an aggregate of appearances, so many [re]presentations of the mind' (A 114). I think the relation between presentations in the mind, and empirical processes is not best viewed as causal or conditional, this raised the question of priorities, but as ontologically ultimately not distinct. Then we can say this relation is identity. In fact this is quite consistent with Kant's position here. He says nature is 'mere appearance' and 'so many presentations in the mind'. I agree with this just so long as he concedes that the appearances and the presentations in the mind are also nature; that is, empirical objects. The use of 'merely' produces an idealist construal of the identity theory, it suggests two incompatible positions: firstly; A is B, and secondly B does not really exist. Suppose we do not read 'merely' this way and say 'B is nothing over and above A'. If 'is' has any force here then it follows also that 'A is nothing over and above B'. Also, 'is identical with' is a reciprocal relation. If A is identical with B then B is identical with A. Otherwise, 'is' is a mistake and should be replaced by 'is a property of' or some similarly weaker expression.

I conclude again that Kant has not exhausted all the empirical possibilities in his search for conditions of experience.

My evidence that a version of the identity theory is consistent with Kant's view of the empirical - transcendental relation partly rests on his remarks at (A 114). There he isolates 'three subjective sources of knowledge': sense, imagination and apperception. These are subjective because they are psychological faculties of the experiencer, also they are transcendental as *a priori* conditions for experience. But

then Kant, crucially, says this: 'Each of these can be viewed as empirical, namely in its application to given appearances'. (A113) In other words, these three faculties can each be understood under either a transcendental or an empirical description. There is no ontological distinction here. The criteria for choosing between descriptions are epistemological. Each faculty is individuated by its function and each of these has both a transcendental and an empirical function. This possibility of redescription using 'transcendental' or 'empirical' is important for showing that transcendental idealism and materialism are mutually consistent. Kant's considered view is though; that a transcendental description is possible of one these three categories is a necessary condition of any empirical description being true of it. He thinks this because of his assumption that the conditions for experience cannot themselves be either experiences or experienced. But if he leaves room for contingent transcendental conditions, which could in principle be experienced, then there is no residual reason for allocating priority to 'transcendental' over 'empirical' in the description of conditions for experience.

Although empirical apperception is a sort of self-consciousness, it is not numerically distinct from what is known by the name as the 'transcendental unity of apperception'. The latter is the thought of one's self identity, and the possibility of this thought makes possible the unity of consciousness over time. Clearly, the neo-Leibnizian term 'apperception' is intended to subsume a variety of type of self-awareness, ranging from the purely formal possibility of self-reference in the TUA, to the more discursive empirical self-consciousness which is a person's alleged acquaintance with their own mental states. There is no need to think of these as the exercise of distinct faculties though. Indeed, there are strong reasons for reading Kant as saying they are the non-empirical and empirical exercise of one and the same capacity for self-awareness. This rests partly on his view that apperception can be understood either under an empirical or under a transcendental description. But, perhaps more significantly, he speaks not only of the TUA as a condition for experience but of apperception as a condition of experience. At A114 for example there is no mention of transcendental unity, just of apperception. This is a strong ground for identifying empirical apperception with apperception in the TUA. If we deny this then we force Kant into the view that empirical apperception qua empirical is a condition for experience. This seems manifestly false. Although something like empirical apperception is arguably a condition for self-experience, indeed on at least one account that is analytic, it by no means follows that empirical apperception is a condition for all experience. There are no doubt many, perhaps the majority-of-moments when a being perceives an object but is not thereby aware of perceiving the object. I see little textual evidence that Kant subscribed to a perennially occurrent or persistent view of self-awareness, nor indeed that he saw any need to postulate such a phenomenon. On the contrary although he does not use the term, his theory of self-consciousness is dispositional.

If we read Kant this way we can say that empirical apperception is exercised in the awareness of one's own mental states and that this same faculty for reflection is

capable of more formal abstract self-reflection also: the consciousness of one's self-identity over time. The TUA then becomes the possibility of the second sort of reflection. But the TUA, empirical apperception, and in the last resort experience are each explained by reference to one faculty or capacity called 'apperception'.

One further distinction is necessary. At (A116) Kant speaks of pure apperception, and the question arises of whether he is mentioning anything not exhausted by either empirical apperception or the transcendental unity of apperception. He says all perceptions have an a priori ground, and in particular empirical consciousness is grounded in 'pure apperception'. (A 116) Kant gives an explanation of pure apperception: 'that is [...] the thoroughgoing identity of the self in all possible (re) presentations'. This seems a rather different claim from either: The TUA is a condition for experience, or: Apperception is employed in empirical consciousness to facilitate the reproduction and recognition of experiences. He is saying; that I am the persistent subject that endures through the course of my experiences is a condition of my having them. This is not just a reaffirmation of the need for the TUA: the TUA was the possibility of the consciousness of one's self-identity. Pure apperception is that self-identity itself.

So the distinction between Pure Apperception (PA) and the TUA is reasonably well made out.

Less well made out, at least on the face of it, is the distinction between PA and the TUC. The TUC was after all the unity of consciousness made possible by the TUA which in turn was itself a condition for experience. Kant has to show that there is a new role for pure apperception which has not already been exhausted by the TUC. I think it is not desirable to read Kant as advocating in PA a distinct psychological reality from the TUC. On the other hand he is drawing our attention to a new feature of what is presupposed by experience, or redescribing familiar facts with a change of emphasis. He is pointing to the identity of the self in all its intuitions. This is a formal point about my experiences: that they are mine. Hence he calls it 'pure', meaning 'non-empirical', 'apperception'.

'All my experiences are mine' is *a priori* and analytic. The transcendental unity of consciousness is supposed to show how this formal condition is realised in our psychology. Although I have just called a Kantian principle 'analytic' Kant thinks it is a necessary truth with an explanatory role in showing how experience is possible. Formally, all my experiences are mine, but materially or metaphysically every experience I have is an episode in one and the same unified conscious that the TUA enables to be called 'I' or 'mine'. Pure apperception is closely related to the TUA even though it is not it. It is that numerical identity over time which the TUA is the possibility of my consciousness of.

Hence Kant explicitly says PA is the thoroughgoing identity of the self throughout experiences. This seems a much more plausible condition for experience than the TUA. It is arguably a real condition of there being any experience whatsoever, not just self-conscious experience. This though rests on the assumption that there cannot be experience without an experiencer, subject. Although this

assumption is deeply rooted in thought and grammar that does not make it true.

[we can only refer to the various perceptions, all of which, from an input of intuitions, to the experience of objects. If anything does explain this transformation it is the existence of the soul. Indeed, it is rather a condition of there existing a disposition to think 'I think [...]' that there should be something to think about, some propositional content of the thought 'I think [...]'. T]

TUA 5

Another, complementary, way of understanding the TUC/PA relation is this. That the TUC should not obtain is one way, just one way, in which PA would break down. In that important sense the TUC is what my self-identity amounts to: it is the particular way in which the formal condition of experience is realised. It is, so to speak, the cash value of the formal remarks.

This mutual dependence between the TUC and PA is to be taken seriously. My self identity over time is, is constituted by, the unity of consciousness, and the unity of consciousness is my synthesising of intuition, or to put it less misleadingly, is nothing over and above a certain synthesising of intuitions. This amounts to one and the same phenomenon being describable in two ways: by the use of a subjective vocabulary or an objective vocabulary.

I shall now try to substantiate this interpretation by analysing Kant's text at (A116-117). He says 'all perceptions are grounded a priori in pure intuition' and explains this partly by the clause 'that is, in the thoroughgoing identity of the self in all intuitions'. (A116) Part of the rest of what he means is that all perceptions are necessarily temporal, a doctrine he assumes established in the Transcendental Aesthetic. By 'the identity of the self' here he means precisely the TUC. But he says 'empirical consciousness' is in pure apperception' (A116). He means that one's perceptions are what make up the unity of consciousness.

Consciousness, and its unity are nothing over and above (ontologically) a set of perceptions. Thus perceptions may be described subjectively, as had by me, as items in that self-same enduring consciousness which is mine; or objectively, as intentionally directed towards objects, or as having intuitions as their content.

This way of reading Kant is borne out by this passage: 'Intuitions are nothing to us, and do not in the least concern us if they cannot be taken up into consciousness'. (A116) He means that objectivity and subjectivity are mutually dependent, so that the conditions for experience of objects ('something presented') are exactly the same here as the conditions for experience or knowledge: *viz* the unity of experiences in one and the same consciousness.

The identity of this consciousness over time is therefore a condition for the perception of objects. He has not abandoned his further view that consciousness of this self-identity, or the unity of consciousness, is an additional condition for experience. He says:

'We are conscious *a priori* of the complete identity of the self in respect of all (re)presentations which can ever belong to our knowledge, as being a necessary condition of the possibility of all (re)presentations.' (A116)

It must be emphasised that consciousness of the identity of the self not to be confused either with introspective acquaintance with a self, nor with a kind of perennial or enduring self-awareness. Kant intends neither of these things. It is the possibility of the thought of ones self identity that he insists upon and this capacity may be exercised intermittently. This has an important consequence for the understanding of 'transcendental' in TUC.

The unity of consciousness is not itself an item for inner experience. If it were so available it would be merely empirical (and 'accidental') but Kant says the unity of consciousness (in this context) is 'transcendental'. Indeed, he refers at A116 to the 'transcendental principle of the unity of all that is manifold in our representations'. This is yet additional support for the view that the unity of consciousness is not anything distinct from the unity of representations. In fact, I think we could call the manifold 'the unity of (re)presentations' here. As unity of consciousness it is described subjectively, as manifold it is described objectively.

Kant is essentially a philosophical monist despite the plurality of psychological and formal structures he distinguishes (and despite his use of the plural form noumena). I call his philosophy of mind 'materialism' even though what is essential to it, the transcendental ground, is described either mentally or physically and either subjectively or objectively, but is intrinsically neither.

The only realism Kant admits is empirical realism. Within empirical realism, unless physical descriptions hold psychological descriptions cannot hold. This is ensured by the categories and the forms of intuition.

Kant not only speaks of the unity of consciousness as 'transcendental' but also claims 'this unity of the manifold in one subject is synthetic' (A116), and 'this synthetic unity presupposes or includes a synthesis'. The synthesis it includes, or partly is, is the synthesising activity of subsuming intuitions under concepts which produces experience of objects. If we are to identify the self, or that unity of consciousness which I am, with anything at all in these passages it is with synthesising activity. The transcendental unity of apperception is a condition of this synthesis. This is because it is constitutive of the unity of consciousness. At (A118) he says the necessary unity of synthesis is prior to apperception, but this means empirical apperception, not the TUA. There is a clear reason why, on Kantian terms, this should be so. Empirical apperception - my awareness of my own mental states would seem to presuppose a unified self as both the subject and the object of that awareness. This point is quite consistent with his view that there exists also a reciprocal dependence of the unity of consciousness, and thus of synthesis, on the TUA.

There is another sense in which 'the synthesis of the manifold in imagination' is 'transcendental'. (A118) We are not conscious of its operations. We know that on

Kant's account intuitions are subsumed under concepts and categories, but this process is not introspectively available to us. This synthesis is therefore not (even weakly) empirical but transcendental. It is also *a priori* in the sense that it can be known to operate without our having to make observations (per impossible) of its operations. Similarly it is 'a priori necessary in relation to the original unity of apperception' (A118) because it partly constitutes that unity. It is part of what that unity is.

There is then a very close link between original apperception and the categories. We should suspect this because synthesis is partly the applying of categories, and synthesising activity is what the unity of consciousness amounts to ontologically. Kant spells out the relation in this remark:

'The unity of apperception in relation to the synthesis of imagination is the understanding.' (A119) (Kant's italics)

The understanding is the faculty which applies categories correctly not dialectically, the synthesis of imagination is the subsuming of intuitions under categories, and the unity of apperception is the self-aware consciousness with which the understanding is identical. Put another way, the unity of my consciousness is partly the twelve categories. This reinforces the claim that the unity of consciousness is a condition for experience. If it variously 'is', 'includes' or 'presupposes' the categories, then it follows that it is a condition of experience, because it has been argued independently by Kant that the categories are themselves conditions for experience.

Similarly for 'the understanding': Kant says, 'All appearances, as data, for a possible experience, are subject to this understanding'. (A119) In other words, the categories must be applicable to them for them to count as experienced at all. This is the force of 'appearances have a necessary relation to the understanding'. (A119)

Kant says with reference to the unity of apperception 'this same unity with reference to the transcendental synthesis of the imagination (is) the pure understanding'. (A119) Now, this makes it clear that Kant is not introducing two sets of faculties, one psychological and one transcendental, but that one and the same set may be understood under either description. In other words the pure understanding does not exist in ontological addition to the understanding, nor is transcendental synthesis numerically distinct from empirical synthesis. On the contrary, either can be understood either as a condition for experience (transcendentally) or as a mental structure at work in producing experience (psychologically). This again, I think, emphasises the monism of Kant's approach to the mental.

I shall now try to go into a little more detail about this monism as a synthesis of subjective and objective. 'Subjective' and 'objective' are not well-defined terms in Kant's philosophy. *Prima facie*, for example at (A121-2), that which is subjective pertains to that which experiences, the owner of experiences, and that which is objective pertains to that which is experienced, the intentional object or content of experience, or what experience is of. I say 'prima facie' here because for a

combination of semantic, epistemological and ontological reasons the distinction is not in the last resort a primitive one for Kant. At (A 121) he explains the necessity for a reproductive faculty of imagination to 'connect' or 'reproduce' experiences so as to be of objects. His next point is:

'But it is clear that even this apprehension of the manifold would not by itself produce an image and a connection of the impressions were it not that there exists a subjective ground which leads the mind to reinstate a preceding perception alongside the subsequent perception [...] which it has passed and so to form whole series of perceptions.' (A121)

'Mind', I take it, is here just a short-hand term for synthesising activity, and perhaps what presupposes psychologically. The crucial expression is 'subjective ground'. 'Ground' (*Grund*) is a spatial metaphor, so we need not take it literally that the reproductive power of imagination 'rests' on some subjective entity. But if there is a ground of some process, then there is a condition of that process. This means, there is at least in principle an explanation of how that process is possible. Kant thinks there is such an explanation available, in terms of subjective rules.

The importance of rules in the use of the imagination is this. The exercise of a reproductive and connective faculty of imagination in perception is not in itself sufficient for experience of a world of enduring physical objects entering into causal interaction. There must be some constraints on what sort of presentation is to stand in relation to any other sort of representation. Additionally, there must be some constraint on what such relations can be; 'part of', 'before' or 'simultaneous with' and so on are possible candidates. The categories, the principles and the schematism deal together with the construction of our experience at this transcendental level and this sort of rule following is designed to explain how empirical judgements are possible. Rather as the schemata and principles show how the application of the categories is possible, so the rules of reproduction show how the application of empirical concepts is possible. If this were not so then Kant would not call the subjective ground 'empirical':

'This subjective and empirical ground of reproduction according to rules is what is called the association of (re)presentations.' (A121)

So the unity of association is identical with the subjective ground and the imagination's conformity to rules. These are three ways of describing the same condition for experience.

Kant's next move is to claim that this in turn (I shall call it 'the subjective ground') stands in need of justification, in fact it requires an objective ground. Kant actually says the unity of association (the subjective ground) requires an objective ground otherwise there would be nothing to guarantee that our acquaintance with appearances would constitute knowledge. He means that without some further

qualification there is nothing to distinguish imagination from perception, or, to put it another way, there is nothing to guarantee that our experiences are of objects. The objective ground will enable our experiences to be not merely states of the subject, but externalist intentional states, that is states directed towards mind-independent objects.

So, what is the objective ground? He says: 'This objective ground of all association of appearances I entitle their affinity'. (A122) It is clear that the objective ground, like its 'subjective counterpart is a capacity for rule following. It too is a possibility of following rules in reproducing experiences in their 'thoroughgoing connection'. (A122) This ground though is *a priori* [and can be apprehended prior to 'all exercise of laws of the . (A122)] Although this objective ground is a capacity for rule following, it is not one which we have the option not to exercise if we have some experiences. He says it 'constrains us to regard all appearances as data of the senses' (A122). He clearly does not mean that we have to mistake what we imagine for what we perceive but that we are automatically prevented from mistaking what we perceive for what we merely imagine. This view is perfectly consistent with his theory that the reproductive imagination is active within perception.

There are clear differences between the subjective and objective grounds. Apart from his calling one 'subjective' and the other 'objective', the first is 'empirical' (A121), the second is 'a priori' and 'antecedent' to the empirical laws of the imagination. This clearly entails the two grounds are not numerically identical. This textual evidence is quite inconsistent with any claim that the grounds are one and the same. However at (A122) Kant says this about the objective ground:

'This objective ground [...] is nowhere to be found save in the principle of the unity of apperception.' (A122)

The unity of apperception referred to here is the TUA. Empirical apperception is ruled out because the objective ground is 'prior' to the empirical, but all appearances must 'conform' to the TUA. The upshot is that neither the subjective nor the objective grounds could exist but for the TUA.

If the relation between the two is not identity then, the connection between them is nevertheless extremely close. In fact all appearances can be thought of as falling into one of two classes; those with a subjective association, or those with an objective affinity, where these two categories are mutually exclusive and jointly exhaustive. That the subjective and objective grounds are numerically distinct precludes an appearance being thought under both. For such an appearance to be experienced at all requires it being thought under one, but only one of the two.

This is part of what I mean when I say Kant's philosophy of mind is a monism. He tries to do justice to psychological dualisms here between subjectivity and objectivity without collapsing one into the other. The TUA makes the distinction between subjectivity and objectivity possible.

At (A122) Kant says 'synthetic unity is [...] objectively necessary'. This can be

taken two ways. He could be just reiterating the point that the unity of apperception is a condition for objectivity, in the sense of 'objectivity' which stands in semantic contrast with 'subjectivity'. But there is nothing to preclude another reading in which 'objective' is used in the sense of 'what is the case'. Then we can say the TUA is 'necessary' for both subjectivity and objectivity. It is objective in a sense that subsumes the two old senses: it makes them possible. This interpretation is borne out by this passage:

'The objective unity of all empirical consciousness in one consciousness, that of empirical apperception, is thus the necessary condition of all possible perception.'
(A123)

'All possible perception' includes, say, both inner and outer sense, but the first occurrence of 'objective' above does not denote the experience of objects. It is predicated of that which makes both inner and outer experience possible. It is a new sort of objectivity. This in turn is borne out by its making possible all appearances 'near or remote' (A123). I take it this spatial metaphor can be unpacked to mean 'subjective or objective', or 'inner or outer'. In either case, the TUA is 'objective' in a sense that is not contrasted with 'subjective'.

The reciprocal dependence of subjective and objective grounds is facilitated not only by the TUA but by the categories. At A126 Kant describes the categories as 'subjective grounds of [...] unity contained a priori in the original cognitive powers of the mind' but then says 'these subjective conditions [...] (are) at the same time objectively valid'. Although the subjective ground in the sense of the association of experiences and the objective ground in the sense of the affinity of appearances are not numerically identical, one and the same ground or set of conditions for experience is describable using the semantically contrasting adjective subjective and objective. Here 'objective' should be read as 'pertaining to the object' or 'belonging to that which experience is of, or intentionally directed towards'. Clearly Kant can then hold this view consistently, just so long as 'pertaining to' and 'belonging to' are not prefaced by 'only' or 'just'.

There is no reason to commit Kant to the strong thesis than there are exclusively and irreducibly subjective or objective structures of experience. In fact we should expect the opposite as subjectivity and objectivity are each made possible by the TUA. There could not be a 'unity of nature' without a 'unity of consciousness' and vice versa. The existence and application of the categories is essential to each.

This synthesis of subjectivity and objectivity in the categories is consistent with Kant's critical epistemology. For example at A125 he claims: '[...] the order and regularity in the appearances, which we entitle nature, we ourselves introduce' but reciprocally; 'only by means of these fundamental concepts (the categories) can appearances belong to knowledge or even to consciousness, and so to ourselves' (A125). We should not understand the manifold of appearances or the unity of apperception as having any sort of existence logically or chronologically prior to the

existence and exercise of the categories. It is not Kant's view that the categories unite or, to use a spatial metaphor, bring together the subjectivity of consciousness and the objectivity or nature. Rather this distinction between subjective and objective descriptions presupposes the prior application of the categories. In a sense terms like 'subjectivity' and 'objectivity' are abstractions from the unity of experience. I think this is the force of the following passage, which is a partial definition of 'categories':

'certain concepts which render possible the formal unity of experience and therewith all objective validity (truth) of empirical knowledge.' (A125)

The fact that the categories make possible the formal unity of experience needs to be borne in mind in distinguishing the TUA from pure apperception. We have seen that the TUA is a disposition to think one's self-identity over time. Pure apperception is, in a sense, that identity itself, or at least, what that identity consists in, how it is realised. This is why Kant says:

'The abiding and unchanging "I" (pure apperception) forms the correlate of all our (re)presentations in so far as it is possible that we should become conscious of them. All consciousness belongs as truly to an all comprehensive pure apperception.'
(A123)

Kant is not saying that there must be occurrent consciousness of what 'I' denotes whenever there is experience. This would be inconsistent perhaps with his endorsement of Humean scepticism about the self and certainly with the purely dispositional account of the TUA. It would make the TUA into a disposition that was always exercised so long as there was occurrent experience. Pure apperception, which is exactly the same as the abiding and unchanging 'I' as the above passage makes clear, just is the unity of consciousness facilitated by the categories at the same time as the unity of the manifold thought as nature. The consciousness of this unity of consciousness requires the TUA or something very much like it. Also, the existence of the unity of consciousness - pure apperception - would not be possible according to Kant unless the TUA obtained. If it were in principle impossible to form judgements of the form 'I think that p' then neither would that formal unity of consciousness called pure apperception be possible. The TUA is a formal condition of the unity of consciousness, but it is the use of the categories which materially or psychologically enables the unity of consciousness to exist. Indeed, the following could be read as suggesting that that is just what the use of the categories is: the uniting of the manifold of intuition with the unity of consciousness:

'[...] concepts which belong to the understanding are brought into play through relation of the manifold to the unity of apperception.' (A124)

But, as the argument above should have established, it is not possible for either the

unity consciousness or the manifold to exist without the employment of the categories. This interpretation is substantiated by Kant's remark at A129 that the source of the categories is not 'the object' nor 'the self'. On the contrary they are the ground of that very distinction.

There is a temptation, which should be resisted, to understand the categories as merely subjective when Kant says they are an a priori and not an empirical source of the law abidingness of nature. Similarly, there is a temptation to think of them as purely objective when he insists they have only an empirical use. Instead we should think of the understanding as the condition of the possibility of both. So although it is 'the lawgiver of nature' (A126) this nature 'is only possible in the unity of apperception' (A127) where this unity is both a unity of experiences and a unity of appearances. This is what he means when he says 'the unity of apperception is thus the transcendental ground of the necessary conformity to law of all categories'. It makes possible a unity that is describable in two ways, or has both subjective and objective properties; subjectively it is a unity of experience, objectively a unity of appearance.

[the soul, and this, in a weak sense compatible with its being transcendental,]

TUA 6

TD in B Version

At first Kant seems to retract this position in the opening pages of the Transcendental Deduction in the B version. For example at (B130) he says the 'manifold of (re)presentation' is nothing more than 'the mode in which the subject is affected', and says we cannot represent to ourselves anything as combined in the object which we have not ourselves combined. It seems as though Kant is producing a subjectivist rewriting of the theory in the A version. This subjectivism is seemingly confirmed by his description of the combination of presentations as 'an act of the self-activity of the subject, it cannot be executed save by the subject itself' (B130). This though would be a misreading.

There are two complementary reasons for Kant's shift of emphasis here. Firstly, it is part of his critical epistemology to adopt a broadly Cartesian starting point. I mean by this that his model of the individual's psychology is fundamentally first person singular and not third person plural in orientation. Secondly, we have seen that one and the same set of structures, all made possible in the end by the TUA, may be described in a subjective or an objective vocabulary. Here he is deploying the subjective one. There is nothing in that to preclude the adoption of the objective one on other occasions which, as we shall see, he often does. That he has not initiated a subjectivist departure from the A version is I think confirmed by his saying there must be a 'unity which precedes a priori all concepts of combination' and this, he warns us, must not be confused with the category of unity.

Section 16 of the B version of the TD is called The Original Synthetic Unity of

Apperception, and I shall take this as another name for the TUA. That it is nothing distinct from the TUA should become apparent in its analysis. Kant defines it in this famous passage:

'It must be possible for the "I think" to accompany all my (re)presentations; for otherwise something would be (re)presented in me which could not be thought at all, and that is equivalent to saying that the (re)presentation would be impossible, or at least would be nothing to me.' (B 132)

For any thought 'p' it must be possible, if it is to be true that p is a possible thought of mine, that p could be prefaced by 'I think' to form the complex thought 'I think that p '. Kant's use of 'possible' here makes it clear that the TUA includes a disposition to self-consciousness. Kant is not thereby committed to the rather implausible doctrine that all my thoughts have the grammatical form 'I think that P'.

There are two Kantian reasons why the 'I think' principle is needed. The first concerns the unity of consciousness. If it were in principle impossible for any of my (putative) thoughts to be preceded by 'I think' then no sense could be attached to their being 'mine' at all. Secondly, it is a tenet of Kantian epistemology that the passive reception of intuitions is by itself insufficient for experience and knowledge. Any such intuitions must be thought, that is, subsumed under the categories, to count as items of experience and knowledge. I have to thereby be aware or conscious of my sensory input for it to be part of my experience. He is not saying I have to be conscious of the fact that I am experiencing, while I am experiencing, in order to experience, but I must have the capacity to do precisely that in order for some experience to be mine. That capacity is a condition of the unity of consciousness which, in turn, is a condition for experience.

There is an interesting equivocation in the passage quoted above about the idea of an unthought presentation. The first view is that unthought presentations are impossible, the second is that they are possible but not items of experience (they are 'nothing to me' as Kant puts it). The more plausible of the two interpretations is the first.

Qua presentations, presentations are presentations to someone or other, necessarily. It would therefore not make sense to speak of presentations being nothing to someone. Of course, we could give an empirical sense to this: colloquially a person is presented with something they find unintelligible, but that is not what Kant means. He entertains, on the second view, the possibility of there existing presentations which are not items of experience, but it is more consistent with the rest of his epistemology to think of this as an absolute or logical impossibility.

Certainly Kant's claim that 'that (re)presentation which can be given prior to all thought is entitled intuition' (B132) should not be taken to imply that unthought intuitions are possible items of experience. We should not give 'prior to' any chronological reading. Rather, he is claiming that intuitions are a condition of thought. He is emphasising one side of the thought/intuition reciprocal dependence.

We should understand 'prior to' as implying both that intuitions are not thoughts and that intuitions supply the 'occasion' of thoughts. Reception is a necessary condition of thinking, but 'prior to' does not imply that intuitions could be either known or experienced without thought.

Finally the equivocation between 'impossible' and 'nothing to' can be made intelligible in terms of the idea of a soul transformation I introduced earlier.

To do this we have to regard the passively received input of intuitions in a rather unKantian way: as purely physiological. These intuitions, not on Kant's view but on my re-writing, are light waves, sound waves etc, but not colours or sounds phenomenologically described. For Kant the transition from intuition to object of experience is effected by thought. On my view the transition from intuition *qua* physiologically described process to event phenomenologically described is effected by the soul. Thus we are faced with a similar dilemma to Kant. There could in principle be a world physiologically identical to our own in which, for example photons bombarded the rods and cones of the retina, but in which no phenomenological colours existed. Should we then say colours are 'impossible' or just 'nothing to' the subject (here, again, physiologically described)?

It is, I think a matter of conceptual stipulation rather than factual discovery. If colours are necessarily phenomenological, and if intuitions are necessarily objects of experience then we should use the 'impossible' formulation. If some sense can be attached to 'colour only describable in the vocabulary of physics' or 'intuition that is not experienced' then we should adopt the 'nothing to' idiom.

Substituting 'physiological input' for 'intuitions' retains two other features of Kant's account. If it is true that there are no unthought (Kantian) intuitions, no uninterpreted intuitions, no intuitions passively sensed in their bare particularity, then this is exactly true also of physiological processes such as light and sound waves. Light waves for example have to contact the retina, an electrical message pass along the optic nerve, and patterns of synaptic firing occur in the soul in order for colour to be perceived. The soul is active (a spontaneity) but the senses passive (a receptivity) in the production of (phenomenological) colour.

In interpreting the passages at (B132) it is important to retain the distinction between the TUA and PA, otherwise the dispositional nature of the TUA will be lost sight of. Kant says:

'All the manifold of intuition has [...] a necessary relation to the 'I think' in the same subject in which this manifold is found.' (B 132)

There is a psychological and a logical way of reading this. Logically, 'I think my thoughts' is analytic. Psychologically we need to read 'necessary' in roughly the sense of 'necessary for' or 'necessary condition for'. Then we can say that the thought of a unitary subject, the very same unitary subject who is receiver of the manifold, is a necessary condition for any intuition being an item of experience.

Both interpretations are complementary, and Kant would not wish perhaps to

make such a sharp difference between them as I have done. Some passages suggest that in his understanding of 'necessary', 'could not be otherwise' collapses into 'could not conceivably be otherwise for us self-conscious subjects'.

In the above passage, the expression 'I think' is a direct reference to PA, and only an indirect reference to the TUA. This is because it is true of every thought I have that I think it, but it is only true of some that I have the meta-thought 'I think that P'. This distinction needs insisting upon. If I am thinking, or for Kant in general if I am experiencing then 'I think' is true of me but its being true of me does not entail its being explicitly formulated in a meta-act of psychological self-judgement.

That last requirement is what the TUA consists in: the possibility of at least the intermittent thought of the identity of one's consciousness. I call the 'I think' above an indirect reference to the TUA just because the TUA makes the PA possible.

There is perhaps an additional justification. There are certainly not two 'I thinks', just two different sorts of exercise of one and the same capacity. We could call one 'conscious' and the other 'self-conscious'. But clearly, from the fact that I am thinking it does not follow that I am thinking that I am thinking, or conscious that I am thinking. One Kant's view though, it must be possible to be thus self-conscious, otherwise no sense could be attached to a unitary I which thinks. Indeed, there would be no guarantee of a unity of consciousness at all.

Kant calls the 'I think' a 'spontaneity' (B132) consistently with his view that thought is active but sensibility passive. His remarks at B132 makes it quite clear that he is not using the 'I think' to refer directly to the TUA at this point but to the PA. He says: about the 'I think':

'I call it pure apperception to distinguish it from empirical apperception or again original apperception.' (B 132)

where pure apperception (PA) 'is that self-consciousness which while generating the representation 'I think' [...] cannot itself be accompanied by any further representation'. (B 123)

So PA generates the 'I think' of the TUA, but it is itself a self-consciousness. Kant's use of 'generates' here should not lead us to suppose he has abandoned his view that the TUA makes the PA possible. Indeed, he says 'The unity of his apperception I likewise entitle the transcendental unity of self-consciousness' (B132). The subject of 'this apperception' is 'pure apperception' so it looks at this point and though Kant is collapsing the PA into the TUA after all. This though is not the most consistent reading to give the text. He is identifying the unity of PA with the TUA. This is a rather loose way of saying that the unity of PA is produced by the TUA. Kant wishes to emphasis the reciprocal dependence of the TUA and PA, and indeed say they are logico-psychological structures of one subject, but in doing this he is tempted to say PA and the TUA are identical. This is a lapse which he should avoid. The reason for this is that even if it is true that A is a necessary and a sufficient condition for B, and B is a necessary and a sufficient condition for A, it does not

logically follow that A and B are numerically identical, though it may well follow that A and B can be coherently thought of as two aspects of one whole. It is the latter relation that Kant needs for the TUA and PA, not the assimilation of one faculty to the other.

Two clear statements at (B132) show Kant's considered view that the TUA and PA are not the identical structure. He says he distinguishes 'pure apperception' from 'original apperception', which I take to be the TUA. Secondly, although the I think has a necessary relation to all my intuitions, he says the I think is a representation which must be capable of accompanying all other representations (B132). If X is capable of accompanying Y it does not follow that it must always do so. If it can sometimes not do so then X is not Y.

While the I think of PA is an occurrent and perpetual condition of experience, the I think of TUA is a dispositional condition of experience which is intermittent in its exercise. It does not follow from this that in any other sense there are two sorts of 'I think'. On the contrary, the unitary subject thinks his/her thoughts, but amongst these is capable of one sort of thought in particular: self-conscious thought of the sort 'I think that P'. It is because of this capacity (the TUA) that Kant calls PA a 'self-consciousness' at B132, not because he thinks that persons are occurrently self-conscious at all times that they are occurrently conscious. A self-conscious being for Kant is a being with a capacity for self-consciousness.

There is another reason why PA is a kind of self-consciousness. Only self-conscious beings can use the language of the self: first person grammatical forms. Here Kant makes the point about the first person plural personal pronoun:

'[...] the manifold (re)presentations which are given in an intuition would not be one and all my (re)presentations, if they did not belong to one self-consciousness.'
(B132) (My italics).

There is a question which Kant does not address here which is whether certain language forms make self-consciousness possible, rather than, as he assumes, the reverse. We still want to know what 'This is a hand' fails to say which 'this is my hand' does say. It could be that the possibility of self-reference is a condition of self-perception. I mean perception of oneself qua the very same person who perceives. It is at least an open possibility that language enables that person who perceives to make reference to him or herself qua perceiver. Otherwise there is a problem about how consciousness of oneself, of that very person who one is, ever becomes self-consciousness: consciousness of oneself *qua* that very same person who is conscious. Arguably the difference is one of description. A person may think of themselves under the description 'this person', or 'this person who I am' where the latter entails, 'this person who is thinking/perceiving etc this person'. If something like this is right then self-consciousness depends on descriptions, or the possibility of descriptions, and so self-consciousness depends in an obvious sense on language.

[In fact all 'objective' above does not denote the experience of objects. It is

predicated of]

TUA 7

Given Kant's commitment to conceptual idealism, it is odd that he offers no positive account of the role of language or categories in self-consciousness. 'Self' for example is not a category. Of course the categories are misused in self-description as he points out in the Paralogisms, but even their restriction to understanding the phenomenal self does not explain the role of language in producing self-consciousness. The nearest Kant approaches to such an account is in the analysis of the 'I think' itself.

'I think' denotes a propositional attitude. We can locate distinct propositions within the 'that' clause which it prefaces: 'I think that P, that Q, that R etc. But we still want to know what the function of the indexical 'I' is. Kant has no explanation of this, but to some extent the 'I think' must be regarded as primitive in his explanation of self-consciousness. It is the possibility of self-reference which is a condition of a unitary consciousness. So, to that degree there is a linguistic condition of self-consciousness. As no explanation is offered of how the I think is possible (and none for Kant is needed) we cannot finally adjudicate in the priority dispute between language and self awareness. Some clarification of how the word 'my' has a use can be gleaned from this passage:

'As my (re)presentations (even if I am not conscious of them as such) they must conform to the condition under which alone they can stand together in one universal self-consciousness, because otherwise they would not all without exception belong to me.' (B132-3)

The unity of consciousness and what we might call the 'reflexivity' of consciousness are each a necessary condition for the use of 'my' in contexts like 'my experiences'. There has to be a unified single enduring consciousness in order for it to be truly called 'mine'. Its being called 'mine' by me presupposes further its awareness of itself, if only in the minimal sense of its ability to formulate judgements of the form 'I think'.

These two conditions, or something very much like them seem to be singularly necessary but not jointly sufficient for 'my' and other reflexive indexicals to have a use. There are all the conditions adduced by Wittgenstein: public criteria for self-ascriptions to have sense, a 'post' within the grammar, the possibility of correctness and incorrectness, the impossibility of utterly incommunicable meaning. These would at least have to be considered in explaining 'my'.

I think the nearest Kant comes to describing sufficient conditions for the use of the first person singular pronoun is at B132 when he says that the 'I think' must be framed in 'the same subject in which the manifold is found' (B132). So, the person who says 'I think that P' is the very same person who thinks that P, or the person who can say 'this experience is mine' is the very same person who has that experience.

This captures the idea of a person referring to his/herself qua that person who refers, rather than just *qua* some person.

A further useful point emerges from the passage at (B132-3) quoted above. The clause within parenthesis; 'even if I am not conscious of them as such' is revealing of two main trends in Kant's thinking on the self. Firstly, it makes it absolutely clear that it is possible for my experiences to be mine without my being occurrently conscious of them as such. Secondly, their being mine results from the fact that they can (Kant's italics) be parts of one and the same self-consciousness. It is the disposition to think self-identity over time that allows us to speak of self-consciousness at all. That in turn makes possible first person grammatical usage.

Kant's precise formulation, which includes 'as such' leaves open the logical possibility that I be conscious of my experiences but not qua mine. There is one clear sense in which this is impossible: I could not mistake my occurrent experiences for someone else. But it is conceivable on Kant's view that there should be consciousness of states of a mind by that mind without that mind having terms like 'I' and 'mine' or 'my' to formulate descriptions under which to make such states intelligible.

This apparent possibility cannot be real if it is true that language, specifically first person grammatical forms or some analogue of them, is amongst the conditions for self-consciousness. Nevertheless, there is prima-facie a mental parallel to the case where x looks at x's body, but fails to realise it is the very same body as his, that is; of x qua observer. This possibility need not entail that x thinks it is someone else's body either. To make sense of Kant's supposition we have to imagine a being who could formulate judgements like 'there obtain occurrent mental states' on the grounds of direct acquaintance with them but who could not formulate 'these mental states are mine' or similar judgements. He has no capacity to make first person psychological ascriptions. I leave aside the issue of whether such a hypothetical being should be called 'a person'.

Because the TUA is partly a linguistic capacity, the 'I think is an expression which allows first person psychological ascription to be formulated, the safest construal of Kant's position is that self-awareness and self-reference are each necessary for each other's possibility, at least in the case of human beings. At (B132) Kant emphasises that empirical apperception is not sufficient for the unity of consciousness. This is the point of:

'[...] the empirical consciousness which accompanies different (re)presentations is itself diverse and without relation to the identity of the subject.' (B132)

The use of 'without relation to' is perhaps unfortunate because it is his considered view that all presentations, to count as such, must be presentations precisely to some unitary subject. What he is saying is that there should exist putative presentations is not in itself sufficient to prove the existence of a unitary subject. So from the fact that there exists a series of presentations, P,... P_n it does not follow that these are episodes

in a single mind. Also, from the fact that there exists a series of meta-mental acts, MP,... MP each of which is an act of awareness of P,... P in a one-one mapping it does not follow that either P,... P or MP,... MP are episodes in a single mind. So although it is in fact true, if they are episodes in a single mind, that the

'thoroughgoing identity of the apperception of a manifold which is given in intuition contains a synthesis of (re)presentations.' (B133)

it is not in virtue of that that we are entitled to speak of a unitary self. That is

'possible only through consciousness of that synthesis.' (B133)

For this argument to work, 'consciousness of' must mean 'direct awareness of'. If it means anything like 'consciousness that' then it fails. This is because there could be consciousness that a series of presentations P,...P_n occurred without its being true that any of them were one's own.

If however in mental acts there is direct access to the occurrence of those acts, consciousness of their occurrence, then in that sense of 'consciousness of' they are 'mine', or belong to a kind of unitary subject at least in principle capable of using 'my', 'mine' etc. Kant says the relation to a unitary subject comes about in this way:

'[...] not simply through my accompanying each (re)presentation with consciousness, but only in so far as I conjoin one (re)presentation with another and am conscious of the synthesis of them.' (B133)

The relation thus established is episodic. Unless such acts of connection were possible we could not speak of a unitary consciousness at all, nor, for that matter, could we speak of an objective unity of the manifold, or a unity of nature. Kant oscillates between a strong and a weak thesis here. The strong thesis is that the identity of the self over time and the unity of consciousness at any one time are made possible by the consciousness of synthesis. The weak thesis is that the awareness of self-identity over time, and the unity of consciousness at any one time is made possible by the consciousness of synthesis.

The weak thesis, on the assumption that terms like 'self-consciousness' and 'synthesis' have a genuine psychological use, is in danger of becoming trivial: I have to be aware of my self-identity for it to be known by me.

The strong thesis though is in danger of being false, and this for a specific reason. Kant tends to invoke 'I', 'my' and 'the subject' in a way that threatens to be viciously circular. Most glaringly he says at (B133) that the identity of apperception contains a synthesis but that this is possible only through consciousness of that synthesis. Or, again, a little later he says the identity of the subject can only come about in so far as I conjoin one presentation with another. There is a severe danger here of invoking precisely what is in need of explanation: The possibility of my self-

consciousness is a condition of my existence as a unitary subject. One, more coherent, formulation is this:

'Only in so far therefore as I can unite a manifold of given (re)presentations in one consciousness, is it possible for me to (re)present to myself the identity of consciousness in [ie. throughout] these representations.' (B133)

What is coherent about this is that it expresses a version of the weak thesis. The uniting of the manifold of representations makes it possible for me to represent to myself the identity of consciousness. It does not, on the weak view, constitute or create or even make possible that very identity of consciousness. This is more plausible than the strong view because Kant has already spent a great deal of time explaining than a quite distinct psychological structure has that role, viz the TUA.

Even on the weak, more coherent, account 'I' is still invoked in a mildly circular way. Kant says 'I' can unite the manifold in 'one consciousness but this one consciousness just is the "I"'. The circularity is only avoided by making this an account of my awareness of my self-identity not of my making my own self-identity possible. 'I am conscious of myself' is coherent. Unless a modal point about identity, 'I am a condition for myself' is not.

There is one strategy, which Kant is aware of, which not only makes the strong thesis internally coherent but also exhibits it as consistent with the weak thesis. This is to do with what it makes sense to say about the self. What we could call the grammar of the self imposes a constraint on the ways in which we can meaningfully talk about the cluster of issues: self-identity, self-consciousness, unity of consciousness. This comes out in the following passage:

'The thought that the (re)presentations given in intuition one and all belong to me is [...] equivalent to the thought that I unite them in one self-consciousness' (B134)

This is not just the reciprocal dependence of subjective and objective descriptions of one unity: as a unity of consciousness, or as a unity of nature, although it is partly that. It is also the point that 'I' can only take on meaning if a self-not self distinction is presupposed. The difference between 'what my self-identity consists in' and 'the possibility of my consciousness of it' tends to collapse because the one is semantically impossible without the other when thoroughly thought through. For example, the thought of my self identity as thinker presupposes the thought of it, if it is to be a topic of discourse at all.

By saying the thoughts are equivalent we should not take Kant to be implying that my self identity and my consciousness of it are actually one and the same: they are not identical. But he is implying that talk of one semantically presupposes the possibility of talk of the other.

This framework of rules governing talk about the self cannot be easily abdicated. It is the grammar within which the philosophical issues are formulated. Kant,

however, has a particular account of why the two thoughts should be equivalent. This enables us to read 'equivalent' as 'logically equivalent' but not as 'semantically equivalent'. P and Q are logically equivalent if and only if P is true if and only if Q is true. P and Q are semantically equivalent if and only if P and Q are logically equivalent and P and Q are synonymous. Kant's claim is that it is true that 'the (re)presentations given in intuition all belong to me' is true if and only if 'I unite them in one self-consciousness is true'. It does not follow that the two statements are thereby synonymous, and Kant does not think they are. In Kant's example, though, it does seem that the two statements share truth conditions. A Tarski type objection could be brought against this contention. For example, 'representation given in intuition all belong to me' is true if and only if the representations given in intuition all belong to me, and 'I unite them in one self-consciousness' is true if and only if I unite them in one self-consciousness. But Kant's view is that one and the same possibility makes both statements true, *viz* that I 'can at least so unite them'. (B 134) The meaning of the self-ascription claim, and its semantic divergence from the claim that I unite my representations is parasitic upon (made possible by) the fact that I can so unite them. The existence of that possibility is the truth condition of both propositions.

We lack clear criteria for synonymy, but that *p* and *q* share truth conditions is not in itself sufficient for their synonymy *p* and *q* may for example have different uses. There are other contrasts tow between Kant's examples. For example in the second I am depicted as active in uniting my presentations, in the first, the claim that they 'belong' to me is neutral with regard to my activity or passivity towards my presentations.

Not only are the two statements not synonymous but 'the thought that I unite [my representations] in one self-consciousness or can so unite them' (B134) is explicitly distinguished by Kant from the 'consciousness of the synthesis of the representations'. (B134) The difference here is between the actual occurrent awareness of synthesis, which he thinks is possible, and on the other hand the thought that I do in fact unite my presentation in synthesis. It is a distinction between the act of self-consciousness which reveals synthesis, and the thought that synthesis occurs.

Despite this distinction, both the consciousness of synthesis and the thought that synthesis occur 'presuppose(s) the possibility of that synthesis'. (B134) This last claim is reasonably uncontentious. If it is true that I am conscious that P then it follows that possibly *p*, just by the grammar of 'conscious that'; rather as, if it is true that I know that *p* then it follows that *p*. Although Kant can be read as drawing our attention to certain rules which govern the use of 'I', 'mine' and other egocentric indexicals, there is a point being made in the philosophy of mind also. The ontological cash value of the grammar is made explicit in this passage:

'[...] only in so far as I can grasp the manifold of the representations in one consciousness, do I call them one and all mine. For otherwise I should have as many coloured and diverse a self as I have representations of which I am conscious to

myself.' (B 134)

Two preconditions of 'mine' having a use are mentioned here. The first is that I, as subject, am intellectually active. This is not spelled out but is implicit in 'I can grasp'. Whatever the complex account of what I am I am irreducibly active in synthesising presentations. This is true even if in the last resort I am ontologically 'nothing over and above' such synthesising activity. The second and more explicit point is that the unity of consciousness is a condition for 'mine's' use. If something is mine then it belongs to something or other. Some persisting particular or process must exist as the owner or possessor of what is mine. The candidate for this role here is the unity of consciousness. A series of presentations must be episodes within one and the same consciousness for them to count as my presentations.

The hypothetical alternative which Kant entertains above is not a possibility on his own account, we could not speak of a self if a series of presentations were utterly 'many coloured' and 'diverse'. By this he does not just mean presentations which differ qualitatively one from another. That they should so differ is not a sufficient condition of their not being mine. He means that unless there exist presentations which are connected one with another (united) in synthesis, and unless the possibility of the awareness of this synthesis existed, then the ontological preconditions for 'my' 'mine' having a use would not obtain. It would not just be true that 'mine' etc would not make sense, but the word could not be produced. Indeed, no determinate thought would be possible at all unless there existed a unitary thinker. This is why Kant says:

'Synthetic unity of the manifold of intuitions as generated a priori, is thus the ground of the identity of apperception itself which precedes a priori all my determinate thought.' (B 134)

'My' is italicised by Kant. If we read the passage with this emphasis he is reiterating the point that there are no un-owned thoughts. But he is also saying there could be no meaningful (determinate) thought whatsoever unless his account of the unity of consciousness were correct, and 'determinate thought' includes but is not exhausted by egocentric indexical thought.

I take it the force of '*a priori*' here is at least that if 'there exists determinate thought' is true then it is a priori that the identity of apperception obtains. It is also his view, though, that the synthetic unity of the manifold is generated a priori. This means that such unity is an a priori condition of experience, one that 'precedes' experience in the sense of by no means being metaphysically abstractable from experience, but remaining a precondition for experience.

That the unity of synthesis is *a priori* in this sense is substantiated by his remark that 'combination does not [...] lie in the objects'. (B134) This means that I cannot passively perceive objects already constituted by the 'combination' (and 'reproduction') of presentations, transcendently speaking. Rather, there being objects of experience at all for me depends upon the possibility of the unity of

presentations in synthesis, and the possibility of consciousness of this process.

In the footnote to (B 134) Kant says 'this faculty of apperception is the understanding itself' (B 134 fn), where in the same paragraph 'apperception' is to be understood as 'the synthetic unity of apperception'. In interpreting this we need to take the term 'faculty' seriously. Then rather than rather misleadingly identifying the understanding with the synthetic unity of apperception we can say that synthetic unity involves the exercise of a capacity, *viz* the synthesising of presentations in combination and reproduction. It is the understanding which is at work in the exercise of that capacity. This is the best interpretation because it is consistent with what he has claimed so far about synthetic unity and the understanding, and it is borne out by this passage:

'the understanding [...] is nothing but the faculty of combining *a priori* and of bringing the manifold of given (re)presentations under the unity of apperception.'
(B134)

Using the 'faculty' reading we can say the understanding is not exactly the same as the synthetic unity of apperception but a capacity within it. Here Kant is speaking in a special sort of psychological idiom, rather than making purely formal claims. This special idiom is, in his view, appropriate to answering the question 'How is the experience possible?'. The best way to describe this idiom is 'transcendental'. It equivocates, as we saw earlier, between the two notions of necessity: 'necessary condition of our experience' and 'could not (logically or metaphysically) be otherwise'. Kant is aware of this equivocation, as is evident from this remark:

'This principle of the necessary unity of apperception is itself [...] an identical and therefore analytic proposition. Nevertheless it reveals the necessity of a synthesis of the manifold given in intuition, without which the thoroughgoing identity of self-consciousness cannot be thought.' (B135)

'All my thoughts are mine' is analytic and necessary. Although not all necessary truths are analytic on Kant's view, those 'thought through identity' are. But Kant does not think the proposition is vacuous. Although analytic it draws our attention to a necessary fact about our experience. Indeed, its very formulation is conditional upon this fact obtaining: synthesis of the manifold and the possibility of consciousness of that synthesis must be possible. This is what he means by saying the identity of consciousness could not otherwise be thought.

It is clear from this that Kant is prepared to blur the distinction, between purely formal and transcendental truths. The analytic statement 'I think my thoughts' is possible because it is true. Put another way, it can be formulated just on condition it is true, and its truth conditions are not utterly internal to it, but consist in certain transcendental facts about us as persons (self-conscious beings).

Kant tries to clarify the distinction between the intellectual and the experiential

components of self-consciousness. He says that 'through the "I" as simple representation, nothing manifold is given' (B135). He is not implying by 'simple representation' that there is introspective awareness of a simple self by denying that one experiences oneself as a 'manifold'. He is denying experience of a self tout court. He allows that I may be aware of my mental states, and that I may be aware that I am a unitary consciousness. What is not possible for me is intuition of a single persistent item within experience denoted by 'I'.

He says 'intuition' is 'distinct from' the 'I' (B135), meaning at least two things. Firstly, he is distinguishing me from my perceptions. Although perceiver and perceived are ultimately made possible by a single set of transcendental conditions, notably the TUA, Kant wishes at an ordinary language level to preserve the distinctions between perceiver, perception and perceived. Then we can read 'I and intuitions are distinct' as 'the perceiver is not the perception'. Secondly, he can be read as reiterating the point that I am not an item amongst my own intuitions, but that which has them.

Despite his claim for ontological conditions of self-identity and psychological self-ascriptions, it remains true that there exist grammatical conditions for self-consciousness. As well as the first person singular form of the 'I think', there is the claim made in this passage:

'I am conscious of the self as identical in respect of the manifold of (re)presentations that are given to me in a intuition, because I call them one and all my (re)presentations, and so apprehend them as constituting one intuition.' (B135)

That 'my' has a rule governed use is a necessary condition of my ability to think my self-identity over time. This, or something very much like it, must be right if we accept Kant's thesis that my self-identity over time cannot be an object of intuition. If I am not continually (or even intermittently) aware of myself as an enduring mental particular, yet can and do think of myself as a persisting psychological self, then a non-empirical explanation of how such thought is possible is necessary. He has already provided an account of the ontological conditions of this thought's possibility, but he is right in thinking there exist also grammatical ones. Otherwise, there would not obtain the distinction between thinking of that person who in fact one is, and on the other hand thinking of oneself qua that person who one is. It is logically possible that there should be beings who were capable of forming sentences like 'This mind is in state 0', and even 'This is the very same mind in state 0 that was in state at an earlier time'. Even so, some additional thought would be had, some additional information added, if such a creature were also capable of formulating 'I am this mind' or 'This mind is mine'.

What this extra token reflexivity consists in, I think, is the ability to think that the very same mind that formulates this judgement, is the mind that this judgement is about. Kant does not explain the force of 'my' in his presentation of the grammatical condition but the above account is at least consistent with what he says.

Kant claims that 'This amounts to saying that I am conscious to myself a priori of a necessary synthesis of representations' (B134) and adds that all (re)presentations must 'stand under this condition if they are to be given to me'. The 'this amounts to' remark is intended to mark the mutual dependency between the ontological and the grammatical conditions for my consciousness of my self-identity. The consciousness of synthesis is a priori quite straightforwardly because the thought that I am self-identical over time and that I synthesise my experiences as a condition of this could not be known to be true through empirical observations. His saying the thought is a priori is consistent with my non-empirical interpretation above.

second.

[Despite these defects in its formulation the]

TUA 8

At B137 Kant spells out more precisely the sense in which combination within the original synthetic unity of apperception is a necessary condition for experience.

Without such combination presentations which are putatively mine 'would not have in common the act of apperception 'I think'. (B137) Unless all my experiences were parts of, or episodes in, one and the same experience they could not be truly called 'mine'. As he puts it 'they could not be apprehended together in one self-consciousness'. (B137) Part of the force of this is that we could not use 'self-consciousness' unless a unity of consciousness obtained. Further, the putative act of empirical apperception whereby I am aware of my own mental states would not be an epistemological possibility either. Indeed, not only are the grammatical and the epistemological possibilities facilitated by original combination, but, arguably, they are mutual conditions of each other's possibility also. Unless I could use 'my' or an analogous first person indexical, I could not be self-conscious. If I could not be self-conscious I could have no use for 'I', 'my' and 'mine'.

The unity of consciousness is not only a necessary condition for intuition, but of understanding. This is because the original unity is a synthesis of subjective and object aspects of experience (as argued above). As Kant puts it:

'all unification or representations demands unity of consciousness in the synthesis of them. Consequently, it is the unity of consciousness that alone constitutes the relation of (re)presentations to an object.' (B137)

To see why this amounts to the unity of consciousness being a condition for the understanding, we need to note that the categories only have an empirical use, and that the understanding is the 'faculty of knowledge'. (B137) Both empirical use, and 'knowledge' require an objective employment: an application to objects. This is why he says 'upon it ('unity of consciousness') rests the very possibility of the understanding'. (B137)

It would be wrong to assume that Kant was insisting upon utterly or irreducibly

subjective conditions of experience of the physical world. Rather the very distinction between subjectivity and objectivity is made possible by the unity of consciousness. This is why he says 'the unity of this act (of synthesis) is at the same time the unity of consciousness' (B137).

In fact Kant uses the word 'objective' to describe the unity of consciousness (at B138 for example). This is a new sense of 'objective', not the same sense in which physical objects are 'objective' but one which makes the subjective - objective distinction possible.

One other way of taking 'objective' here is possible. He says:

'The synthetic unity of consciousness is [...] an objective condition of all knowledge. It is not merely a condition that I myself require in knowing an object but is a condition under which every intuition must stand in order to become an object for me.' (B 138)

On one reading 'objective' is used to point out that the unity of consciousness makes possible an objective world in the sense of one that is other, or not part of my psyche. On that reading 'not pertaining to my mental state, but a possible object of experience for me' is roughly what 'objective' means. Although consistent with the first reading it is not exactly the same as it because this second sense of 'objective' stands in semantic contrast with 'subjective'. Both theses are in Kant's text but the former is more profound. The second can be construed as either a part of it or as a consequence of it.

The insight upon which the necessity of the unity of consciousness rests, is a grammatical one. This is that there can be no use for first person ascriptions unless there is a use for third (or perhaps second) person ascriptions. But on Kant's view, the epistemological distinction between what I am psychologically and what is 'external' to me is possible only through synthesis within the original unity of consciousness.

Kant says that the proposition that 'in the absence of synthesis the manifold would not be united in one consciousness' (B138) is analytic, and 'says not more than':

'all my representation in any given intuition must be subject to that condition under which alone I can ascribe them to the identical self as my representations.' (B138)

so the second formulation must be analytic too: if P is analytic and P says no more than Q then Q is analytic or entails something analytic. Kant does not explain 'says no more than' but I take it that if P says no more than Q then Q is at least a translation of P. That leaves it open that if P says no more than Q, Q may say more than P. Q says at least P. In this particular case, the second formulation can be construed as analytic, if it is construed formally to mean 'all my experiences are mine'. But, to say that P is analytic and that Q is analytic, and to say that Q is a translation of P is to imply that

the conjunction of P and Q is a logical truth. Now, even if Kant's account of the transcendental conditions of experience is correct, it is not a logical truth that it is correct. So it is not analytic either, and so, nor can the conjunction of the two conditions be a logical truth either. This is because no logical truth can be the conjunction of a contingent truth and a logical truth.

There could in principle be alternative accounts of the ownership of experiences. On one view, if it is analytic that Q, then it is just a mistake to look for truth conditions for Q by devising an ontology of the self. On another view there is no such thing as the unity of consciousness, just numerically and qualitatively distinct states of one and the same soul. So, Kant really has to choose between formal conditions and ontological conditions for self-conscious experience. The former will turn out to be analytic, if they involve analysis of concepts like 'self' and 'experience'. The latter though will be contingent, and even empirical despite being conditions for experience. Despite being putatively transcendental (in Kant's sense) there is not one condition which could not be replaced by a rival candidate.

This still leaves open the question of whether Kant's initial claim, that in the absence of synthesis the manifold would not be united in one consciousness is analytic, is true. This claim is analytic if we accept Kant's definition of 'the manifold united in one consciousness' as just being what synthesis is. If we read the claim as analytic we have to abandon any merely causal thesis, for example that synthesis causes the unity of consciousness, but this is acceptable because 'cause' for Kant should not have any transcendental use. There are though two reservations about allowing the claim as analytic. One is that if P is analytic then P is a necessary truth. Kant is thereby committed to the view that synthesis is necessary for the unity of the manifold in consciousness is a necessary truth. But, it is logically possible that some faculty other than synthesis should be a condition for that unity. He says synthesis is active, but if a truth it is not a necessary truth that a person's finding their experience intelligible is an active process. Indeed, possibly experiences being episodes in a single soul is what makes such unity possible.

So, even if Kant's theory is correct, its logical status is not that of a tautology, nor even of a necessary truth if 'necessarily true' here means 'could not be false' or 'could not be true'.

The other reservation is that if the claim is analytic then its truth should be decidable by conceptual analysis of its subject and predicate parts. But, it is not clear that someone who understood 'synthesis' thoroughly would thereby be acquainted with the concepts of the unity of the manifold in a single consciousness, as 'synthesis' has been defined so far. At least this is no more feasible than that a person should acquire the concept of 12 by thinking the unit of 7 and 5: an example Kant says is synthetic.

But, again, if we are prepared to accept Kant's stipulative definition of 'synthesis' then this second reservation may be waived. The whole complex: "'without synthesis the unity of the manifold in consciousness would not be possible" is analytic', would then be an example of the translatability between subjective and objective

descriptions discussed above.

I have argued that the most consistent view to be extracted from the Transcendental Deduction, in both versions, is that the TUA makes possible both subjective and objective descriptions one and the same set of processes. Indeed, all other psychological distinctions have the TUA as their transcendental 'ground'.

There is a passage at B139 which appears *prima facie* inconsistent with that interpretation. Kant says the TUA is 'entitled objective' and must be distinguished from the subjective unity of consciousness, which is a determination of inner sense. (B139) This now sounds as though 'TUA' is the name of an 'objective' faculty which stands in need of semantic contrast with the concept of a 'subjective' faculty.

This appearance is misleading. Certainly the TUA must be distinguished from the subjective unity of consciousness, but so also must it be distinguished from the empirical unity of apperception. The empirical unity of apperception is the objective correlate of the subjective unity of consciousness, just as outer sense is the objective correlate of inner sense. The TUA makes possible both the subjective unity of consciousness and the empirical unity of apperception. This is quite consistent with the view that the TUA 'is that unity through which all the manifold given in an intuition is united in a concept of the object' (B139). This is the claim that the TUA is a condition for the objectivity of experience, that is, for experiences being of objects external to the subject's psychology.

All experience, inner or outer bears a necessary relation to the 'I think'. It is a logical condition of their being experience at all that it be someone's experience. There are no 'un-owned' experiences. For these reason he says 'only the original unity is 'objectively valid' (B140).

One final ambiguity needs to be eradicated. At (B 140) Kant says the empirical unity of apperception has 'only subjective validity'. This is yet another sense of 'subjective', quite distinct from those used so far. 'Subjective' means here is 'relative', or 'in some degree varying from one person to another'. Kant means that the content of empirical apperception is contingent and shifting, but the I think of the TUA is formal and unchanging. That he means 'relative' (or 'relative to a subject') by 'subjective' here is substantiated by the example he gives. He says: 'To one man, for instance, a certain word suggests one thing, another, some other thing' (B140). So the unity of consciousness's empirical content, in each case, is not what makes it a unity. Some consciousnesses' have just that empirical content is not what bestows unity on it. Some other content would do.

So the consciousness of each person is numerically distinct from every other yet this is not in virtue of such consciousnesses containing either similar or different thoughts. As a matter of contingent fact, persons do think different sorts of thoughts if 'what is given' to any consciousness is not 'necessarily and universally valid'. (B140) Kant's view is that even if (as is not the case) each person though the same sort of thoughts, it would not be in virtue of that qualitative similarity that each person's consciousness was a unity. Conversely, from the fact (which obtains) that some persons think qualitatively dissimilar thoughts (at any one time) it does not

follow that each does not possess a unitary mind. This emphasises that the unity of consciousness rests not on empirical but on transcendental ground.

There is also another sort of objectivity which is made possible by the TUA. We have seen that the TUA makes objective experience, experience of objective particulars, possible. It also allows objective judgements, judgements about objective particulars, to be made. I shall read 'judgement' to mean 'propositional attitude' and sometimes just 'proposition', but note that for Kant judgements have psychological as well as logical properties. For example, he says

'a judgement is nothing but the manner in which given modes of knowledge are brought to the objective unity of apperception'. (B141)

So judgements are episodes in minds, but they also have truth values. If there is at least one judgement then there is at least one judge. If there is at least one judgement then that judgement is either true or false. We judge correctly or incorrectly.

Now, there are two separate issues at work in Kant's text that need to be separated out. The first is: What allows judgements to be truth valued? The second is: What enables judgements to be about items other than mental states of the subject (judge)? Kant assumes that by answering the second question he has also answered the first, but on one reading this is not so. This is because a judgement may be either true or false if made about the subjects mental state, or either true or false if made about some other item. The assimilation results from Kant's analysis of 'is'. He says 'is' is employed to distinguish the objective unity of given representations from the subjective. (B 142) Now, the fact that judgements have the logical form 'X is F' does not in itself guarantee that 'x' denotes a particular independent of the mind of the subject. On the contrary, 'x is F' may be the logical form of some psychological self-ascription. If 'is' expresses any sort of objectivity here it is a sense of 'objective' closely related to 'true'. On that reading, the claim that 'x is F' is just the claim that 'x is F' is true. But, as we just noted, the truth of 'x is F' is a distinct notion from the referent of 'x'.

To make Kant's account more plausible we need to take seriously the claim that the 'is' 'indicates their (judgements') relation to original apperception and its necessary unity' (B142). How is this done?

There are two senses in which the TUA makes objective judgement possible. Firstly, the TUA is a condition of any judgement. If a judgement is 'nothing but' the way in which knowledge is brought to the 'objective unity of apperception', then unless the TUA obtained judgement could not exist under that description. Secondly, because the TUA enables the distinction to be made between subjectivity and objectivity, it enables further a distinction to be drawn between two sorts of judgement: judgements of the form 'x is F' and judgements of the form 'it seems to me that x is F'. Kant thinks that the onus is on him to establish the possibility of judgements of the first form when *prima facie* a person's experience consists of a series of presentations. This is the importance of the claim at B142 that the cupola 'is' is used to distinguish between the 'subjective' and the 'objective' unity of

presentations.

Kant thinks we need a guarantee that the series of experiences each unitary mind has is of a persisting world of mind-independent objects. At B142 he is only prepared to define 'judgement' as an 'objectively valid' relation:

'a judgement [...] is a relation which is objectively valid and so can be adequately distinguished from a relation of the same (re)presentations that would have only subjective validity.' (B 142)

If we read 'objectively valid' as partly meaning 'true' then it does have a role. This because 'x is F' is true if and only if x is F. x's being F, as opposed to just seeming F, is made possible by the categories and the principles (B142).

In the above passage Kant excludes a relation of representations with only subjective validity from the class of judgements. 'It seems to me that X is F' is such a relation. But, such relations must obtain, or at least be thinkable for 'x is F' to have sense. It is obviously not a condition of the truth of 'It seem to me that x is F' that 'x is F' is true, but it is clearly a condition of 'x is F' having a use that x could in principle seem to some subject to be F.

The TUA is therefore a condition of judgements about presentations, as it is a condition for the having of those presentations.

The example Kant chooses to illustrate the dependence of the objectivity of judgement on the TUA substantiates the above readings of 'subjective' and 'objective'. His subjective example is 'If I support a body I feel an impression of weight'. The objective example is 'This body is heavy'. Even the subjective example presupposes the possibility of making an objective judgement, viz, 'This is a body' but that does not vitiate the contrast between 'It seems to me that x is F' and 'x is F'.

The objectivity of the judgement 'x is F' is to be understood in this way:

'What we are asserting is that they [two (re)presentations] are combined in the object, no matter what the state of the subject may be.' (B 142).

where this combination is not just a Humean constant conjunction. After all, two presentations could be repeatedly conjoined in the subject's experience without any objective (mind independent) particular being thereby presented.

Two final points about judgement: In the above passage the term 'subject' must be read as 'subject's psychology'. Kant tends to neglect the human body in his discussions of the self but a person might make an objective judgement (mind independent) about his own body which was not uninfluenced by the (physical) state of the subject. We could not then accept 'no matter what the state of the subject' because a physical state of the subject would provide its truth conditions. If we read 'subject' purely psychologically though, this difficulty is avoided.

The other point is that Kant has still not clearly distinguished, on the one hand, between the question of what enables propositions to be true or false, and, on the

other hand, what enables them to be about mind dependent or mind-independent particulars. This leaves open the question of how a certain sort of proposition (Kant might not wish to call them 'judgements') is possible, namely, first person psychological ascriptions, ie judgements of the form 'x is F' or 'x seems to me to be F' where 'x' denotes some occurrent mental state of the subject (speaker). This would reveal Kant's view on the sense of such ascriptions and on their putative incorrigibility. He does not pursue the matter in the Transcendental Deduction.

Although the conclusion of Kant's argument at (B142) is that the TUA makes judgement possible, there are indications at (B 143-4) that judgements, and the categories which are used in their formulation, are themselves essential conditions for the unity of consciousness. They are not more fundamental than the TUA, but clearly if A is the ground of, that is makes possible, B and C, that in no way precludes B making C possible also. We would just have to stipulate that A and B are jointly necessary for C but that neither is singularly sufficient.

He begins the argument by emphasising that the TUA remains a condition for the unity of consciousness:

'The manifold given in sensible intuition is necessarily subject to the original synthetic unity of apperception, because in no other way is the unity of intuition possible.' (B143)

But then he claims that the logical function of judgement is to bring the content of the manifold under 'one apperception'. (B 143) It is the subsuming of the manifold under categories in the acts of judgement which ensures the manifold is 'brought into one consciousness' (B143). The process thus described, though the word is not used at B143 is synthesis.

So, although the TUA is analytic, the ontological cash value of 'unity of self-consciousness' consists in the application of categories to intuitions in synthesis. So although 'All my experiences are mine' is a necessary truth, there is a psychological story about the relationship between me and my experiences. Their combination and reproduction according to the synthesis consciousness consists in. This is why Kant can say:

'A manifold, contained in an intuition I call mine is (re)presented, by means of the synthesis of the understanding, as belonging to the necessary unity of self-consciousness.' (B 144)

Synthesis is now mentioned, and connection with the TUA made explicit. Synthesis allows the manifold to be presented as mine, even though its being mine, and its thinkability as mine both depend on the TUA. Synthesis, we can think of not as a formal but as a psychological or ontological condition of my unitary self-consciousness.

Despite this Kant does not allow us to forget the TUA is the ultimate ground of

each of these distinctions. All empirical consciousness depends upon the possibility of formal self-consciousness, the thinkability of my self-identity over time:

'the empirical consciousness of a given manifold in a single intuition is subject to a pure self-consciousness *a priori*.' (B 144)

There is one remark, at (B145-6), which could be misinterpreted. He says 'our understanding [...] can produce a priori unity of apperception'. This, if Kant is to be consistent, can be a reference to the TUA only if we retain his view that the TUA is the ultimate transcendental ground. The TUA cannot be ultimately dependent on the categories because if A is the ultimate ground of B then B cannot make A possible. The answer is, I think, that there is only one unity of consciousness but it can be understood under an empirical or a transcendental description. As empirical consciousness the understanding is essential to its possibility. But as transcendental its possibility is essential to the understanding. This proviso would have to be made: It is the understanding which 'brings' the manifold of intuition under the unity of apperception.

So, in an important sense, the 'I think' is only thinkable if there is some experiential or intellectual content to be denoted by the that clause in 'I think that p'. There must be synthesis of intuitions for them to be thinkable as mine, and the application of the categories to those intuitions single set of transcendental conditions, notably the TUA, Kant wishes at an ordinary language level to [pr y'] in his presentation of this grammatical condition.

[The Soul as Transcendental Ego THIS SECTION TO END OF BOOK

It is possible to believe the proposition: I am numerically identical at a later time with a being at an earlier time despite discontinuity in the history of the consciousness of that being between t1 and t2. But, is it possible for me to know such a fact just through being acquainted with mental states or, less question beggingly, by the obtaining of some mental states as objects of thought?

There is a need to invoke an inner perceiver or transcendental subject of self awareness. There is need to invoke such an entity or structure to account for one's knowledge of one's self-identity over time. A Humean ontology of the self as a bundle of perceptions is inadequate to explain the idea of self. What we need to do is to explain how the bundle can be 'mine' (someone's) without being perceived by an inner observer as owner. Instead of 'bundle of perception' I shall speak of 'set of mental states'. Two sub-questions arise: what makes them a set? and what makes this set mine, rather than, say, yours? It is the same facts that answer both questions. These facts are ultimately metaphysical, and just in that respect, not in Kant's sense, transcendental.

They amount to this: the transcendental self is the soul. I am only willing to call the soul a - transcendental condition of experience in a minimal and fairly unKantian

sense. It includes the sense of 'transcendental' as 'what makes knowledge or experience possible'. It is logically possible that a Berkeleyan idealism should be true where experience makes the soul possible, and experience is grounded elsewhere, in God and the soul perhaps. I put forward the - transcendental self is the soul as an empirical hypothesis, a hypothesis that should be metaphysically testable.

If we now ask: What makes some mental states a set, in the sense of parts of one and the same experience? then the answer is they are states of the same soul. The soul effects a qualitative transformation of sensory input (physiologically described) into a presented world of physical objects with secondary qualities and mental states, commonsensically described. Each soul transforms its physiological input into what I shall call a mental field. There is no such thing as awareness or consciousness of this mental field, but its existing or obtaining is all that can really be meant by saying persons are conscious or have awareness. Now, numerically distinct souls transform numerically distinct physiological inputs into numerically distinct mental fields. So, in that sense, the mental field is logically prior to the set or mental states.

The idea of experiences as distinct from one's experience arises in two ways. Firstly it is possible to think about experience in different ways: employ various criteria to discriminate one aspect of one's experience from another. Secondly, different physiological inputs and different soul transformations issue in different sorts of mental states, and secondary qualities. If it is argued the second is only one particular case of the first I should allocate some priority to it still as it is metaphysically testable. On both accounts the logical relation between experiences and mental field or experience is part and whole.

That then is how the soul as the transcendental self accounts for certain mental states being a set, or certain experiences being parts of one and the same experience.]

Kant's Holism

by

Stephen Priest

INSERT IN KD A 106 FF

Kant says that 'all necessity, without exception, is grounded in a transcendental condition.' (A106) What does this mean? Is it true?

I

If A grounds B then A makes B possible. This is ambiguous as it stands because

the modal 'possible' could commit us to either, if not for A then B would be logically impossible, or if not for A then it would be a contingent fact that B is impossible. Kant will want to rule out the second of these but work with a notion of impossibility weaker than logical impossibility if we mean by that: it is logically impossible that p if and only if p is contradictory. If p is necessary then if p is true then p could not have been false, and if p is false then p could not have been true. So 'necessary' means either 'not only true but could not not be true' or 'not only false but could not not be false'.

It is part of Kant's philosophical logic that the members of a sub-class of necessary truths have that status precisely because their negations are contradictory. For example analytic truths are self-contradictory to deny. This is consistent with his holding that if p is necessary then it is always the case that there exists a reason for p's necessity, but it is not obvious why all necessary truths require a transcendental ground if the contradictory status of their negations provides a sufficient explanation of the necessity of a sub-class. It is best to take Kant at his word here and read him as committed to: if p is necessary, and even if we can already explain what p's necessity consists in, p is grounded in a transcendental condition. This thesis commits him to; if p is necessary then there is some possible explanation of what p's necessity consists in, and, some possible explanation of how it is possible for p to be necessary. This leaves it open that one theory might execute both tasks.

I shall read 'transcendental' minimally, this way: if p is a transcendental claim then if p were false then knowledge or experience would not be possible, but p expresses no empirical proposition. By an empirical proposition I mean one which may be confirmed or refuted by observation (including Kant's inner sense), because p is about what may in principle be observed. I add this 'because' clause because, clearly the fact that some observation takes place could be a good (indirect) proof that some transcendental (ie non-empirical) proposition is true.

So now we can read the whole of the A106 claim this way: if p is necessary then p is necessary because some proposition q is true and q is not empirical.

Kant does not have a good independent argument for this principle but he uses it to establish the need for the transcendental unity of apperception. The A 106 principle is followed immediately by:

'There must therefore be a transcendental ground of the unity of consciousness in the synthesis of the manifold of all our intuitions'
(A106)

What entitles Kant to the 'therefore'? A problem is that p's necessity might be the explanation of p's possibility. If p is necessarily true then p is true and if p is true then p is possibly true. If we derive p's possibility from p's necessity then it is not clear that there is a need for any transcendental ground for p once we have analysed what p's necessity consists in. That is a problem for the principle in general, it might not be a problem for making the need for the TUA depend upon the existence of the unity of

consciousness.

We need to decide next what exactly the necessity is which needs to be grounded. By 'unity of consciousness' here Kant means empirical unity, or psychological unity and clearly it is not a necessary truth that there exists some unified empirical consciousness. What is necessary here, at least on Kant's account, is that if there exists a synthesis of the manifold of all our intuitions then there exist a unity of consciousness. The conjunction of these two claims is putatively a necessary truth. So, accepting the principle, it follows that there is a transcendental ground of that necessity: some value of p on which it will come out as necessary that if there is synthesis of the manifold of intuition then there exists a unity of consciousness. This thinking motivates the search for the transcendental unity of apperception.

It would seem from this argument that the need for a transcendental ground for experience is a special case of the principle that all necessity requires a transcendental ground. It is at least a consequence of that principle because Kant's reasoning is like this: All necessity requires a transcendental ground. Experience would not be possible unless a certain necessary truth held *viz* that if there exists a manifold of intuition then there exists a unity of consciousness. Therefore experience requires a transcendental condition. The form of this reasoning; if not P then not Q, but if not Q then not R, so if not P then not R, is valid but Kant has provided us with little reason to believe the first premise.

Kant thinks he has established the need for a transcendental ground and provides an argument for the claim that the transcendental ground cannot possibly be empirical, and a reason for identifying the ground with what he calls the 'transcendental unity of apperception'. This is the argument:

- (1) 'Consciousness of self according to the determinations of our state in inner perception is merely empirical, and always changing'
- (2) 'No fixed and abiding self can present itself in this flux of inner appearances'
- (3) 'What has necessarily to be represented as numerically identical cannot be thought as such through empirical data'
- (4) 'To render such a transcendental presupposition valid, there must be a condition which precedes all experience, and which makes experience itself possible'
(A107)

This is a Humean argument with a Kantian conclusion. The first premise, (1), is the claim that if I introspect then the knowledge I thereby obtain is empirical. 'Empirical' admits of a narrow and a wide construal. Narrowly, empirical knowledge is knowledge acquired only through sense perception (outer sense). Broadly, empirical knowledge is knowledge acquired only through experience (outer and inner sense). Clearly, in (1) Kant is working with the wide construal of 'empirical' on which it

comes out as true that if p is established only by the exercise of inner sense then it follows that p is empirical. That much is unproblematic.

An unclarity is introduced into the first premise by a failure to distinguish between consciousness and what consciousness is of in inner sense. Kant says consciousness of self is always changing but for the second premise he will need it to be the appearances which are always changing (ie are 'in flux'). The ambiguity is between my consciousness of my mental states, and the mental states themselves. This is an important ambiguity and its dispelling is quite neutral with regard to the central options in the ontology of mind. We can think of the various ways in which one may be conscious of oneself: be anxious about one's capacities, conscious of one's mood, hope one's memories are accurate, believe they are and so on. These acts of consciousness are prima facie quite distinct from their intentional objects even if they are in complex ways dependent upon them (for example, for their individuation). We can hold the intentional act constant and alter the object, or hold the object constant and alter the intentional act.

Now, Kant has said the consciousness of self is always changing, and that commits him to holding that the intentional act changes. But what he needs is the intentional object (in this case, the mental states) to be changing. He has no argument for either of these views and it is plausible to suppose that the reader is simply supposed to regard it as self-evident that no enduring psychological self may be found in introspection, once he or she introspects.

This seems to me not self evident. If I introspect then there is something which is present to my consciousness and which perhaps I consist in. We could call it 'the field of my awareness' or some such, borrowing this idea from the use of 'field' in 'visual field', but extending it to denote the unified consciousness that I feel myself to be, not just the visual sub-portion of it. I see no way of adjudicating between Hume and Kant on the one hand and myself on the other on this matter, and am prepared to hypothetically entertain the truth of their view to decide the validity of Kant's argument.

Kant says the consciousness of inner sense is 'always changing'. This claim is rather stronger than he needs and is in any case implausible. Nothing commits us to the view that if a person is conscious of one of their own mental states 'M' then M is always changing, or that that person's consciousness of M is always changing. I can attach sense to 'always changing'. For example, if M is always changing, then at any time that M exists M is losing some property or gaining some property or both, but from the fact that this claim makes sense it does not follow that it is true, and, again, I see no a priori nor any empirical objection to someone's being in M, being aware of being in M, and there being some time during which neither M nor the consciousness of M gains or loses any property. It seems to me this sometimes happens.

We may safely weaken Kant's claim and substitute another which will supply the necessary concept of mental change for the first premise. This is: There is no mental state of mine which lasts as long as I do. Suppose for example, I last from some time t_1 to some later time t_2 , then any mental state that I am in will last a shorter time that

duration $t_1 \dots t_2$. This is plausible because even if, as I suggest, if I introspect I am conscious of a kind of consciousness that I might consist in, there is nothing to suggest I am in this state unless I introspect, and even if I do, from the fact that I might be identical with my unified consciousness it does not follow that I am thus identical.

I have tried to cast doubt on the truth of the first premise, and I think the second premise (2) is very possibly false too. Notice however that (2) follows logically from (1) and so Kant has taken a step towards a valid argument for the non-empirical nature of the conditions for experience even if the premises are dubitable. If 'self' in (2) is the name for some candidate object of inner sense which I consist in then it follows from the fact that no object of inner sense which I consist in is unchanging, or no object of inner sense lasts as long as I do, that I am not identical with some item I encounter in inner sense. I do not consider the merely logical possibility that I consist in one and only one, always changing, mental state. I assume here that if A is B then A lasts just as long as B does. If your intuitions about your own consciousness are Humean rather than like mine you may consider the move from (1) to (2) sound.

Kant inserts a sentence in the argument at A107 saying that this discursive changing awareness one has of one's own mental states is called 'inner sense' or 'empirical apperception'. This is rather carelessly put because it sounds as though 'inner sense' and 'empirical apperception' are two names for the same faculty. In fact, inner sense yields experiences and empirical apperception subsumes those experiences under concepts. The distinction parallels that between intuition and understanding. In fact, it is a special case of it.

If we wish to construe Kant sympathetically here we may point to the mutual dependence of concept and intuition and say that their joint exercise may be partially but correctly described as 'inner sense' and partially but correctly described as 'empirical apperception'. Still, it would be a mistake to think that 'empirical apperception' and 'inner sense' differed only in sense and not also in reference. The point of Kant's terminological insertion is that he will soon wish to distinguish the transcendental unity of apperception from both inner sense and empirical apperception. It is important for the conditions of experience being transcendental that they are not falsely assimilated to certain conceptual and empirical faculties.

The third premise needs clarification. It is about the self, about what I am. The notion of identity deployed here is as follows. Suppose I am whatever has my experiences, whatever else I am. Then consider any two arbitrary experiences of mine which are not simultaneous. Call them E1 and E2. It follows that either E1 occurs before E2 or E2 occurs before E1 and it may or may not be true that some time elapses between the having of E1 and the having of E2. If we want to explain what it consists in for E1 and E2 to be my experiences, then part of this must consist in mentioning the fact that the self which has E1 is numerically identical with the self that has E2. This is the force of Kant's 'necessarily' in (3).

Now, whatever it is that has the role of self, whatever it is that has my experiences, cannot be thought empirically according to Kant. I take it this means that

no sentence comprising only empirical names and predicates could, logically, be a description of the self in this transcendental sense. By an 'empirical name' I mean the name of a property which is in principle observable. This is what Kant means by 'cannot be thought through empirical data'.

So, no thinking of empirical names and predicates is thinking about the self as that which is numerically identical over time and the owner of my experiences. Why should Kant think this? There is no doubt a point about the elusiveness of the self here: that which has one's experiences cannot be an item that falls amongst those experiences. The self is irreducibly subjective and active whereas the objects of one's experiences are, *qua* objects, passive and objective.

It seems to me this 'cannot' cannot be a logical 'cannot'. I see no objection to that which experiences at some time being an item falling within some experience at a different time and so on. For example, suppose the brain is that which experiences. Suppose my brain is that which has my experiences. Then it is a contingent fact about us that our brains do not perceive themselves. There is no reason in principle why they should not. There is no reason why subject of experience should be in principle so elusive as to not be empirical.

Kant insists that the conditions for experience cannot be empirical. There is an independently plausible thought which supports Kant's thesis here. Suppose X exists. Suppose there is some condition of X's existence which could be importantly mentioned in explaining how it is possible for X to exist. Then, plausibly, that condition cannot be any part of X (except modally, because anything is the explanation of its own possibility). That apart, if there are conditions for experience then it looks plausible to maintain that those conditions are not part of experience. Kant does not deploy this argument overtly but it may be supplied as a suppressed premise in the argument at A107.

There is a further and perhaps stronger reason than the alledged elusiveness of the self for finding the third premise persuasive. On Kant's view of necessary truth no empirical truths are necessary. But that I am self identical over times t1 and t2 if some experience at t1 is mine and some experience at t2 is mine is arguably a necessary truth. I do not rule out the logical possibility of my ceasing to exist for some finite duration between t1 and t2 but existing at both t1 and t2, but this cannot not affect the identity of myself at t1 with t2 necessarily holding if the two experiences are mine.

Now, if it is true that no sentence composed of empirical names and predicates uniquely is a necessary truth then what my self identity over time consists in cannot be expressed empirically. But is the assumption about the contingency of such empirical sentences correct? Suppose I am a wholly empirical object. I mean that any aspect of the whole that I am may be constructed using only empirical semantics. Suppose in an attempt to establish empirically my self identity between t1 and t2 I stay awake. I pay attention to the spatio-temporal continuity of that body that I am always associated with and so on continuously at all times between t1 and t2. It would not follow from this that I am self identical between t1 and t2. That could only be a deductive inference, and it is no experience which would confirm or refute it.

If we construe Kant this way then it comes out as true that no empirical thought, no thinking of sentences containing only empirical names and predicates and capable of only observational verification and falsification, will be a thinking of that necessary truth.

If we find this argument based on the premise that there are no empirical necessities persuasive, or if we find the intuition about the elusiveness of the self persuasive, then we have good grounds for accepting the third premise. (3) does not logically follow from (1) and (2) but it does not need to for the argument at A107 to be valid.

We may think of (1) and (2) as spelling out what certain sorts of empirical data consist in, viz empirical data with which we become acquainted through self-consciousness, and as establishing that all introspective self-knowledge is empirical. This means that no matter how thorough a subject's introspective scrutiny of their own mental states, if they come to know some true sentences by this scrutiny, none of those sentences is a necessary truth. Kant needs this otherwise the self which is thought as necessarily identical over the time it experiences could so be thought 'through empirical data'.

The conclusion, (4), is on one construal analytic, and so not the informative result Kant would really wish for. It is analytic because of the meaning of 'transcendental'. If P is a transcendental proposition then P's truth is a non-empirical condition which makes experience possible. It is therefore a conceptual truth that if P is to be true then there must be something non-empirical which makes experience possible.

If we rewrite the conclusion so that we have 'necessary truth' instead of 'transcendental supposition' then we can make the argument yield a conclusion to Kant's purpose. Then, as Kant no doubt intends, (4) makes reference to (3) by claiming there must be a transcendental ground for the necessity (3) reports. Kant is entitled to derive this conclusion if we accept his principle that there is a transcendental condition for every necessary truth.

So, construed this way the argument at A107 succeeds in establishing the need for a transcendental ground of experience. My self-identity over time is a condition for experiences at different times being mine. This is expressible by the necessary truth that my experiences at different times are mine. This fact cannot be expressed empirically. Every necessary truth has a transcendental ground so this necessary truth has a transcendental ground.

What Kant seeks to do now is identify the transcendental ground. To do this he deploys an argument to show that the transcendental ground is the transcendental unity of apperception.

II

Why the transcendental ground of experience is the Transcendental Unity of

Apperception

I extract the premises and conclusion of Kant's argument from the text at A107 in this way:

- (1) 'There can be in us no modes of knowledge, no connection of one mode of knowledge with another, without that unity of consciousness which precedes all data of intuitions, and by relation to which representation of objects is alone possible'
- (2) 'Even the purest objective unity, namely, that of the a priori concepts (space and time) is only possible through relation of the intuitions to such unity of consciousness'
- (3) 'The numerical unity of this apperception is thus the a priori ground of all concepts, just as the manifoldness of space and time is the a priori ground of the intuitions of sensibility'
- (4) 'This original and transcendental condition is no other than transcendental apperception' (A107)

The first premise expresses the truth that my self-identity over time is a condition for my having any knowledge, experience, or thought of objects. The force of this is possessing knowledge, or having thought about objects, presupposes not only having experiences but having experiences at different times and thinking of these as grouped into systematically related wholes. There is a holistic assumption at work in the text here: A putatively unique or solitary item of knowledge or experience would not count as such, or at least would not count as knowledge or experience of objects. It is one's experiences in their relation to one another which is partially constitutive of their being experiences of objects (of entities which exist independently of experience of them). Similarly it is sentences in logical and semantic relation to one another that constitutive knowledge claims.

If we accept this holism then we have to rule out the possibility that punctual selves could have just one item of knowledge each, or just one thought or experiences of an object each. So, if there is knowledge thought or experience of objects then there exist numerically distinct experiences over time. But if these are to facilitate any one person having knowledge thought or experience then those experiences must belong to that person. Hence, my self-identity over time is a condition for my knowledge.

(1) is fully consistent with the presupposition of the third premise of the A107 argument that if experiences at t_1 and at t_2 are mine then it is I who exist at both t_1 and t_2 . This thought is used by Kant in (1) as a necessary condition for my having any knowledge or thought or experience of objects whatsoever. We have here then one

kind of condition of experience; self-identity over time.

Kant wishes to claim rather more than this however, even in (1). He wants to make a claim not just that I am identical over time but a further claim about what I am. I am at least partly a unity of consciousness. We should raise the question here of what Kant means by 'unity of consciousness' and ask what entitles him to use this expression in a characterisation of what I am.

What has to be ruled out is any concept of an empirical unity of consciousness. This is clear because of the argument that the conditions of experience are transcendental, and not empirical, and by Kant's insistence that the unity in question 'precedes all data of intuitions' (my italics, A107). 'Precedes' here has no chronological connotation. It is a term used to formulate transcendental claims. I read it this way: A precedes B if and only if B is made possible by A and A is not any part of B. This implies that the unity of consciousness in question is not introspectively available. It rules out the possibility of its being amongst the data of inner intuition because it is not amongst the data of intuition tout court.

It is worth pausing at this point to decide whether an empirical unity of consciousness could play the conceptual role Kant has in mind for this non-empirical unity. Kant is antipathetic to any concept of empirical unity which violates the Kant-Hume principle that I cannot find myself in introspection as sustained mental state. Suppose, however, *pace* Hume and Kant, I am directly and intuitively aware of the continuity of my own consciousness when I introspect, despite changes in my thoughts and perceptions. There are reasons for thinking that even this would not meet the requirement Kant has on interpreting 'unity of consciousness'. This is because the unity he has in mind is purely formal. If an experience at t1 is mine and if an experience at t2 is mine then it follows that I am self-identical between t1 and t2 whatever I am, or whatever else is true of me. So, the expression 'unity of consciousness' carries with it here no strong implications for a particular ontology of the self. Any ontology of the self-that meets this purely formal requirement would in principle be adequate at this stage of Kant's account.

So, 'unity of consciousness' here is used simply to mention the fact that my experiences at different times, in order for them to count as mine, must have a single owner.

Premise (2) asserts that the unity of consciousness is transcendently more fundamental than another transcendental condition for experience: the unity of the forms of intuition, space and time. A is transcendently more fundamental than B if and only if experience is made possible by A and experience is made possible by B but B is made possible by A and A is not made possible by B. The unity of the forms of intuition is the necessary spatio-temporality of all our intuitions.

It was established in the Transcendental Aesthetic that because our forms of intuition are space and time all our intuitions are either temporal or else spatio-temporal. This necessarily follows from our having just those forms of intuition. (Kant allows that in principle there may be beings with other forms of intuition, or even intuitive understandings, so the necessity is one that follows from our

transcendental psychology).

Kant does not express himself so precisely in the second premise as in the Transcendental Aesthetic in at least one respect. He says in (2) the unity of the concepts of space and time is made possible by the unity of consciousness. I think he really intends to make the different and stronger claim that the unity of our intuitions in a single spatio-temporal framework is made possible by the unity of consciousness. He fails to distinguish clearly here between space and time and our concepts of them, something he is at pains to do in the Transcendental Aesthetic. If we do make that distinction we can say that both the unity of space and time and the unity of the concepts of space and time are made possible by the unity of consciousness. Indeed, it seems plausible to argue that the unity of space and time is a condition for the unity of the concepts of them, but Kant does not argue for this dependence.

Nor does Kant produce a clear argument for making the unity of intuitions in space and depend on the unity of consciousness, but one can be supplied. If it were not true that putative individual experiences at $t_1...t_n$ were experiences in a single mind, or experiences with a single owner then there would be no unity of experience tout court. A fortiori, there would be no unity of experience within any forms of intuition, for example space and time.

The soundness of this argument depends on Kant's view that space and time are transcendently ideal, because there seems to be no logical or a priori objection to some string of putative experiences existing over $t_1...t_n$ and being spatially interrelated, and being temporally interrelated even though they are not 'owned'. That we could not truly call these 'experiences' is an important part of Kant's point. We could no doubt call them 'events', but it is the unity of of consciousness that entitles us to call them experiences rather than events not their spatio-temporal relations. To put it another way, suppose we accept there are two kinds of condition for experience: (a) If a series of episodes is to count as a series of experiences they must be owned by a mind and, (b) if a series of episodes are to count as experiences they must be spatio-temporal. Kant wishes to make (b) depend on (a), or claim that (a) is more transcendently fundamental than (b). Unless we accept that space and time are transcendently ideal there is no good reason for this order of priorities. This means we have to insert the main conclusion of the Transcendental Aesthetic as a premise to obtain this order of priorities by a valid argument.

A quasi-Kantian reason could be supplied for making self-identity over time a more fundamental condition for experience than spatio-temporality. This would require taking seriously the conceptual constraints on our actual starting point in doing philosophy, and refusing to allow that it makes much sense to talk about space and time as transcendently real, as objective features of a mind-independent reality. To take a God's eye view is not an option open to us. Our only option is to talk about space and time as they appear to us, as we know them as features of the world as we actually experience it. There clearly there could not exist the spatio-temporal unity of experience in that sense unless there were a unified consciousness within which experience could exist.

Kant's reason for accepting the second premise is just that intuitions are experiences. Experiences would not be possible without a unity of consciousness and so no spatio-temporal intuitions would be possible without a unity of consciousness. This is valid but it does not rule out the possibility that the order of transcendental possibilities should run the other way: A unified consciousness would be impossible without experiences. Experiences are events. Events are spatio-temporal, so there would be no unity of consciousness without space and time. Kant has not ruled out the possibility of a mutual dependency between the unity of space and time and the unity of consciousness. I take it he thinks the Transcendental Aesthetic chapter simply precludes this.

Premise (3) draws an analogy between the unity of concepts in the transcendental unity of consciousness and the unity of intuitions in space and time. This analogy is perhaps rather forced as Kant has suddenly switched from talking about the unity of experiences to the unity of concepts. However, it is reasonably clear how the analogy is supposed to go. In the Transcendental Aesthetic he argued that putatively numerically discrete times are parts of one and the same time, and putatively numerically distinct spaces are parts of one and the same space. It follows that any part of time is temporally related to every other part of time, and any part of space is spatially related to every other part of space. It follows that any temporal item is temporally related to every other temporal item and any spatial item is spatially related to every other spatial item. There is only one space and only one time and the places and times within them therefore form a system. The analogy is that all my experiences are parts of one and the same experience which is called mine, just as different times are parts of one and the same time called 'time' and different places are parts of one and the same space called 'space'. Kant blurs the analogy by talking suddenly about concepts, but it can be made to work for any of my mental contents, and if concepts are (*inter alia*) mental contents it can work for those too.

There is another way of reading the third premise. This makes the point of analogy precisely the relation of being 'an a priori ground for'. So just as space and time are an *a priori* ground for all intuitions, so the unity of consciousness is an a priori ground for all concepts. A grounds B transcendently just in case A makes B possible and no proposition reporting this relation is empirical. Clearly, if it is true that every proposition is (exclusively) either empirical or a priori, it follows that P is *a priori* if P's truth is a transcendental ground for B.

It is worth explaining Kant's point about concepts here. Concepts need a transcendental ground for much the same sort of reason as experiences. If a concept that I possess is exercised at t1 and again at t2 it logically follows that I at t1 am numerically identical with the person at t2. Parallel assumptions about holism are also at work. It is perhaps not self contradictory to maintain that a concept might be exercised just once by a punctual being, but in fact our concepts form a complex logically and semantically interrelated whole and depend on their roles in judgements for their meanings. Arguably this is not possible unless a concept may be exercised repeatedly in different contexts. This logically presupposes the identity of a self as

concept user over time. If the holistic assumption is true then the identity of the self logically follows.

The conclusion (4) goes through easily on one very minimal construal of transcendental apperception: Whatever it is that arrests the regress of transcendental conditions falls under the description 'transcendental apperception'. Something arrests the regress just in case it transcendently grounds whatever is in need of a transcendental ground but it is not itself in need of a transcendental ground. We need now to decide what Kant thinks falls under this description.

III

What is the transcendental unity of apperception?

Kant expands a piecemeal answer to this question throughout both versions of the Transcendental Deduction. For clarity's sake I shall say at the outset that Kant will claim that it has at least the following essential properties: It is a formal requirement on experience, that is, a requirement of transcendental logic, not empirical psychology. It is part of what the self's identity over time consists in. It is a kind of disposition to self-consciousness. Propositions expressed by sentences characterising it are a priori in logical status. With these preliminaries in mind we may examine Kant's arguments for his answer to the question.

- (1) 'Unity of consciousness would be impossible if the mind in knowledge of the manifold could not become conscious of the identity of function whereby it synthetically combines it in one knowledge.'
- (2) 'The mind could never think its identity in the manifoldness of its representations, and indeed, think this identity a priori, if it did not have before its eyes the identity of its act whereby its subordinates all synthesis of apprehension (which is empirical) to a transcendental unity.'
- (3) 'The original and necessary consciousness of the identity of the self is thus at the same time a consciousness of an equally necessary unity of the synthesis of all appearances according to concepts.'
- (4) 'This transcendental unity of apperception forms out of all possible appearance, which can stand alongside one another in one experience, a connection of all these representations according to laws.' (A108)

The first premise is the claim that a certain unity of consciousness would be impossible unless a certain kind of self-consciousness exists. The unity of consciousness is the empirical unity of consciousness, but the self-consciousness is not, or not straightforwardly empirical. It entails my self-identity over time as a

condition for my experience. Kant wishes to elaborate on this claim by saying that I must in addition be conscious of my self identity over time for experience to be possible. We may think of these conditions as each singularly necessary for experience, but clearly my self-identity is also a necessary condition of my consciousness of my self identity. This is fully consistent with my self identity being a condition for experience because it will turn out that my consciousness of my self-identity partly consists in having experiences. (It is in fact just the special case of self-conscious experience).

Is this true? It might be conceded to Kant that my self-identity over time is a formal condition of a series of experiences counting as mine, but it might be objected that the further requirement, that I be conscious of this identity is redundant. After all, are there not certain beings, for example non-human animals, whose experience is unified in the requisite way but who are not conscious of their identity over time?

This line of objection misses Kant's point. He is saying that it must be possible, must in principle be possible, for a self to be conscious of its continued identity over time if its experiences are to be a unified whole. If this were impossible then, Kant assumes, the reason for this could only be that we were not talking about a self-same owner of experiences at all, and then we should have to give up talking about experiences in this case altogether.

A perhaps more plausible objection to Kant's reasoning is that it is logically possible that not the only reason why it might be in principle impossible for a self to be conscious of its identity over time is that it is not in fact a unified consciousness. Perhaps the being simply lacks the requisite mental apparatus and it is part of its nature not to be self-conscious. Kant does not have a reply to this in the text but I think he would say that then we would have to give up talking about a self altogether. Selves are necessarily self-conscious. We could talk about a being (as I have done) but not a self. This is consistent with his statement in the *Anthropology* that self consciousness distinguishes humans from animals.

Premise (1) contains a claim about knowledge which gives additional support to the view that a unified consciousness is a self consciousness. If a mind has (propositional) knowledge then arguably it is self-conscious. For example if I know that *p* then this might be because I have remembered that *p*, perceived that *p*, wished that *p*, judged that *p*, hoped that *p* and so on. The idea of psychologically atomic items of knowledge is perhaps not self-contradictory, but arguably we acquire, retain, revise and jettison our beliefs in the context of our propositional attitudes generally. Now, Kant thinks this presupposes that I can know that it is I, one and the same mind, which knows who perceives, desires, etc. If this were in principle impossible then we could not talk about 'knowledge' here. This is what Kant means when he says that the mind must be conscious of the identity of function whereby it combines the manifold and comes to possess knowledge.

A point needs to be made about Kant's use of 'mind' here. It carries no strong ontological connotation. It is whatever thinks, and whatever is self-identical over time. Kant is not concerned here to elaborate an ontology of mind, only to spell out

the formal conditions for experience and knowledge.

Premise (2) is the claim that premise (1) is only true on condition that a certain possibility obtains. The possibility is expressed by a metaphor. Kant says the mind must 'have before its eyes' the identity of the act whereby it subordinates synthesis to a transcendental unity. The transcendental unity just is the self's self identity over time argued for at A107. Synthesis means 'empirical synthesis': the application of concepts within experience to make it intelligible. 'Subordinates' is unclear but makes sense if it is another way of picking out the dependence of synthesis on the mind's identity over time. The question now is what the act is that Kant thinks subordinates synthesis to the transcendental unity of consciousness.

This cannot on the face of it be any straightforward mental act because all of those stand themselves in need of a transcendental ground. A more plausible construal is to put the stress on 'it's' in 'it's act', so then we have: the mind must be able to be conscious of the fact that its mental acts are its own, in order for them to be subordinated to a transcendental unity. This would be fully congruent with Kant's view that a mind which was in principle incapable of self consciousness could not be a unified consciousness, or a 'self'.

If we read Kant this way, then we can unpack the 'before its eyes' metaphor as a means of saying that a mind that is a self must be self conscious. So if a mind has before its eyes its identity between different mental acts it is conscious of its continued identity over time between those acts. The metaphor is a misleading one, because it suggests this self-consciousness has to be a kind of experience (a kind of inner seeing). But we know that the transcendental unity of consciousness is purely formal, and we know that no exercise of inner sense can of itself generate a concept of the persistence of the self over time. That was the conclusion of the 'Humean' argument at A107. If we say a self must be conscious of its numerical identity over time in order to be worthy of the name then this consciousness need not be any experience. It could for example be the thought that I, experiencing now, am identical with a person experiencing at some earlier time. Clearly any such thought will be partly about experiences, but it is not thereby an experience.

If the possibility Kant thinks he needs for the first premise is read this way, then the second premise as a whole can be read as saying something true. (2) claims a certain dependence between a self and its experiences as a condition for a certain kind of self-consciousness. It amounts to this: I could not be conscious of my continued self-identity over time unless, firstly, I had experiences at different times, and, secondly, I was conscious that those experiences were mine. That is the dependence running one way. Conversely, I could not be conscious of numerically distinct experiences occurring at different time as all mine unless I could be conscious of (think) my self-identity over time. That is the dependence running the other way. The second premise does the work of justifying the first because the first is just a version of the dependence running the first of these two ways.

Kant says I could not know a priori that I am self-identical over time unless I could be conscious that my experiences are mine. This is valid because if I could not

know that I am self-identical over time unless I could be conscious that my experiences are mine, a fortiori I could not know this a priori. Kant does not need to give this reason because he has already established that I cannot know my self-identity *a posteriori*, and if I know my self-identity, and if *a priori* and *a posteriori* are collectively exhaustive logical categories, then I can know my self identity a priori.

The third premise, (3) is a rephrasing of the second premise as an identity statement. Construed literally it is false. From the fact that consciousness of my numerical identity over time is a condition for my consciousness of my experiences as mine it does not follow that my consciousness of my experiences is my consciousness of myself, in the sense of that which has them. Nor does it follow that the thought of my self identity is one and the same with my consciousness of my experiences as mine. We could however obtain this identity with the addition of a premise which Kant has argued for elsewhere: the Humean view that from an empirical point of view I am nothing over and above my experiences. If I am the unity of my experiences, if I am my experience as a whole, the my consciousness of my self identity is a consciousness of my experiences because that is what 'self' refers to here.

A problem with this construal is that the consciousness of my numerical identity over time is a thought not an experience. It is, in fact, knowledge of a necessary truth. But my consciousness of my experiences is an experience, at least so long as we accept the Kantian view that I now I experience because I have a faculty of inner sense.

This objection can be met by insisting that the consciousness (thought) of the identity of the self is a consciousness (thought) of the unity of the consciousness. This is because the unity of my consciousness consists in the contents of my consciousness having a single owner. My self-identity over time is constitutive of the unity of my consciousness.

Kant also says the unity of the consciousness and the identity of the self are each necessary. He does not say what conception of necessity he has in mind. He could mean that each of the sentences reporting the unity of consciousness and the identity of the self expresses a necessary truth. This would have to be a kind of Cartesian necessity, so that 'I am self identical over time' and 'I am a unified consciousness' are necessarily true just so long as I think them. Alternatively, Kant might mean that if one of them is true then the other necessarily follows. This would make their conjunction a necessary truth, but the truth of some conjunction is not a sufficient condition for the truth of either conjunct. Kant seems to be saying in this case that each is singularly necessary. Another construal makes 'necessary' here mean 'necessary for experience' or 'transcendentally necessary'. This would have the merit of allowing my self identity over time and the unity of my consciousness to be necessary for experience. Kant says they are equally necessary. This is acceptable if they are mutually dependent and amount to different ways of expressing the same fact; the transcendental unity of consciousness. This precludes the possibility that one

is transcendently more fundamental than the other. Each of these two sentences: I am self identical over time, and, I am a unified consciousness is necessary and sufficient for the other.

We therefore have to read 'original' this way; My self identity is original. 'Original' means 'transcendentally fundamental'. The unity of consciousness is 'equally necessary' (for experience). The unity of consciousness is what my self-identity consists in.

I have counted (4) as the conclusion of the argument at A because there is a construal of it on which it logically follows from the conjunction of (1), (2) and (3). The transcendental unity of apperception has been explained so far as at least my self identity over time. All possible experiences can 'stand along side one another in one experience' just in case the transcendental unity of apperception holds. If by 'law' in (4) Kant means exactly this formal rule: Any experiences are parts of one and the same experience just in case they have one owner, then (4) is entailed by the argument so far. This is a reading which maximises the plausibility of Kant's view, but a drawback with it is that Kant uses the plural 'laws' and the transcendental unity of apperception has been so far formulated by a single sentence. So we need to look further. I take this up in the section The Function of the Transcendental Unity of Apperception.

INSERT IN KD B 132 FF

IV

What is the Transcendental Unity of Apperception? The Argument in B

There is a shift of emphasis between the two editions in the explanation of what the transcendental unity of apperception is. In the A edition the emphasis is on the identity of the self over time as a condition of experience, but in B it is the possibility of self-consciousness which emerges as definitely constitutive of the ground of experience. We shall have to decide what the logical relations are between these two conceptions and in what sense it might be possible to holistically combine them in a single principle, but I turn first to the B edition arguments.

The 'I Think'

Notoriously;

'It must be possible for the 'I think' to accompany all my representations; for otherwise something would be represented in me which could not be thought at all, and that is equivalent to saying that the representation would be impossible, or at

least would be nothing to me'
(B131-2).

APPENDIX FN1

Patricia Kitcher addresses the problem of in what sense the 'I think' of apperception may be correctly said to exist:

'The thinking self cannot be phenomenal (because it would be causally determined) and the doctrine of apperception cannot be about a noumenal self, because this doctrine is known, so the I of apperception is '*... given, given indeed to thought in general, and so not as appearance, not as thing in itself (noumenon), but as something which actually exists, and which in the proposition 'I think', is denoted as such (cf. B157).* But this solution is unavailing: 'existence' is a category. Kant applies another bandaid: 'the existence here (referred to) is not a category' (B423a). By his own or any other standards, Kant passes beyond the bounds of intelligibility at this point. It is totally unclear what the claim that the thinking I exists is to mean'.
(Kitcher, 1984, 123)

It is right that the 'I think' of apperception is not phenomenal and not noumenal. It is also right that 'existence' is a category. But it is not right that Kant thinks the 'I think' exists, unless this just means that the self possesses formal properties, as well as a phenomenal and a noumenal dimension. What Kant says in the passage Kitcher quotes is that the I of apperception is '*given to thought in generalas something which actually exists*' (my italics). Now, from the fact that something is given to thought as actually existing it does not follow that it exists. And, in fact it is Kant's view that it is a philosophical mistake to hypostatise the formal subject into a real thinking subject of experience. There is a tendency of reason, a tendency to stray from common sense to metaphysics, which partially consists in the thinking of the real existence of the subject. But two points need to be noted about this. It is a tendency Kant thinks we should resist, if we are not to be misled philosophically. And it is a tendency which makes use of the categories. In fact it misuses them. We need to view 'the existence here (referred to) is not a category' in a wider context. In the footnote at CPR 378, B423 the sentence is immediately qualified by:

'The category as such does not apply to an indeterminately given object but only to one of which we have a concept and about which we seek to know whether it does or does not exist outside the concept.'

Kant means that misthinking the I of apperception as a real thinking subject does require the (mis)use of the category of existence. But that category does not in fact or truly apply to the 'I think' because the 'I think' is a formal condition of experience.

(As he puts it in the footnote: 'The I think' precedes the experience which is required to determine the object' CPR 378, 423a). There remains the question of why 'I think' should be 'given as' something which exists if it does not. To understand this we need to bear in mind Kant's remark that 'I think' has an empirical (as well as a transcendental use). For example he says 'I have called the proposition 'I think' an empirical proposition' and 'The 'I think' expressed an indeterminate empirical intuition ie perception' (CPR 378, B423a). This can be taken in two ways. 'I think' construed empirically might mean I this man or this human being think. That is consistent with his view that whole persons think, or the thought, 'I think' can appear as one thought amongst other. But on either of these (complementary) empirical construals one thing is clear. No inner or subjective 'I' appears to consciousness. Read this way Kant is not inconsistent, and there is not need to postulate Kitcher's 'transcendental, phenomenal, empirical (that is, existing) thinker' (Kitcher, 1984, 138).

So my reading of the transcendental unity of apperception differs from Kitcher's:

'Consciousness requires self-consciousness in other words it must be possible for all my representations to belong to a thinking self (a contentually interconnected system of mental states).' (Kitcher, 1984, 143)

The possibility Kitcher requires here is, on Kant's view, a necessity. It is in fact analytic on Kant's theory that all my mental states belong to a thinking self. The possibility for Kant is that of the 'I think' prefixing any of my thoughts, as a transcendental condition of their being mine. It is this dispositional and formal character of the 'I think' which is essential to Kant's concept of the transcendental unity of apperception, but which does not feature in what Kitcher labels 'the principle of apperception' (Kitcher, 1984, 143). This omission leads Kitcher to see the unity of the mind as a matter of content not form. (See above, also Kitcher 1982, (a).) But clearly that two or more mental states have qualitatively similar contents is neither a necessary nor a sufficient condition of their belonging to the same mind. The 'interconnected system', to use Kitcher's expression, exists because of the formal properties of the self expressed in the transcendental unity of apperception. It is that same possibility of the 'I think' accompanying thoughts that effects the transition to talk about self, as well as unified mind.

Sydney Shoemaker is then clearly correct to argue that the transcendental unity of apperception is essentially a doctrine about self-consciousness, and not about synthesis. (See Shoemaker, 1984, esp.151 ff.) My dispositional reading of the transcendental unity of apperception is consistent with a suggestion of Shoemaker's about the connection between consciousness and self-consciousness:

'Synthesis must be such as to best a considerable degree of rationality on the subject of the mental states, and ... in creature of any appreciable degree of conceptual sophistication, and more particularly those capable of conceiving of themselves this

rationality requires at least an appreciable degree of self-awareness. One reason for this is that it is only to the extent that a creature knows what its beliefs are that it is in a position to modify them in the light of new experience, in the way require by rationality.' (Shoemaker, 1984)

Shoemaker refers to McGinn (1983, 20-1) in support of this view. Clearly, a creature both rational and self-conscious in this sense need not be perennially self-conscious. Not every thought that it has needs to be an occurrently self-conscious thought. For example, it is not a necessary condition for the self-consciousness of such a creature that it know what all of its beliefs are: only some relevant subset of them. There is nothing either to preclude our holding that the subject of such rationality is partly constituted by the possibility of such self-consciousness. The proposal is one to which Kant would be sympathetic.

INSERT IN TD AT A 106 FF

Kant's Empirical Self

by

Stephen Priest

I examine some of Kant's arguments for his conclusion that the conditions for experience cannot be wholly empirical. I then evaluate his thesis that the ultimate condition for experience is a formal possibility of self consciousness and not an empirical or metaphysical self.

I

When Kant asks how experience is possible, this breaks down into the sub-issues of how inner and outer sense are possible, how the differentiation between perceiving subject and external world is possible, and how it is possible for subjective presentations to be ordered experiences of a world of interacting physical objects. In different passages in the Transcendental Deduction in the first (A) edition of *The Critique of Pure Reason* (1) Kant has the justification of different features of experience in mind. He begins by introducing us to a regress of conditions for experience, where each newly revealed condition is also a condition for the conditions revealed so far. Kant's theory of the self is in the last resort designed to arrest this regress of conditions, and exhibit the mutual dependency of all subjective-objective distinctions on a fundamental condition which is neither subjective nor objective but purely formal.

The unity of consciousness is one such condition for experience which is neither subjective nor objective. Indeed, the distinction between empirical subject and

empirical object, inner and outer sense, are made within the unity of consciousness, so in an obvious sense these distinctions within experience presuppose a prior experiential unity. But Kant thinks there exists a further condition, a condition for the unity of consciousness:

'All necessity, without exception, is grounded in a transcendental condition. There must, therefore, be a transcendental ground of the unity of consciousness, in the synthesis of the manifold of all our intuitions, and consequently also of the concepts of objects in general, and so of all objects of experience, a ground without which it would be impossible to think any object for our intuitions.' (CPR 135-6, A106)

There is a condition, then, which is not only a condition of experience, but also the condition of all other conditions of experience. This Kant calls the Transcendental Unity of Apperception. This is the most fundamental concept in Kant's theory of self (and arguably the key concept for making sense of *The Critique of Pure Reason* as a whole). The above passage introduces the need for the transcendental unity of apperception (henceforth TUA). Kant says that all necessity is grounded in transcendental conditions. I take it this means that each variety of necessity has transcendental conditions.

Clearly there are different senses in which Kant could be taken to maintain there are conditions necessary for experience. For example, a condition for experience might be a necessary condition for experience, so that, if that condition did not obtain, then there could not be experience (in some sense of 'could'). Or, 'If there is experience then condition 'C' obtains might be a necessary truth: for example if the negation of that sentence contains a contradiction. But Kant has said that 'all necessity without exception' is grounded in transcendental conditions. So the interpretation of 'necessary' here is not relevant to deciding why the condition is necessary. If it is true that the unity of consciousness is a condition of experience, then it is true that there is a condition of the unity of consciousness, whether or not it is true that the unity of consciousness is a necessary condition for experience, or whether it is a necessary truth that the unity of consciousness is a condition for experience. The two readings are in any case mutually consistent, and Kant gives us no grounds for preferring one to the other.

Prima facie 'All necessity without exception' presents Kant with a problem, because if we ask whether it is necessary that the TUA is a condition for the unity of consciousness, and if Kant's reply is Yes, then that too will stand in need of some transcendental ground and the regress of conditions will not be arrested. Kant is aware of the problem, and it should be born in mind in what follows.

If we accept the premise that wherever there is necessity there is a transcendental ground, and if we further accept that some notion necessity is involved in the unity of consciousness being a condition for experience, then we may allow that Kant has validly inferred that 'there must therefore be a transcendental ground of the unity of consciousness'. But nothing obviously follows about the nature of this ground from

those premisees alone. By 'transcendental' we may infer that some sentence of set of sentences must be true if the unity of consciousness is to be possible, but the argument so far gives little clue as to the semantics of those sentences. In particular, it does not follow without further premisees that the condition is 'In the synthesis of the manifold of all our intuitions'. That might seem a plausible area to conduct the philosophical investigation, but Kant has so far given us no compelling reason for searching there. He uses 'must' to assert that the condition must necessarily be of a certain nature.

But whatever the nature of the TUA its centrality to the *Critique of Pure Reason* is evident. It is in fact a condition of the synthesis of the manifold of intuition, and Kant says 'consequently' of concepts of objects in general. The relationship between intuition and category is mutual or reciprocal and here Kant is simply fastening on to one side of the dependence: that of categories on synthesised intuitions. But clearly the TUA is a condition both of the categories, and of the synthesis of intuitions irrespective of their reciprocal dependence; so it makes possible the unity of two of the major themes of the critical philosophy: the use of reason, and the use of the senses. When Kant says the TUA makes possible 'all objects of experience' and 'any object for our intuitions' he clearly intends 'any' and 'all' to cover items experienced in both inner and outer sense.

Kant thinks the distinctions between mental and physical and subjective and objective depend on the distinction between inner and outer sense. This means that unless there were a distinction between inner and outer sense there could not be distinctions between mind and matter, or my mental state and yours, or in general, between what only pertains to the psychology of the subject and what pertains only to the external world. So, if the TUA makes the unity of consciousness possible, and if the unity of consciousness makes all experience possible, then clearly the TUA also makes possible all the distinctions within the structure of experience. This is what I mean when I say the TUA is the fundamental concept in *The Critique of Pure Reason*. The discussion of every other topic in that book presupposes the TUA as its ultimate condition. It should be clear then that Kant has left far behind the issue of how experience of an objective enduring world is possible given our experiences are subjective and relatively non-enduring. This is a Kantian problem - treated for example in the Principles chapter and the Analogies, but it is not the most fundamental problem. Indeed, its very formulation, with its confidence in subjectivity and scepticism about objectivity, makes use of a subjective-objective distinction which relies in the last resort on the TUA.

So Kant is not ultimately a subjectivist trying to show how objectivity is possible. Nor is he an objectivist trying to show how subjectivity is possible. His philosophical project is to show how both subjectivity and objectivity are possible. This project is quite consistent with his exhibiting mutual dependencies between subjective and objective features of experience, and this allows commentators to adopt competing positions about Kant's subjectivism or objectivism. It needs to be shown that the transcendental unity of apperception is identical with the condition for all conditions

for experience. It needs to be shown that Kant believes this, and then his demonstration of it needs to be examined. directly after the paragraph at PCR 135-6, A 106 in which Kant diagnoses the need for a condition of experience and of all other conditions, he says:

'This original and transcendental condition is no other than transcendental apperception.' (CPR 136, A107)

This is unequivocal textual proof that the TUA is the condition Kant is seeking. When he speaks of an original condition he means: an original condition is a condition of all conditions. An original condition has no condition: there is not any condition of an original condition. The TUA, Kant will argue, is an original condition in this sense. When he speaks of a transcendental condition, he is speaking of what makes experience, or knowledge possible. He has earlier clearly identified transcendental knowledge with knowledge of how knowledge is possible. It remains to be decided what transcendental apperception is, and more especially what the transcendental unity of appreciation is. But first we should note that Kant dispels a prima facie plausible candidate for the role of original condition of experience.

II

The Empirical Self

Kant deploys an argument to show that the original condition of experience cannot possibly be the empirical self. The empirical self is oneself as one is aware of oneself through inner and outer sense and as one thinks of oneself as presented through inner and outer sense. This is the argument:

- (1) 'Consciousness of self according to the determination of our state in inner perception is merely empirical, and always changing.'
 - (2) 'No fixed and abiding self can present itself in this flux of inner appearances.'
 - (3) 'What has necessarily to be represented as numerically identical cannot be thought as such through empirical data.'
- (C) 'To render such a transcendental presupposition valid, there must be a condition which precedes all experience, and which makes experience itself possible.'
(CPR 136, A107)

I have just said the empirical self appears in outer as well as inner sense. This is correct, but here Kant is concerned with the empirical self of inner sense and (empirical) apperception only. This is clear from his remark about 'this' flux of inner

appearance. Such consciousness is usually named 'inner sense' or 'empirical apperception' (CPR 136, A107). We may concede to Kant the first premise because it asserts that inner sense is an empirical consciousness of self. In inner sense I appear to myself as a series of mental states. This counts as empirical because it is an experience of myself, and it is an experience of myself as I appear to myself, not as I am in myself. Given that I am not always in the same mental state but that I am always in some mental state at least during the exercise of inner sense, then it follows that one's state is 'always changing', so long as this is not taken to preclude my being in the same mental state for any length of time at all. Kant clearly does not mean this. He means that if I am in a mental state then I will cease to be in that mental state and I will be in a qualitatively distinct mental state. When he says inner sense is a 'consciousness of self' he means there exists awareness of this process: the sequence of mental states. (This awareness is not itself permanent. There is nothing to suggest Kant thinks we perpetually exercise inner sense, and that inner sense is not perpetually exercised is itself consistent with Kant's view of one's mental states as 'always changing').

If we accept this picture of empirical self-consciousness then we should accept the second premise also. It straightforwardly follows from the fact that any mental state will be replaced by another mental state that there exists no mental state of mine which persists so long as I do. So there is no mental state with which I may identify myself. If only mental states appear to inner sense, then if a self is to appear to inner sense then that self must be a mental state. But a self predates, lasts during and postdates any (except its first and last) mental states, so no self is identical with any of its mental state if these exist in a temporal sequence. So premise (2) is true if premise (1) is true.

The description 'what has to be represented as numerically identical' is one component of a definition of 'self'. The subject of a sequence of mental states, that which has them, has to be thought of as remaining identical through those mental state to count as the subject of all of them. That is the force of 'necessarily'. Unless such a putative subject were self identical over time we could not talk about a self here. The second part of the third premise is such a self identical self cannot be something introspected - cannot be an empirical object of inner sense - because as argued by the first two premises no such enduring self is presented to inner sense. So the third premise follows from the first two. If they are true then there cannot be any (introspectively) empirical concept of the subject of experience. Kant says 'as such' because he does not preclude the logical possibility of an empirical subject. We do have an empirical conception of ourselves as enduring empirical subjects. But this conception cannot be of what does the work of being the real subject of our mental states.

The conclusion is that there must exist a condition for experience which by implication is not an item to be found within experience. I take it a condition which 'precedes' experience is just the same as a condition which 'makes experiences itself possible'. 'Precedes' contains no chronological connotation here. The 'transcendental

supposition Kant mentions is the subject of experiences, or better, a condition which will fulfil the role that we typically think of as taken by a subject of experience: being the owner of experience and making experience possible. This conclusion follows from the three premises so long as we read the putative self-identical self as being a condition of their being experience, not just that which has experience. Kant is entitled to this assumption so long as it is true in some sense that there are no 'unowned' experiences, and the owner endures throughout those experiences.

Kant's argument is valid so, if the premises are true, it is true that the empirical self cannot be the original condition of experience. If the quasi-Humean grounds for the premises are conclusive then the argument is sound, and it is the case that the original condition is not empirical. If we believe in an original condition for experience it in any case is highly plausible to deny that it is empirical. To accept that would be to make some item within one's experience the condition of all of it.

III

The TUA as the Original Condition for Experience

Kant next deploys an argument to show that the original condition of experience is identical with the TUA as he will define it:

(1) 'Unity of consciousness would be impossible if the mind in knowledge of the manifold could not become conscious of the identity of function whereby it synthetically combines it in one knowledge.'

(2) 'The original and necessary consciousness of the self is thus at the same time a consciousness of an equally necessary unity of the synthesis of all appearances according to concepts, that is, according to rules.'

(3) 'The mind could never think its identity in the manifoldness of its representations, and indeed think this identity a priori, if it did not have before its eyes the identity of its act, whereby it subordinates all synthesis of apprehension (which is empirical) to a transcendental unity, thereby rendering possible their interconnection according to *a priori* rules.'

(4) 'This pure original unchangeable consciousness I shall name transcendental apperception.' (CPR 136-7, A107-8)

The first premise establishes the identity Kant seeks between the role putatively filled by a self-identical subject over time, and the original condition for experience. Until this point in the Transcendental Deduction we have been introduced to the two

notions separately. He has argued that the empirical self cannot be the subject of its experiences, and he has argued that there is an original condition of experience. Now he asserts that the possibility of a mind's consciousness of its self identity over time is a condition for its experience. This raises a problem that needs to be dealt with immediately; what does Kant mean by 'mind' here? There is no reason to suppose that Kant is committed to any sense of mind stronger than 'set of unified experiences' at this stage. It would, I think, be a mistake to read mind to mean 'transcendental ego' or still less 'soul' here or even 'noumenal self'. There is no evidence Kant is postulating a mental entity, and much evidence that he is at pains not to.

Kant introduces the idea of an original condition by 'unity of consciousness would be impossible if'. Unless a certain fact obtained there could be no unity of consciousness. If that is true, and if there exists a unity of consciousness then that fact must be true. The necessary fact is that the mind (a set of unified experiences) must be able to be conscious of 'the identity of function' in synthesis. To understand this we need to know what a function is. I understand a function here to be an act of synthesis: a unifying mental act by which a mind finds intelligible the objects of its experience. So, suppose some mind thinks of or perceives some object then that will consist in a series of mental acts; call them F1, F2 etc.

These mental acts or functions of synthesis constitute the object as an object of experience. Now, by 'identity of function' Kant does not mean that F1, F2, etc have to be the same kind of mental act. He does not mean that they have to be qualitatively identical. He means they have to be the acts of a single mind; episodes within one and the same unified consciousness. Even this though does not sufficiently spell out all the necessary conditions for a unity of consciousness. It must be possible for a mind to be conscious that all its separate acts of synthesis belong to itself for it to count as a mind - a unified consciousness - at all. Clearly it is a sufficient condition of the existence of a unity of consciousness that a unified consciousness be aware of an object of experience (or synthesis the manifold). But for this to be possible that mind must be able to be aware that it is the self-same owner of those mental acts. Why does Kant think this? To see why consider the converse putative possibility. Then it would be impossible even in principle for a mind to be conscious of its mental states as its own. But the only plausible reasons for asserting this would be reasons for denying that there existed a unified consciousness at all. So being a unified consciousness partly consists in the possibility of being aware that one is one. Being able to know that one is one is a condition of being one. The first premise is the first step in the identification of the original condition of experience with the TUA because the concept of awareness of self identity over time will turn out to be an important part of what the TUA consists in. It is important to notice that the TUA cannot be explained in terms of synthesis. Rather, the converse is true. Synthesis is possible on condition the TUA holds.

The second premise spells out in a slightly different fashion the relationship between the unity of consciousness and the consciousness of that unity; it says it 'is' it. This cannot be the 'is' of strict identity because (with dubious coherence) the

consciousness of it would just be it, and the construction 'consciousness of' or 'awareness of' would be redundant. The relationship is just that outlined above: the possibility of the consciousness of the unity of consciousness is a condition for the unity of consciousness. Clearly 'X is a condition for Y' and 'X is Y' do not express the same proposition (except where something is a condition of itself), so we should read 'is' here just to express that close dependency. Just as 'mind' raised a *prima facie* problem of interpretation in the first premise, so 'self' raises a similar problem in the second. By 'self' here Kant is not committed ontologically to anything distinct from the unity of consciousness, but the unity of consciousness understood under a special description: the unity of consciousness understood as being capable of being conscious of itself as a condition of itself.

Now we may read Kant as saying that the original and necessary consciousness of the self, that is; the possibility of the unity of consciousness' self-awareness, is automatically an awareness of the unity of the synthesis of all appearances. It is an awareness of those unified mental functions which make up the unified consciousness. Being aware of the unity of consciousness is being aware of the set of acts of synthesis by which the manifold of appearance is made intelligible as the objects of experience. Now we can read 'is' as the is of strict identity: the self, the unity of consciousness, and the unified acts of synthesis are not numerically distinct one from another, so what falls under one of these descriptions falls under the other two. Kant says that the consciousness of self and the unity of synthesis are 'equally necessary'. This must mean 'equally necessary if there is experience, or if there is a unity of consciousness', because clearly it is a contingent fact that there exists a unity of consciousness: there might not have been such a unity. And clearly it is equally a contingent fact that there is an original condition for experience; there might not have been any experience. But if there is experience, then it is necessary that there is a unity of consciousness. And if there is experience, and if there is a unity of consciousness, then it is necessary that there is an original condition of experience. To say that they are 'equally' necessary is not to suggest that the unity of consciousness could obtain (still less necessarily obtain) without the original condition of experience. 'Equally' here does not imply the unity of consciousness and the TUA are equivalent in the transcendental hierarchy. Kant means that they are necessary in the same sense. Given that there is experience, it is necessary that there is the unity of consciousness, and it is necessary that there is the original condition of experience. The first necessity exists because unless there were a unity of consciousness there could not be an application of the categories to experience, and the manifold could not be understood as a rule governed world of objects. The second sort of necessity exists because unless it were at least in principle possible for the unity of consciousness to be aware of its self-identity over time, we could not speak of a unity of consciousness at all.

The second premise follows from the first because it amplifies or makes explicit what the unity of consciousness is.

While the second premise explains a condition for the identity of the mind as a

unity of consciousness, the third premise introduces a condition for the consciousness of that identity. The condition of its thinking its identity a priori is its awareness of the act of relating the acts of synthesis to itself as a unified consciousness. So a transition has been effected from this claim: that a condition of a unified consciousness is its awareness of its self identity, to this claim, a condition of that awareness is a further awareness of its relating its acts of synthesis to itself as that unified consciousness. Why does Kant think this additional condition has to be met? Why does Kant not think it sufficient for the unity of consciousness' thought of its own identity that it merely be able to consciously relate its acts of synthesis to itself as that unity of consciousness? Why in addition must this be a self-conscious process? Several reasons may be advanced, but the most plausible is this. Suppose a mind is conscious of its self identity over time partly by the exercise of a particular mental act: the thought of its identity across some set of mental states of which it considers itself the subject. Then that mental act will be one mental act amongst others - a self-conscious mental state. If that mental act is itself to count as part of the self-same unified mind which is the object of its attention then it in turn must in principle admit of being the object of a further act of reflection. If this were not the case - if this were not a possibility - then Kant would have admitted an exception to this rule: the TUA makes experience possible. This is because there would then exist at least one mental state that was not transcendently grounded in the original condition: *viz* any mental state which realised the possibility of the unity of consciousness' awareness of its self-identity over time. Clearly Kant cannot allow this if he is to maintain that the TUA is the original condition of experience, because it would amount to a violation of the first conjunct of the definition of 'original condition'. It would not be the condition of all conditions, because it would not be the condition of the possibility of a mind's consciousness of its own identity, and that is a condition: a condition of experience.

It is worth raising the question of whether this meta-claim expressed by the third premise commits Kant to a regress of conditions beyond the TUA, which it is Kant's purpose to prevent. This appears *prima facie* to be so, because must not that meta-possibility of the awareness of a possible act of awareness of self identity itself stand in need of a further possible act of self-awareness and so on *ad infinitum*? The regress is in fact illusory because the possibility of the self awareness of a self-conscious mental state is just one instance amongst others of the generalisation that the TUA expresses: that it must in principle be possible for a mind to be aware of its own self identity. This entails only the weak claim that any mental act may in principle be an object of reflection-not the strong claim that every mental state be an object of reflection. *A fortiori* although any self-conscious mental state must in principle admit of being an object of reflection it does not logically follow that every self-conscious mental state be an object of reflection.

I have used expressions like 'self conscious' and 'reflection' here to unpack Kant's metaphor; 'the mind has before its eyes the identity of its act'. The metaphor suggests a kind of self-conscious experience because seeing is experiencing, but there are

reasons for supposing Kant does not have in mind or at least does not just have in mind an operation of inner sense. This is because inner sense only reveals to a person their mental states, it does not of itself provide the thought of the belonging of those mental states to one and the same mind, still less the consciousness of that's mind's identity over time. So the experiential interpretation of the metaphor should be played down in favour of this: The mind must be able to think its self identity in the acts whereby it consciously relates its mental states to itself as a unity of consciousness. If the realisation of that possibility requires a kind of self-conscious experience - something's appearing to the mind's eye to pursue Kant's metaphor - then no harm is done by allocating that introspective operation to inner sense so long as we realise the exercise of that inner sense alone is not what the existence of the possibility consists in.

Additionally, Kant says the thought is *a priori*. This precludes its being exhausted by an act of inner sense because all the acts of inner sense are empirical. In any case, Kant has argued at length that the empirical self is not the transcendental condition for experience, so he cannot admit that any empirical act of self-awareness is of itself a condition of the unity of consciousness. Kant explicitly states that the synthesis of apprehension is empirical, but we know that the transcendental unity of consciousness is *a priori* so it is hard to see how any empirical act of self-consciousness could do the work of consciously relating the two.

I conclude that Kant has presented persuasive grounds for believing that the conditions for experience cannot be wholly empirical. It is far less clear that they can be only formal if they are to ground the reality of our experience. What the ontology of a self-conscious mind consists in remains an unsolved philosophical problem. (2)

(1) All references to the *Critique of Pure Reason* are to *Immanuel Kant's Critique of Pure Reason* trans. Norman Kemp Smith (Macmillan, London, 1978).

(2) Patricia Kitcher calls 'the problem of "too many selves" the coherent reconciliation of the noumenal self, the phenomenal self and whatever is denoted by the first person singular pronoun in 'I think'. (See, for example, her *Kant's Transcendental Psychology* (Oxford University Press, Oxford, 1990) p.139. Kitcher's problem is intricate, but the form of a solution is: The noumenal self and the phenomenal self are numerically identical. Just one self exists as it is and as it appears. The use of 'I' in 'I think' is abstract. It denotes whoever uses it whatever the true ontology of that user. This form of solution reconciles Kant's *prima facie* inconsistent claim that

'The I think expresses an indeterminate empirical intuition, i.e., perception. Something real that is given, given indeed to thought in general, and so not as appearance [phenomenon], nor as thing in itself (noumenon), but as something which actually exists, and which in the proposition I think is denoted as such'

Stephen Priest *Light on the Subject* (2016 Draft)

(B422-23n) (quoted by Kitcher p. 139)

3

HEGEL'S
MASTER AND SLAVE DIALECTIC

Stephen Priest

Stephen Priest is Senior Research Fellow in Philosophy in the University of Oxford. He lectures in the United States, Europe, and Japan, and his work has been translated into six languages.

Stephen Priest
Faculty of Philosophy
University of Oxford
10 Merton Street
Oxford OX1 4JJ
England UK

http://www.philosophy.ox.ac.uk/members/senior_research_fellows/stephen_priest

Preface

In Part One I provide an elementary overview of Hegel's system, locating the Master

Stephen Priest *Light on the Subject* (2016 Draft)

and Slave Dialectic within it. The overview is derived from three decades of lecturing on Hegel.

Stephen Priest

Oxford
Trinity 2012

Contents

Part One: Introduction 1

Part Two: A Commentary on the Chapter 'Self-Consciousness' in
Hegel's *Phenomenology of Spirit* (1807) 21

Part One: Hegel's System

(1) Two Hegels

OK, today I'll tell you about Hegel on phenomenology and dialectic. Now, to

understand this we need to realize that there are two aspects to Hegel's system.

Firstly, Hegel engages in what he calls 'Phenomenology', and the *Phenomenology of Spirit*, the Jena 1807 work, is the most important statement of Hegel's phenomenology. Phenomenology is *the description of what appears to consciousness*, so a phenomenology of *spirit* is a description of how spirit, or consciousness, appears to itself and that's what that book is about.

Now, the other part of Hegel's system or the philosophical method that Hegel's employs, apart from phenomenology, is called 'logic'. Now, logic is not logic. For example, by 'logic' Hegel does not mean the study of deductive relationships between sets of propositions. He doesn't mean the exhibiting of inference relations between truth valued items or any thing like that, and it's been a source of dismay and disappointment to positivists and analytical philosophers to discover the contents of the books by Hegel which have the word 'logic' on the cover. Hegel means something quite different. By 'logic', Hegel means *dialectic*.

Now, what does he mean by dialectic? Well, there are two ways, or many ways, of understanding this but two are as follows: firstly, dialectic is *the demonstration of the dependencies between universals*, where universals are concepts or, if you like, universals are *categories* in a sense analagous to Kant's or Aristotle's categories, or partly analagous. I should say there are many differences between Kant's concept of a category and Hegel's of a category. For example, Kant thinks, in the *Critique of Pure Reason*, there are only twelve categories. Hegel thinks there are many more of them. Kant thinks that the categories can only be used to understand the world as it appears, but Hegel thinks that the categories can be used to understand the world as it really is. Kant thinks the categories are fixed in number and nature. Hegel thinks the categories are historically changing. So there are enormous differences between Kant and Hegel on categories. But they have one very important point in common: As concepts, or as universals, the categories are *constitutive of what is*, in an idealist sense. So that's one way of understanding what dialectic is: the exhibition of dependencies between universals or Categories.

A second way of understanding dialectic is as the ancient philosophical technique of *establishing the identity of opposites*. Dialectic, in this sense, is clearly recurrent in different philosophical traditions. To the extent to which it makes sense to talk about traditions in philosophy, there is dialectic as the exhibiting of dependencies between opposites, for example, in Plato's philosophy. Remember the argument from opposites in Plato's dialogue called the *Phaedo*, the dialogue between Socrates and Sebes, and Semias. Or, dialectic can be found in ancient Chinese philosophy in the mutual dependency of Yin and Yang, and in Taoist philosophy. Or, dialectic is in Marxist politics and philosophy, perhaps essentially. Dialectic in this sense of exhibiting the unity of opposites can be found in the work of the Neo-Platonist philosopher Plotinus in the *Enneads*. Well, Hegel also understands dialectic in this sense. Hegel thinks philosophy is constituted by oppositions, or opposed concepts, such as: subjective and objective, finite and infinite, free and determined, mental and physical, sacred and profane, and ultimately, Being and Nothingness. Hegel believes

that in philosophy we operate with pairs of opposites, For example: free and determined, mental and physical, sacred and profane, space and time, right and duty etc. Doing philosophy involves privileging one of these concepts over its opposite, For example: stating that people are free rather than determined. Dialectic will show the *dependencies* between these *prima facie* opposed concepts. Seemingly mutually exclusive categories actually require one another for their intelligibility and their application. Each is found in the other as both its limitation and its fulfilment. The opposed, or antithetical, categories are relieved or synthesised or combined in a new category or 'synthesis'.

Some terminology. Hegel uses a very specific vocabulary. He has dismayed many thinkers with his tendency to state paradoxes (eg: Schopenhauer Isaiah Berlin, Karl Popper). Dialectic is a method of philosophy.

Hegel claims the attempt to prove one position and refute its opposite is antagonistic or 'one-sided'. Because philosophy operates in these terms, it makes no fundamental progress. Hegel argues we should employ the technique of reasoning known as 'dialectic'. This involves showing binary oppositions to be mutually dependent. There are no rights without obligations. There is no subjectivity without objectivity, no individual without society, not up without down, and conversely. We can only achieve an adequate understanding of humanity, and its place in reality, if we realise that humans partake of both opposites. Consider the view known as compatibilism, according to which both determinism and libertarianism are true. Hegel can be understood as a compatibilist in all philosophy. He is against the view that positions can be refuted conclusively refuted in a way that entails a concept but not its opposite finding an application. Someone who has adopted a position without endorsing its opposite has only a partial grasp of the truth. Hegel claims his philosophy is holistic. It synthesises partial views into a larger whole.

This is an example of dialectic from the *Philosophy of Right* (1821). Hegel says there is a dialectical relationship between rights and obligations. This means that (i) rights and obligations are opposites and (ii) rights and obligations are mutually dependent. Why does Hegel think this? (This is an intuitive example, not used by Hegel.) Suppose you lend someone five pounds. A number of facts follow from this on a Hegelian view: (i) you have a right to the return of the five pounds, (ii) the borrower has an obligation to return the five pounds, and (iii) the right and the obligation arise simultaneously, being both constituted by the lending of the money having taken place.

Hegel thinks the relationship between right and obligation (in this kind of context) is what he calls "identity-in-difference." This means (i) rights and obligations are different and (ii) rights and obligations are identical. A paradox: 'identical' could mean qualitatively or numerically identical. If rights and obligations were numerically identical, rights would be obligations. If they were qualitatively identical, this would mean rights were *like* obligations. Hegel could also be read as claiming that rights and obligations are mutually dependent.

There are difficulties with the example: What exactly constitutes *lending*? The

mere act of *giving* five pounds could also constitute an act of generosity, or foolishness. What is *giving*, over and above the change in spatio-temporal location of (say) the five pound note? This shows the difficulty in making inferences from a physical event to abstract realities such as rights and obligations. Suppose someone thought there are no such things as social or political realities which entail *value*. (They could state, like Wittgenstein at the beginning of the *Tractatus Logico-Philosophicus*, that 'the world is all that is the case' i.e. the world is the totality of facts and no more.) How can there be room, on such a view, for obligations, rights, wrongs, the *contractual*? Unfortunately, things we believe in often disappear under philosophical scrutiny. For example: God, space, time, the self, matter, society. Rights and obligations do not seem to be an exception. They are analysed away by philosophical analysis.

Hegel believes that (in this kind of case) the establishment of right and obligation, and distinction between them, is through the *lending* of the money. A single relationship was established. The relationship between rights and obligations is like that between shape and size in the case of physical objects, or that between up and down in the position of physical objects. Rights and obligations are intimately connected, *essentially* connected. Identity-in-difference entails a relationship of mutual dependency.

Of course, the particular cases could all be questioned, For example: You could say there *can* be size without shape, as in the case of *numbers*. There can be shape with no size, in the case of geometrical figures. Particular counter examples then cast doubt on the global thesis, the applicability of dialectical reasoning. You could start to unravel the idea of identity-in-difference.

Now I'll just say something about this distinction between phenomenology and logic, in Hegel's system. Remember when I was talking about Kant I said that it's often claimed that Kant reconciles or synthesises rationalism and empiricism, in so far as it's true that there are two discernible movements, or tendencies, in the two hundred years of European thought before Kant wrote, one of them called 'empiricism' and the other 'rationalism'. In so far as that's right, it seems to be true that Kant synthesised or combined them. At least, Kant thought so. But it seems to me that in drawing the distinction between phenomenology and logic Hegel has *reintroduced* the distinction between rationalism and empiricism. Empiricism is the phenomenology and rationalism is the logic.

Now, this remark is one which I'm sure would have appalled Hegel because he thinks of himself as a *synthetic* philosopher or, paradigmatically, somebody who produces speculative syntheses in philosophy. But I'm saying that one of the greatest syntheses of all is never achieved in Hegel's philosophy.

(2) Locating the Master and Slave Dialectic

Hegel made two attempts to write his system, in roughly 1807-1815, and roughly 1815-1831. Each involved attempts to write phenomenology and to write dialectic.

The earlier period includes the *Phenomenology of Spirit* (1807) and the *Science Of Logic* (1812-14). In the later period, Hegel wrote a series of works he calls the *Encyclopaedia of the Physical Sciences*. Three important volumes are: the *Logic* (or *Lesser Logic*), *Philosophy of Mind*, and *Philosophy of Nature*.

What is phenomenology? In 'Phenomenology of Spirit', 'Spirit' is a translation of the German word *Geist*. *Geist* can also be translated 'mind', 'consciousness', 'essence', or 'morale'. *Geist* can be translated as any or all of these of these. 'Geist' has theological and supernatural, as well as humanistic, connotations.

Phenomenology is *the description of what appears to consciousness, the description of appearances*. We obtain radically different Hegel's depending on how we translate or interpret 'Geist'. This ambiguity is part of the richness of Hegel's thought and means it finds disparate applications, in politics, in theology, in philosophy of history, in understanding conflict, in understanding self-consciousness.

The young Hegel, of the 1790s, was interested in theology. The older Hegel, authoring, say, *Philosophy of Right* (1821), seemingly abandons theology to focus entirely on philosophy. (Interestingly, Hegel, in that book, does not take it for granted that ethics is prior to politics.) However, we can often read the work of the later of Hegel as carrying theological entailments, whatever Hegel's professed views.

Hegel has a sophisticated theory of self-consciousness which is also complex, obscure and difficult to understand. Three of Hegel's major works each contain a section wholly devoted to the subject. These are *The Phenomenology of Spirit* or 'Jena Phenomenology' of 1807, *The Philosophy of Mind* and the Berlin *Phenomenology*. The latter two books form part of the so called 'Encyclopaedia of the Philosophical Sciences' 17/27/30, a series of writings which Hegel regarded as the most definitive exposition of his system and to which the badly written and tortuously argued 'Jena Phenomenology' is a prolegomena.

Not only does the topic of self-consciousness occupy a great deal of space in the Hegelian texts but the concept is a key one for understanding Hegel's philosophical system. Here our primary concern is with what light Hegel can shed on the concept of self-consciousness and not with what light that concept can shed on Hegelian metaphysics but, it is in fact almost impossible to untangle what Hegel says about self-consciousness without making frequent reference to the system in terms of which he intended the concept is to be understood.

In brief outline, then, this is the Hegelian system. Reality for Hegel is ultimately and fundamentally mental or spiritual. In this sense it is true that Hegel is an Idealist - not in the sense that he denies that there are physical objects or that he thinks the objective world is constituted by human sense perception. None of this is true of him. He is an idealist in the sense that Marx is a materialist. Marx allows that persons are conscious but holds that reality is ultimately material, Hegel that there is matter but that reality is ultimately mental. This ultimately spiritual realist in Hegel is called 'Geist', which is the German for 'mind' or 'spirit'. *Geist* [is not] equated with God in Hegel's system and, in the interest of Protestant orthodoxy, he rejected the label 'pantheist' for his philosophy. Despite these disclaimers, *Geist* is best

understood as a sort of all pervading aspect of reality; reality in the sense of 'everything there is' or 'the universe as a whole'. Geist is infinite according to Hegel, not in the sense in which the number series is infinite but infinite in the sense in which there isn't anything which isn't it. Geist is all pervading.

Geist, or the cosmic spirit, has two important features. Geist is embodied and Geist is conscious. Geist is embodied in nature. 'Nature' is Hegel's word for the physical universe that is found in everyday sense perception and investigated by the natural sciences. It is essential to note at this point that the embodiment of Geist does not commit Hegel to a mind - body dualism. The embodiment of Geist is to be understood on Aristotelian rather than on Cartesian lines. Nature (matter?) is the expression of Geist - the form or shape of the cosmic spirit. Geist only exists in so far as it is embodied and nature only exists in so far as it is the embodiment of Geist.

The relation between Geist and nature exhibits a recurrent theme in Hegel's philosophy - that of 'identity-in-difference'. This can be stated briefly as follows; (i) If x is a necessary condition for y and if (2) y is a necessary condition for x (either logically or causally) then x and y are in this sense 'identical' - they are two aspects of one reality. Notice that Hegel is not involved in a breach of the logical principle of identity here - he is not saying 'x is y' or 'y is x' he is saying x and y are parts of a single larger whole (which as aspects they may exhaust). This reconciliation of seemingly incompatible perspectives is an important part of Hegel's philosophical technique and is known as 'dialectical reason'.

The other major feature of Geist is consciousness. Geist is conscious according to Hegel, to the extent and only to the extent, that persons are conscious. It is clear, then, that human existence is of major importance to Hegel's metaphysics. The point can be put this way - persons are conscious and persons are part of the universe; therefore the universe is conscious. Not only this, but to the extent that persons are conscious of nature, of the physical world, the universe is conscious of itself. Person's consciousness of ????????? is Geist's self-consciousness. It is at this point in Hegel's metaphysics that the central importance of self-consciousness can be seen.

His view is in fact that history, or reality in time, is the process of Geist's coming to self-realisation. Geist has posited his own ignorance. To the extent that persons are ignorant so is the universe ignorant of itself. The historical process is precise the process of the universe recovering its knowledge of itself.

As is in keeping with what has been said so far, human experience in history is the 'cutting edge' of this cosmic process. All the products of human action all the artefacts, political and social institutions are, to greater or lesser degree expressions of Geist's progress to self-knowledge. Different historical phases have been better or worse expression of Geist's progress and Hegel's, often tremendously rich analyses of art, religion, law, politics, social institutions and the history of philosophy tend to chart this progress. It was Hegel's perhaps over optimistic and pretentious view that Geist was nearing self-realisation in his own day and that his own philosophical system was the most appropriate and complete expression of this progress. The final stage of history is called 'Absolute Knowing' by Hegel and this is the title of the final

chapter of PhG.

Such then, in barest outline, is Hegel's system. The view I have given is a very metaphysical one; one that sketches Hegel's view of reality as a whole. This is not to deny that there are other readings of Hegel. It is possible, for example, to present a very 'humanistic' Hegel based on the consciousness of individual persons, or a view of Hegel the social and political theorist, Hegel the Philosopher of History, Hegel the pre-Marxist, Hegel the first Existentialist etc etc. Much of what he says about self-consciousness is incomprehensible though if his metaphysics is neglected.

Armed now with an overview of his system and a statement of what dialectical reason consists in we can turn to Hegel's theory of self-consciousness. As is frequently the case, the contents pages of Hegel's books are a useful guide to the structure of his thought. The headings for 'self-consciousness' in the 'Jena Phenomenology' are;

B. SELF-CONSCIOUSNESS

(iv) The Truth of Self Certainty

- A. Independence and dependence of Self-consciousness; Lordship and Bondage.
- B. Freedom of self-consciousness; Stoicism, Scepticism, and the unhappy consciousness.

The equivalent section in the Berlin Phenomenology is as follows;

B. Self-Consciousness

- 1) Characterisation
- 2) Freedom
 - i) Abstract
 - ii) Free Self-Certainty
 - iii) Objectification
- 3) Division of Self-Consciousness
 - a) Desire
 - 1) Drive and Activity
 - 2) Self seeking and destruction
 - 3) Satisfaction and Identification
 - b) Recognitive Self-Consciousness
 - 1) Struggle
 - i) imperiousness
 - ii) life and death

- iii) coercion and subjection
- 2) Mastery and Servitude
- 3) Community of Need
 - i) slavery
 - ii) service
 - iii) communal provision

- c) Universal Self-Consciousness
 - 1) recognition
 - 2) acceptance
 - 3) rationality

and in the Philosophy of Mind;

- (b) Self-Consciousness
 - a) Appetite
 - b) Self-Consciousness Recognitive
 - c) Universal Self-Consciousness

Synthesising Hegel's views of self-consciousness in these three works, a common structure can be extracted. This structure can be understood in two different ways' (1) as the charting of a chronological development of self-consciousness and in the individual and the universe or (2) as the description of the logical structure of self-consciousness; an outlining both of the preconditions for self-consciousness and an explication of what self-consciousness is. Hegel no doubt intended the two modes of understanding to be complementary - History is a chronological progress of universal self-consciousness for Hegel, yet, in a sense, the universe is only becoming what it really is by this process. The goal is implicit in the origin. Thus the second explicative interpretation is equally valid as a conceptual and dialectical unpacking of the notion of self-consciousness.

The structure which can thus be understood is threefold - each of the three stages marking either a 'moment' or component aspect of self consciousness (interpretation (2) - or else a chronological stage in a developing self-consciousness interpretation (2). The three stages are as follows:

1. Self-Certainty.
2. Master and Slave Dialectic.
3. Universal Self-Consciousness.

This threefold structure broadly parallels his treatment of 'consciousness' and there is a common theme; in both explications we are offered to begin with the most rudimentary and commonsensical level of understanding the concept and this is gradually shown to be inadequate. Thus the analysis of 'consciousness' began with

‘sense certainty’ in the PHG - set of utterly taken for granted beliefs about the world found in naive realism and sense perception. In a parallel way, Hegel begins the explication of self-consciousness with ‘self-certainty’.

In the *Phenomenology of Spirit*, there is a crucial middle chapter, the so-called ‘master-and-slave dialectic’. Here, Hegel argues that *self-consciousness is essentially social*. Without a society, there can be no self-conscious beings. This seems a strange and radical claim. Self-consciousness seems very private, individual, and intimate. However, Hegel argues that it depends upon the existence of an external world, and of other conscious beings in the external world. It also depends upon an antagonistic conflict or struggle for recognition with conscious beings in the external world.

By saying this, Hegel politicises self-consciousness. He has brought about an enormous break with the philosophical ‘tradition’ before him. Before Hegel, people's views of self-consciousness were more or less Cartesian. In Descartes, there is no suggestion that a conscious being becoming conscious requires any encounter with another conscious being. The Cartesian holds that if he believes he is in a mental state, that belief must be true. There is no room for introspective error. The self-conscious mind is translucent; there is no unconscious mind. Cartesian self-consciousness is individual and atomic: It pertains to individuals one by one. Hegel argues against this concept of self-consciousness not only by stressing the importance of conflict with other people for self-consciousness, but by ultimately arguing that self-consciousness is cosmic and *one*.

The Master and Slave passages in Hegel have been massively influential on later philosophy. It influences Marxism in the notion that consciousness is class consciousness. The bourgeoisie or capital owning class is the master, the proletariat or working class is the slave. In the revolution advocated by Marx and Engels the master becomes the slave and the slave the master, an inversion anticipated at a more abstract level by Hegel. The existential phenomenology formulated by Sartre in *Being and Nothingness* (1943) and by Merleau-Ponty in *Phenomenology of Perception* (1945) implies that self-consciousness depends upon the body in a constitutive encounter between self and other. Self-consciousness is consciousness of the self, but the very awareness of self, the carving oneself out from the world, depends on perceiving the body of the other.

That some kind of distinction is marked between self and other beings is fundamental to social and political thinking. Sartre and Merleau-Ponty argue that conflict between people produces the kind of self-consciousness that each of us has. A being who had never encountered an other could not be self-conscious.

Whether or not Wittgenstein actually read much Hegel, the conclusion of his private language argument is essentially Hegelian. There is no solipsistic or utterly individual self-ascription of mental predicates. In what became *Philosophical Investigations* (1952) Wittgenstein argues that (i) language is public (languages depend upon rule following by a community of speakers), ii) understanding requires language, and therefore iii) understanding one's own mind is something public. Introspection, the seemingly quiet and private moments of our lives, depends upon a

public framework of rule following. As in Hegel, the seemingly inner and individual depends on the outer and social.

Peter Strawson has probably not been influenced by Hegel. (He would not have enjoyed Hegel's style.) However, the conclusion of Strawson's *Individuals: An Essay On Descriptive Metaphysics* (1959) chapter 4, 'Persons', is Hegelian. Strawson claims that a necessary condition for ascribing a state of consciousness to oneself is that one is able to ascribe that state of consciousness to other beings. You can only meaningfully say 'I am depressed' if you could also say 'He/she is depressed'. The concept of depression cannot be acquired solipsistically. Psychological concepts span the gap between undergoing experiences and observing behaviour. The distinction between mental and physical predicates depends upon the concept of a person. 'Person' is logically primitive with regard to the distinction between 'mind' and 'body'.

There are three main sections of the *Phenomenology of Spirit*: 'Consciousness', 'Self-Consciousness' and 'Absolute Knowledge'. The first section, 'Consciousness', is an examination of empiricism (the doctrine that the world is as we perceive it through our senses), rationalism (the doctrine that the world is as we think it to be) and scientific thought (understood by Hegel as a synthesis of empiricism and rationalism). 'Self-Consciousness' contains the Master and Slave Dialectic. 'Absolute Knowledge' is about existence as a self-conscious totality: a passage with a mystical quality. Thought and Being ultimately *coincide*, according to Hegel. This means that in Absolute Knowledge there is no difference between what is *being* what it is and what is *knowing* what it is. Because reality is a kind of self-consciousness, its self-knowing is its being and its being is its self-knowing.

There are both humanistic and theological readings of Hegel. There are two importantly different interpretations of *Geist*: (i) *Geist* as something that would be removed from the world if humans were removed, roughly, human consciousness and its expression, or as (ii) something transcending humanity: God. Hegel claims that *Geist* resists scientific explanation. Nations are distinguished by their mentalities. *Geist* can mean the sequence of historical mentalities. For example, Hegel distinguishes between the Greek, the Oriental and the Germanic 'worlds'. *Geist* in the humanistic sense is comparable to Marx's concept of *praxis*. Marx asks what distinguishes humans from animals. He offers the answer: labour. *Praxis* (the Greek word for 'action') is the manipulation of matter by humans to create houses, roads etc. Humans alter the world to suit their ends. They then live in a human world, a human-manipulated world. Is *Geist* God? *Geist* does not seem reducible to its human expression. On the other hand, Hegel argues *Geist* cannot exist without its expression. (For example, there is no purely theoretical legislation; there has to be real legislation.) It looks, then, as though *Geist* cannot be God in any straightforward orthodox sense. In the great monotheistic religions, God creates freely. Hegel's *Geist* creates essentially. *Geist* is free to create what it does create, but *Geist* cannot not create.

(3) Self-Certainty

The chapter 'Self-Consciousness' begins with the heading 'The Truth Of Self-Certainty'. What follows is an attempt to answer the question: *What is the fundamental relationship between the individual and the social?* It implies the sub-questions: *How do I demarcate myself from the world as an individual?*, and *How do I demarcate myself from other individuals, as a member of society? How is it possible for there to be individuals? How can an individual be oneself?* Hegel attempts to establish how there can be individuals and societies.

In the *Phenomenology of Spirit*, Hegel uses the word 'consciousness', to mean 'being that is conscious'. The word should be understood in an abstract way. Hegel uses the term 'self-consciousness', to mean 'that which is self-conscious'. A self-consciousness could be a human being, or a group of human beings, such as a country, state, or club. The unit is left deliberately vague. Hegel hopes the model he advocates will have diverse applications. The notion of being a self-consciousness is not one we can avail ourselves of at the *beginning* of the Master and Slave Dialectic. Hegel is asking: How is self-consciousness possible?

Hegel argues *being* a self-consciousness requires being *accepted as* a self-consciousness. This claim may be understood on many levels. *Being one* presupposes being *recognised as one*. For example: Being a student at the University of Oxford requires being accepted or recognised as a student of the University of Oxford. Being a political state requires being accepted as a state by other states. This is true of any organisation the existence of which entails self-consciousness. Being something is not intrinsic but rests on political presuppositions to do with relations with other things of the same type.

There are stages through which a being becomes self-conscious. Understanding this involves explaining some of Hegel's paradoxical remarks. Look at p. 104 of the Miller translation of the *Phenomenology of Spirit*:

'Consciousness makes a distinction, but one which at the same time is for consciousness not a distinction.'

Hegel thinks this statement encapsulates self-consciousness, expresses its essence. *What* distinction is drawn here? The distinction is between subject and object, where the subject is *that which is conscious*, and the object is *that which consciousness is consciousness of*. In the case of self-consciousness, consciousness is both subject and object. For this reason, Hegel says that the distinction is for consciousness *not* a distinction. The distinction that consciousness draws *within* self-consciousness is only a distinction within itself: *What it distinguishes from itself is only itself*. If it distinguishes itself from itself, in a sense, this is a distinction that cannot be drawn. Anything is necessarily self-identical. On another level, there is an epistemological distinction, but not an ontological one. On another level, in self-consciousness,

consciousness is presented with consciousness but, phenomenologically, or strictly speaking, it is not thereby presented with the *distinction* between that which is conscious and what consciousness is consciousness of. Again: Consciousness draws a distinction which, *for consciousness*, is not a distinction. On the next page, Hegel says (p. 105): 'What it distinguishes from itself is only itself'. The distinction does not appear to consciousness. In this sense it is false that the putative distinction is a distinction. Consciousness is not presented with a distinction because, within consciousness, consciousness is presented only with itself as object.

Hegel claims that only a being that feels desire can be self-conscious. Desire is a necessary condition for self-consciousness. Why does Hegel think self-consciousness presupposes desire? Could there not be self-conscious beings without desire? Commentators find it hard to explain Hegel on this issue.

Whatever is self-consciousness is conscious of itself: conscious of being, and conscious of being what it is. Self-consciousness therefore presupposes a self/not-self distinction. There cannot be self-consciousness without whatever it is that is self-conscious, carving out a 'me/the-rest-of-the-world' distinction. Self-consciousness distinguishes two mutually exclusive and jointly exhaustive portions of the universe: the part that *one is* and the whole of the remainder, which *one is not*. Hegel thinks the self/not-self distinction can only be established by *desire*, because desire establishes an *asymmetry* between one's own state of mind and the rest of the world: on the one hand, the state of the desirer, a state of lack, and, on the other hand, a state of the world as possible source of fulfilment. This is an interesting and insightful idea. However, we need to know the strength of this presupposition. In what sense of 'impossible' is it impossible for a self-conscious being to not experience desire? It seems hard to rule out a priori the drawing of a self/not-self distinction based on say, where pain is or is not felt, where thought does and does not arise, or where freedom can and cannot be directly exercised.

Hegel thinks desire establishes the self/not-self distinction. In some sense of 'sufficient', desire is not only necessary but sufficient for self-consciousness at a fairly rudimentary or abstract level. Because there are conflicts in desires, self-consciousness presupposes conflict. Although desire seems very private, intimate and introspective, desire presupposes something as public as the conflict between desires. That desire is a prerequisite for self-consciousness is another reason why self-consciousness will turn out to be political. The clash of desires operates at every level. If A desires that *p* and B desires that *not p*, their desires are irreconcilable, irrespective of any particular ontology.

Desire which sets up a mismatch between one portion of the world and the rest of the world. It establishes a lack in one's own being and the possibility of fulfillment in the rest of Being.

(4) Master and Slave

'Master and Slave' is section A of the chapter 'Self-Consciousness'.

How is self-consciousness possible? How can there be individuals? Hegel believes self-consciousness depends on the existence of a society. Hegel asks how the distinctions self/not-self, subject and object, can be drawn.

The self/other distinction is the distinction between me and somebody else, or perhaps me and anybody else. It presupposes the self/not-self distinction, but the self/not-self distinction does not presuppose the self/other distinction. This hierarchy is presupposed by the establishment of the individual and the social.

How is the self/other distinction established? Self-consciousness presupposes an encounter with *another consciousness*. Why does Hegel think this? Could there not be a solitary self-consciousness? An animal or a computer could develop self-consciousness in isolation, surely. Hegel believes that in order to be self-conscious it is necessary to be conscious of oneself *as x*, where *x* belongs to some type. To put it another way (following a thought of Wittgenstein, Elizabeth Anscombe and Jonathan Bennett) being conscious of oneself is being conscious of oneself *under a description*, or some designation or other. You are not conscious of yourself *full stop*. You are conscious of yourself *as a man, as a woman, as a thinker, or as a human being*, and so on. On this view, there is either *no such thing* in the human case as pure self-consciousness, or, if there is, this is only possible if there is consciousness of oneself under a description. Self-consciousness is always contaminated and interfered with by language, by description. Subsuming oneself under some concept or other is a necessary condition, or prerequisite, for being self-conscious at all.

These points are highly controversial. We live in a Kantian age in which emphasis is placed on exposing conceptual preconceptions: Perceiving or thinking of anything involves, consciously or not, interpreting it. However conditioned, or linguistically and pragmatically structured, our experience might be, it has not been shown that there might be, to take two seemingly very different cases, Buddhist disclosure of emptiness, or 'no-mind', or a disclosure of the soul to itself. To follow Hegel, we have to at least provisionally endorse the neo-Kantian thesis that either all experience is conceptually structured or there cannot be any experience unless some experience is conceptually structured.

Hegel raises the question: What is the origin of these concepts we deploy in being self-conscious? If being self-conscious entails being capable of the thought 'I am *x*', we want to know how the concept of *x* originates. Hegel thinks these concepts have a third person, public origin. We learn them from encountering other conscious beings. Crudely put, self-consciousness requires the possibility of thinking: 'I am one of those'. This kind of thought requires the existence of the public. Hegel is claiming two things: (i) *I am not* the other, but (ii) *I am like* the other. I am numerically distinct from any other conscious being. I am qualitatively similar to any other conscious being. Self-consciousness presupposes that (i) I am capable of thinking that I am numerically distinct from other people, and (ii) I am capable of thinking of myself as qualitatively similar to other people. This is part of the account of why self-consciousness presupposes others: a *conceptual* derivation through the encounter with the other.

In the case of self-consciousness, *being* one requires being *accepted as one*. Being conscious of oneself as a man requires recognition or acceptance by other men. It is not possible to *privately promote oneself* to such a role through a kind of inner freedom, a first person singular *fiat*. We do not have the power to choose our own identities in this atomistic, private or introspective manner. Despite the sections on the body, and being-for-others, in *Being and Nothingness* Hegel seems to be taking a view diametrically opposed to Sartre's existentialism, an extreme form of individualism. The early Sartre thinks I define myself as what I am through my free choices. Hegel claims that it is not possible to freely and individually *turn oneself into* a particular being. Both being x and being conscious of oneself as x requires acceptance by other people. (The English word 'bother' includes the word 'other.')

The master and slave dialectic is a struggle for recognition by consciousness. At a sociological level, self-consciousness are being subverted and confirmed by one another throughout the human world. There is a struggle for recognition between anything which can be thought of as entailing consciousness, or being emergently self-conscious: individuals, classes, countries. Personal identity, in the sense of who or what one is, is not private. One's own personal identity is in the hands of the other. It follows that people, simply in virtue of being self-conscious, have a hold over one another. They are a threat to one another. One consciousness might deny recognition to, or withhold acceptable from, another consciousness. These social and political facts are presupposed by self-consciousness. They are more explanatory of self-consciousness than for example, physical facts.

The Master and Slave Dialectic takes the form of a fable, a stylised, abstract, story designed to instruct us in the politics of self-consciousness. Hegel says that in an encounter between two consciousnesses, one is always dominant and the other subservient. There is a struggle for power and recognition, in which one consciousness is master and the other slave. In the struggle for recognition, because there is a power struggle between master and slave, the master threatens the slave with death. At this moment, there is a liberation of the slave and an enslavement of the master. The master is the slave and the slave is the master. The master realises that if he kills the slave he will no longer be a master. A master is essentially *master of*. Hegel describes the complex relations between master and slave, dialectically and phenomenologically. The notion of identity-and-difference applies to the relationship. Although the slave is oppressed by the master, the slave has power over the master's identity *as master*.

As a consequence of the threat of death, the slave is liberated in the sense that he becomes more fully self-conscious in the face of his doom. Look at pp. 117-118 in the Arnold Miller translation. Facing death is *necessary for self-consciousness*. This again seems an extraordinary claim. The commentators make little sense of it. Not all of us have faced death, yet we seem to be self-conscious. This seems to show straightforwardly that risking one's own life is not any kind of necessary condition for being self-conscious. What does Hegel mean?

Self-consciousness is *consciousness of one's own existence*. But this is only

possible when contrasted with *consciousness of one's own non-existence*.

Dialectically, Hegel argues we can only have a concept of our existence by having a concept of our non-existence. Rather as it would be hard to imagine someone understanding the concept of 'here' but not that of 'there', or 'me' but not 'you' (or 'he', 'she' 'it'), in the same way, someone with an understanding of existence must also have an understanding of non-existence. You might say, because only the concept of one's own non-existence is necessary for self-consciousness the *real* risk of death is not necessary. Theoretical knowledge, or knowledge of the bare possibility of not existing is good enough. Hegel does not think this is right.

There is ordinary everyday self-consciousness and existential self-consciousness. The latter is a state of being aware, even shocked by one's own existence: so to speak, *feeling it in one's bones* that you exist: a stark and immediate insight. This kind of experience is arguably caused by the threat of death and, perhaps, only by the threat of death. This is a second way of understanding Hegel's claim that self-consciousness presupposes awareness of death. One feels one's own existence more strongly when threatened with extinction. By threatening the life of the slave, the master has thus subverted his own power: rendered himself less self-conscious than the slave.

The slave is not killed in the fable, but the master attempts a new strategy for repressing the slave's identity, and making his own identity more secure. He makes the slave manufacture commodities for his master. This strategy also fails, because in making objects the slave expresses his own consciousness in material form. Although the objects are for use by the master, they express the slave's consciousness. The slave perceives his self-consciousness reflected in matter: a form of exteriority not available to the master.

The master realises he can never eliminate the slave's freedom entirely. The master has an ideal of freedom, which is that the slave would *freely* obey the master. In democracy, the slave freely obeys the master. In totalitarian societies the slave obeys the master without the option of not obeying. However, that the slave obeys the master *freely*, entails the possibility of the slave not obeying the master. Supposing somebody coerces somebody into a sexual relationship. In this case, the person being abused does not have the freedom not to be abused. But suppose the abuser has it as an ideal that the other would *choose* to enter into the sexual relationship. If that were ever to come about, the possibility of refusing the sexual relationship would mean that the master was no longer the master: the master's mastery is in constant jeopardy. A master in constant jeopardy is not fully a master. Hegel is saying there is a kind of paradox in the situation of the master: He wants to be obeyed freely, but this entails *the very real possibility of not being obeyed*. This is, perhaps, a power structure prevalent in human affairs.

The Master And Slave Dialectic ends with the master becoming the slave and the slave becoming the master, a rather pessimistic conclusion because presumably the process begins over again.

The master and the slave are both self-consciousnesses. A self-consciousness could be a man, an animal, perhaps a computer, a society, a country. Hegel thinks of

groups and communities as self-conscious in the way we would normally think of individuals as self-conscious. Why? Arguably, communities behave more like self-conscious individuals than, say, like physical objects. Arguably, communities are made out of conscious individuals but, if Hegel is right, we should not be naïve about the presuppositions of either ‘conscious’ or ‘individual’ here. Even if societies are not ‘constituted by’ individuals, or we cannot avail ourselves of ‘individual’ without ‘society’, the communities and institutions to which we could apply Hegel’s phenomenology seem to *entail the existence of* self-consciousness. I say ‘seem’ because there could be a community of computers or robots engaged in a master and slave dialectic but wholly devoid of consciousness. The relationship between Hegel’s philosophy and artificial intelligence has yet to be explored.

Turn to page 110 in the Miller: ‘A self-consciousness exists for a self-consciousness’. The notion of *existing for* expresses not just consciousness if but also the notion of existing entailing a prerequisite:

If *b* ‘exists for’ *a* then if not *a* then not *b*. Another item of vocabulary: the distinction between being *in-itself* and being *for-itself*. Consciousness *in-itself* is consciousness as it is intrinsically, or in abstraction from its relation to others, but consciousness *for-itself* is consciousness directed towards objects.

Hegel casts doubt on our notion of self-consciousness as atomic and entering only occasionally into power struggles with other beings. The conflict between consciousnesses makes possible their transition to being self-consciousnesses. If without this conflict, the transition is impossible then the conflict, if not perennial is inevitable if there are to be any self-consciousnesses.

Geist: If we do not read this term theologically, we can understand it as consciousness in general or human consciousness. Conceptual schemes attributed to nations, groups etc. are then produced through conflict.

(5) Time and Subjectivity

Is time political? To what extent is time part of the objective world and to what extent a human construct?

Time is not a simple concept. It can be analysed into at least the following five logically related concepts: (i) duration, (ii) before, simultaneous with and after, (iii) past, present, future, (iv) beginning and ending, (v) change. We could distinguish natural from clock time: Natural time is time as measured by natural events, i.e. night and day, the seasons etc. Clock time is the division of days into hours, of hours into minutes, and minutes into seconds. It depends on a mechanical, or at least, an intellectual device. Clock time seems to have a monastic origin. (The word ‘clock’ is derived from the French ‘cloche’, meaning ‘bell’.)

Hegel raises the question: *Is consciousness in time or is time in consciousness?* To say that consciousness is in time appears common-sensical and, at least as far as the Newtonian and Einsteinian paradigms are concerned, scientific, because before any consciousness existed, presumably there were aeons of time when life existed,

before that aeons of time when there was no life etc. Presumably if there was some cosmic catastrophe that destroyed all conscious beings, there would then be aeons of time when the planet would continue to exist without consciousness.

On the other hand, Hegel thinks there are very good reasons for thinking that *time is in consciousness*. The past does not exist because the past does not exist now, he claims. Similarly, the future does not exist because the future does not exist now. Historical events did exist, but they don't; they have elapsed. The past is no longer; the future is not yet. Only the present exists; the time is always now. A curious phenomenon. Why is it always now? Perhaps 'now' means 'simultaneous with me'. However, is language so important here? Could definitions make it the case that now is the time it is? Could language make it the case that now is only ever the time it is (if 'only ever' is aloud some kind of meta-sense here)? How could language acquire such metaphysical power?

Hegel feels that this phenomenological observation (an observation based on what appears to consciousness) suggests that time is within consciousness. In particular, the past does not exist outside memory, and the future does not exist outside prediction or conjecture. Of course, the operations of memory and conjecture take place now, in the present. Conjoining the commonsensical or historical or Newtonian notion, with the phenomenology, Hegel concludes both that *consciousness is in time*, and *time is in consciousness*. How can both these claims be true? Normally, we think that 'a is in b' rules out 'b is in a'. An example of dialectics; opposites must be reconciled. What does 'in' mean in 'in consciousness' or 'in time'? 'In consciousness' means 'appears to consciousness' or 'is a phenomenon'. 'In time' means 'temporal' or 'rightly subsumed under temporal predicates'. Once we notice the difference in the concepts expressed by 'in' the tension, if not dissipated, is less. Dialectic rests on ambiguity.

Consider the concepts making up the concept of time. It can be argued that past, present and future belong to consciousness, or are at least part of human reality. A past is *someone's* past, a present is *someone's* present, and a future is *someone's* future. This raises the question of whether there can be objective time if past, present and future are introduced into the universe by human existence. Commonsensically, objective time is real, because events before consciousness existed, ordered by a before-after relation, or in an 'objective' time sequence. 'Objective' here entails 'irrespective of how we think them to be' or even 'irrespective of whether there are any minds. Conversely, an objective time sequence does not exist as past, present and future: The tensed ordering of events depends on human reality. Saying this, we can mark a distinction between natural/objective time and subjective/human time. The first includes time sequences and *before-after relations*; the latter includes *past, present and future*.

Clock time is superimposed on natural time by human conceptual schemes. A conceptual scheme is a mental map by each of us locates him or herself in time. This scheme includes a distinction between 'before now' and after 'now', as well as 'today' and perhaps a mental image of the days of the week, and months and years. It

seems that certain components of this conceptual scheme are political. In modern capitalist society, time is bought and sold as a commodity. A person in employment receives their salary not only for work, but also for time spent. Time is bought. We talk of "spending" and "wasting" time. It is a commodity not in the sense of being a physical object or a set of physical object, but in the sense of being measured according to a human conceptual scheme. It is divided into units so it can be bought and sold.

The wearing of a wristwatch can be thought of as political. Certainly two hundred years ago, it would have been peculiar for someone to wear a wristwatch and check the time regularly. A way of being is incorporated into political time.

In *Being And Time* (1927), Heidegger bases his arguments on Hegel's view of time. He also introduces a crucial new concept, *being-toward-death*. As humans, we obviously have a consciousness of our own mortality. Being-toward-death is more than this. Heidegger is writing at the level of Fundamental Ontology (not psychology or even epistemology). Being-towards-death is the *ever present* trajectory of human existence: an existence which entails the perpetual possibility of extinction, being someone in a way that implies the inevitability of being nothing, being in a way that implies not being. This awareness gives one a concept of one's life as a whole. The self-conception of a life as 'my life' is provided by consciousness of death. Taking this awareness of our own finitude, we can understand why in capitalist society time is so valuable. In giving one's time for wages, one is giving a portion of one's life-time. This is possibly one of the things that makes work unpleasant.

Consider the word 'deadline'. It is tempting to suppose that the word is felt as something psychologically oppressive because it is felt as a mini-rehearsal of one's own death, the big D; not because of the word but because death is the last deadline.

Conspicuous differences between the Mayan calendars developed in Yucatan between 400 and 1400 AD, and the much earlier calendars of the ancient Egyptians (living around 1500 BC). Are we political prisoners of time? Human time is thoroughly political. But is there a kind of human time that is not political? Hegel and Heidegger think there is.

Is there existential time? Heidegger says there is a kind of time that is 'primordial' with regard to the rest of time; it is presupposed by the rest of human time. Hegel anticipates Heidegger's notion of primordial or existential time. In his 1812-14 volumes the *Science Of Logic* (which include little science or logic) Hegel employs dialectical reasoning to arrive at the conceptual scheme presupposed by all conceptual schemes. Obviously there are radical differences between communities as culturally removed as those of 21st century England and the American Hopi Indians. Hegel tries to arrive at the most essential and universal of concepts. He is not naive as a historian. Hegel arrives at two fundamental concepts, Being and Nothingness. It doesn't matter whether you are an ancient Greek or a medieval Hindu, you won't be able to make much sense of anything if you cannot draw a distinction between being and not being. Unless you have the concepts of being and nothing you do not have any concepts.

Hegel says the relationship between being and nothingness is dialectical. The two are opposites. They are also identical. At this point many philosophers of a fairly logical or scientific frame of mind simply close the book. Being and nothingness are opposites semantically. Psychologically, to think of being is not to think of nothingness, and vice versa. Ontologically, being is filled and nothingness is empty, as Sartre would say. Yet the two are mutually dependent as concepts. Quite often, when Hegel writes 'identical' (*identisch*) he means 'mutually dependent': If not *a* then not *b* and if not *b* then not *a* but *a* is not *b* in the sense of 'numerically identical with' *b*. Here, however, he does seem to mean 'numerically identical'. The attempt to think nothingness results in thinking being, and *vice versa*.

Because being and nothingness are antithetical concepts, they have a synthesis or a reconciliation. This is *Becoming*. Both Heidegger and Hegel think that becoming is the essence of existential time. In the last ten pages of his book, Heidegger endorses Hegel's notion of becoming.

What is becoming (*Werden*)? Because it is the synthesis of Being and Nothingness, it *fuses* the two. In Hegel's thought, becoming is the *movement* or the *oscillation* between Being and Nothingness. In becoming, there is beginning to be, and there is ceasing to be. By beginning to be is meant absolutely beginning to be, or beginning to be *ex nihilo*. By ceasing to be is meant absolutely ceasing to be, becoming nothing, as opposed to changing into something else. It is counter-intuitive that the ideas of absolutely becoming to be and absolutely ceasing to be could be reconciled. But Hegel and Heidegger claim to be investigating a kind of time more fundamental than the political time we have been conditioned to think about; we must expect strange conclusions.

When I speak, there is beginning to be and ceasing to be at every moment. Sounds come into being and go out of being constantly. Beginning and ending are constitutive of becoming. In beginning there is a transition from nothing to becoming; in ending there is a transition from becoming to nothing.

Hegel and Heidegger do not mean *x* begins, lasts a certain period of time, and then ceases to exist. In the case of becoming, there is no duration between the coming to be and the ceasing to be. They have to be thought of as simultaneous. Otherwise it would not be a case of becoming.

In *Being and Time*, Heidegger says that becoming, in precisely this transitional sense, is what our existence consists of; mentally, physically, in every respect. This is primordial or existential time is what Heidegger calls 'temporality'.

An interesting political question: Could a society survive the destruction of its means for measuring time? Could 21st century Britain survive the destruction of all clocks? A thought-experiment would be revealing of the importance of clock time.

Hegel insists that philosophy is historical, not just in the obvious sense that philosophy like anything else has a history. He means philosophy is *essentially* historical. Different kinds of philosophy are dominant at various historical periods, for example, analytical or synthetic philosophy. Analytical philosophy involves analysing concepts and breaking them down into simpler constituents. Synthetic

philosophy involves the reverse: construing new concepts from old ones. For example: construing the concept of becoming from 'being' and 'nothingness'.

Philosophical movements are products of historical developments, For example: Liberalism as a product of the French Enlightenment and its emphasis on the individual. Kant's writing is a reaction to this, where it is claimed that it is not the individual, but the group that acts in history. When the individual acts, it is because he is representative of the group. Hegel feels this anti-individualistic view is a consequence of the Napoleonic wars. The preoccupation of political theorists with nationalism also dates from this period, Hegel claims. The awakening of Nationalism by the French Revolution, and its harnessing by Napoleon and its export over Europe, causes the introduction of the concept of the nation into political theory. Philosophy is not just affected by history. Philosophy is the an abstract expression of the historical events of an epoch. If one wished to argue against this position, one could point out that the people we select out as 'philosophers' are much more frequently those who *reject* the mentality of the age, those who think outside the existing paradigm.

Part Two: A Commentary on the Chapter 'Self-Consciousness' in Hegel's Phenomenology of Spirit (1807)

'Der Gewissheit seiner selbst' ('Self-Certainty') is Hegel's term for our pre-philosophical confidence in what we are. The chapter of *Phaenomenologie des Geistes* called 'Die Wahrheit der Gewissheit Seiner Selbst' calls this confidence into question (1). Here I offer a reading of the dozen paragraphs immediately preceding the celebrated chapter 'Selbstaendigkeit und Unselbstandikeit des Selbstbewusstseins: Herrschaft und Knechtschaft' (127-136) ('Independence and Dependence of Self-Consciousness: Lordship and Bondage'(111-119)) by which they have been rather overshadowed and which, nevertheless, they make possible.

Gewissheit

'Gewissheit' is used by Hegel to describe the manner of conviction generated within a form of consciousness, for example, in 'die sinnliche Gewissheit' (69 ff.) or 'Gewissheit und Wahrheit der Vernunft' (157 ff.). Although 'Gewissheit' is the ordinary German word meaning 'certainty', it may also be translated, *inter alia*, as

'proof', 'certitude' or 'assurance'. The Hegelian expression 'der Gewissheit seiner selbst' is more naturally translated as 'self-assurance' or 'self-confidence' than by Arnold Miller's 'self-certainty' which is awkward English. If we understand '[die] Gewissheit seiner selbst' as denoting certainty about oneself then Hegel may be read as a critic of Cartesian doctrines about the incorrigibility and transparency of self-knowledge. If we understand 'der Gewissheit seiner selbst' as denoting self-confidence or self assurance then we may read Hegel as replacing a naive pre-philosophical view with a richer, philosophical, view. The two interpretations are mutually consistent and on either or both of them the transition to the 'Herrschaft und Knechtschaft' chapter is facilitated by these passages.

Hegel thinks there is a tension, indeed some incompatibility, between what we think we are and what we are. This tension will reveal the inadequacies of the form of consciousness called '[die] Gewissheit seiner selbst'. Arguably, we tend to think that what we believe about ourselves is fundamentally correct (when we are not engaged in self-questioning) so the phrase 'self-certainty' is to that extent apposite. (2) We are self-certain in the sense that we take for granted the accuracy of our self-image.

The Paradox of Self-Consciousness

Hegel begins 'Die Wahrheit der Gewissheit Seiner Selbst' by distinguishing self-consciousness from consciousness in general; the moments (*Momenten*) or aspects (*Gestalten*) of which he has been charting so far in the book. He contrasts them by saying on the one hand;

In den bisherigen Weisen der Gewissheit ist dem Bewusstsein das Wahre etwas anderes als es selbst. (120)

'In the previous modes of certainty what is true for consciousness is something other than itself.' (M104)

and on the other hand;

Nunmehr aber ist dies entstanden, was in diesen fruheren Verhaeltnissen nicht zu Stande kam, naemlich eine Gewissheit, welche ihrer Gegenstand, und das Bewusstsein ist sich selbst das Wahre. (103)

'But now there has arisen what did not emerge in these previous relationships viz a certainty which is identical with its truth; for the certainty is to itself its own object, and consciousness is to itself the truth.' (M104)

In any consciousness that is not a kind of self-consciousness, consciousness is consciousness of something other than itself. In self-consciousness, consciousness is consciousness of itself. In consciousness that is not self-consciousness there is room

for error (perhaps empirical error, or error in counting). In self-consciousness, on one level, there is no room for error because it is inevitable consciousness which is presented to consciousness. Hegel uses *wahr* ('true') and *Wahrheit* ('truth') both in his own external appraisals of forms of consciousness, and for the internal forms of knowledge and conviction they yield. It is therefore essential to distinguish then between 'certain or true' in the elucidation of the perspective *of* the consciousness Hegel is talking about, and in Hegel's own perspective *on* that consciousness. Here Hegel is describing *from the inside* what appears to consciousness and what appears to self-consciousness. They are distinguished by their objects.

On the face of it, Hegel's next remark is paradoxical

Es ist darin zwar auch ein Andersein, das Bewusstsein unterscheidet naemlich, aber ein solches, das fur es zugleich ein nicht Unterschiedenes ist. (120)

'In this [self-consciousness] there is indeed an otherness: that is to say, consciousness makes a distinction, but one which at the same time is for consciousness not a distinction' (M104).

as is

[...] das Unterschieden des Nichtzuunterschiedenden, oder Einheit des Unterschieden. (122)

'[...] the distinguishing of what is not to be distinguished or the unity of what is distinguished' (M106).

There are several ways of explaining this paradox as not self-contradictory.

Firstly, Hegel thinks of consciousness (at this stage) as a *relation*. If I am conscious then I am conscious of something or other. Thus, on this analysis, there is something paradoxical about self-consciousness in particular. There is a distinction because *something is related to something* but, in a different sense, no distinction because *something is related to itself*. I am related to myself in self-consciousness as subject and object but I am not on that account two consciousnesses. No breach of the law of identity is involved because, to borrow Aristotle's phrase I am not related to myself in the same respect.

Secondly, a distinction can be drawn between ontology and epistemology. In an ontological sense the being who is self-conscious is self-identical. (This is misleading for Hegel but can be accepted at present). There is an epistemological distinction between that which is conscious and what consciousness is conscious of, but not an ontological distinction. The structural relation presupposes that there is no distinction at the ontological level.

Thirdly, using the vocabulary of 'Subjekt' and 'Objekt' we can say subject and object are numerically identical in self-consciousness. If I am self-conscious I am

both the subject and the object of my experience. Even physically, if I look at a part of my body I am subject *qua* that which looks but object *qua* what my looking is at.

Fourthly, an internal/external distinction can be used to show that 'consciousness makes a distinction', that is, the subject/object distinction is drawn by the particular form of consciousness under consideration, and does not have to be ascribed to this consciousness by Hegel or his readers in the position of appraising this form of consciousness.

Fifthly, in self-consciousness I am both active and passive; active in so far as I am conscious, passive in so far as I am what my consciousness is consciousness of. That I am self-identical, however, is a precondition of of this structure.

Once its coherence is explained in any or all of these ways this proposition can be seen as a precise description of the structure of self-consciousness:

[...] es nur sich selbst als sich selbst von sich unterscheidet. (121)

'[...] what it distinguishes from itself is only itself' (M105).

According to Hegel, one element has so far been omitted from the description: the *subjectivity* of self-consciousness; or, which amounts to the same thing for Hegel, the structure of the self, or I, in self-consciousness. Hegel says;

Ich ist der Inhalt der Beziehung, und das anderes dasselbe ist; es ist es selbst gegen ein anderes, und greift zugleich ueber dies andre ueber, das fur es ebenso nur es selbst ist. (120)

' the 'I' is the content of the connection and the connecting itself. Opposed to an other the 'I' is its own self, and at the same time it overarches this other, which for the 'I' is equally only the 'I' itself'. (M104).

Using the subject - object vocabulary, this passage can be understood as follows: The I, or self, is not just one of the structures of self-consciousness but *all of them*: the subject, the object, and the relation between the two. Using the term 'experiences' the point can be made this way: The I, or self, is *what experiences, what experience is experience of and the experience itself*. It is appropriate for Hegel to introduce the first person singular pronoun here because it is a necessary truth that that which could not in principle be described in first singular person terms could not be self-consciousness. Non-language users might be self-conscious, but only on condition that they are the sort of beings that could become language users. Despite this public or social constraint, self-consciousness is irreducibly subjective. Any being that is self-conscious is necessarily thereby related to itself. If any description excluded this utterly individualistic possibility then it could not be a description of self-conscious.

Hegel's distinction between consciousness in general and self-consciousness in particular is illustrative of Hegel's philosophical method. Consciousness is

aufgehoben (relieved, synthesised, abolished, retained) by self-consciousness. *Aufgehoben* is the term employed in dialectical reason to describe the reconciliation, and exhibiting as partial and particular truths, of two perspectives or concepts and their subsequent subsumption under a new and broader concept or perspective. Here all that was outlined by Hegel as true of consciousness is retained as a part description of self-consciousness. All the moments, or dialectical components, or stages, of consciousness are parts of self-consciousness.

Hegel describes the first appearance of self-consciousness:

Betrachten wir diese neue Gestalt des Wissens, das Wissen von sich selbst, im Verhaeltnisse zu dem Vorgehenden, dem Wissen von einem Anderen, so ist dies zwar verschwunden; aber seine Momente haben sich zugleich eben so aufbewahrt; und der Verlust besteht darin, dass sie hier vorhanden sind, wie sie an sich sind. (121)

‘If we consider this new shape of knowing, the knowing of itself, in relation to that which preceded, viz. the knowing of another, then we see that though this other has indeed vanished, its moments have at the same time no less been preserved, and the loss consists in this that here they are present as they are in themselves.’ (M105).

Self-consciousness is necessarily not the knowing of an other but necessarily a self-knowledge. Everything described under 'Die sinnliche Gwissheit; oder das Diese und das Meinen' ('Sense-Certainty: or the "This" and "Meaning"') and the subsequent chapters has here been retained.

For example; sense experience can, of course, be part of self-consciousness. Similarly, the attitudes and ways of knowing adopted in being conscious of an object which is not oneself can be employed to be conscious of the one one is. Hegel says the only difference is that, considered as aspects of self-conscious, these moments or aspects of consciousness can be considered *as they are*, not just *as they appear*. This is because, as rational contemplators of the structures of consciousness, we are not actually thereby engaged in being conscious of objects in the sense under examination. If we were, our view would be ‘internal’ or that of a participant. We would, in fact, understand them as experienced and not as thought. This is perhaps a tenuous argument as merely thinking about something is a mode of being conscious of it and thinking about something is no more likely to be free of interpretative preconception than experiencing it directly. If the trend to neutrality or objectivity is made less strict, say one of degree, then it might be plausible to maintain that rational contemplation of something is *less likely* to be distorted than experience of it. That too would need argument.

Hegel next presents an argument to show that self-consciousness is consciousness of itself. This is not just the intuitively plausible and perhaps analytic claim that if someone is self-conscious then they are conscious of themselves. It is the further claim that if someone is self-conscious then they are *conscious of their self-*

consciousness. Prima facie this further claim would not seem to follow logically from the first. It would seem that from something is conscious of itself it cannot be concluded that it is conscious that it is conscious of itself, or conscious of the consciousness of itself, or any such meta-claim. There also seems a danger of an infinite regress. If self-consciousness entails consciousness of the self-consciousness it is, does that not entail that the self-consciousness it is conscious of is conscious of the self-consciousness it is, and so on?

Hegel says the structure of self-consciousness is a combination of unity and distinction;

Es ist als Selbstbewusstsein Bewegung; aber indem es nur sich selbst als sich selbst von sich unterscheidet, so ist ihm der Unterschied, unmittelbar als ein Anderssein aufgehoben. (121)

'As self-consciousness it is movement; but since what it distinguishes from itself is only itself as itself, the difference as an otherness is immediately superseded by it.' (M105)

Self-consciousness supersedes this otherness in the sense that the 'distinction' is *aufgehoben*, or maintained, as part of what self-consciousness consists in. Otherness is a 'distinct moment' for self-consciousness (M105). Hegel says:

[...] aber es ist für es auch die Einheit seiner selbst mit diesem Unterschiede, als zweites unterschiedenes Moment. (121)

'[...] but there is also for consciousness the unity of itself with this difference as a second distinct moment' (M105)

In other words, it is the unity of both (a) the unity of subject and object and (b) the structural difference between them that characterises self-consciousness more fully. The relation between subject and object in self-consciousness is identity-in-difference.

Another interpretation of Hegel's view here is perhaps complementary to that put forward so far. The distinction that Hegel mentions is a self/ not-self distinction. It is certainly a necessary condition for any being being self-conscious that that being draw a distinction between what is they are, on the one hand, and everything they are not, on the other. Hegel can then be taken as defining the structure of self-consciousness in contrast to that of ordinary empirical sense experience of what is not oneself. This interpretation is born out by this passage:

Das Bewusstsein hat als Selbstbewusstsein nunmehr einen gedoppelten Gegenstand, den einen, den unmittelbaren, den Gegenstand der sinnlichen Gewissheit und des Wahrnehmens, der aber fuer es mit dem Charakter des

Negativen bezeichnet ist, und den zweiten, naemlich sich selbst, welcher das wahre Wesen, und zunaechst nur erst im Gegensatzee des ersten vorhanden ist. (121-2)

'Consciousness, as self-consciousness, henceforth has a double object; one is the immediate object, that of sense certainty and perception which however for self-consciousness has the character of a negative; and the second viz itself which is the true essence, and is present in the first instance only as opposed to the first object.' (M105)

The 'negative' role of sense certainty suggests perhaps a merely contrastive role for sense experience which self-consciousness is not engaged in but requires, in order to be what it is. The obvious drawback with this interpretation is that the self-not self distinction does not coincide with the thought-sense experience distinction. I can be conscious of myself through the senses and I can merely think about what is not myself.

So 'sense experience' cannot do the work of bestowing meaning on 'self-consciousness' contrastively.

A more plausible interpretation is this. To be fully self-conscious it must be possible, at least in principle, to know that one is self-conscious. This is because, within what we could call 'bare' self-consciousness, the subject-object/unity of self-consciousness is not revealed. When I am 'barely' self-conscious I am only conscious of *that which in fact I am* (whether my consciousness is through sense experience or thought). What is required for 'bare' self-consciousness to become 'fully fledged' self-consciousness? Unless the further requirement is fulfilled that I know that *I am the being I am conscious of*, ie, that it is myself that I am conscious of (and not just, *a hand or this body here*, where this hand is *in fact* the hand of the self-same person who is conscious of it) I am not fully self-conscious. I lack an essential component of what we usually mean by 'self-consciousness'. But what is the meaning of this 'I am the being I am conscious of'. To understand this we would have to know what 'I am' means and what 'I am this' says that 'this is this' does not. We are far from understanding these meanings.

It is not just necessary to draw a self/not-self distinction to facilitate self-consciousness. This is necessary but not sufficient. The being who is self-conscious must that they are both the being who is conscious and the being who his consciousness is of. This is the force of Hegel's conclusion that the 'second' object of self-consciousness is 'itself - which is the true essence' (M105).

Hegel says:

Dieser Gegensatz seiner Erscheinung und seiner Wahrheit, hat aber nur die Wahrheit, naemlich die Einheit des Selbstbewusstseins mit sich selbst, zu seinem Wesen. (121)

'This antithesis of its appearance and its truth has however for its essence only the truth viz, the unity of self consciousness with itself' (M105)

The antithesis is between the partial perspective of the person who is conscious of themselves, but does not know it, and the fuller, truer, perspective of the person who is conscious that they are self conscious, and know that they are what they are conscious of. This latter, including and retaining the structure of the former, is the 'true essence' of self-consciousness, that is, what self-consciousness really is. Rational reflection, here as elsewhere for Hegel is a more reliable guide to the truth than experience.

Hegel describes self-consciousness as a 'being that is reflected into itself'. This image, perhaps of mirrors facing each other, is designed to capture an aspect of self-consciousness which is not present in an analysis of self-consciousness as mere consciousness of oneself as yet another object for sense experience. It is a poetic way of making the point that that which is self conscious is what self consciousness is consciousness of.

Desire

A controversial and obscure element in Hegel's account has not been discussed so far: the role of desire (*Begierde*). In some sense, a being must *feel desire* in order to be self-conscious. This must be a causal, or metaphysical, claim and not a logical one as it is not self-contradictory to suppose that a being has no desires whatsoever but is nevertheless self-conscious. The concept of desire employed by Hegel can be understood on two main levels; the ordinary human level and the universal or metaphysical level. At each level, it is used in a different sense.

At the human level, Hegel holds it is a causally necessary condition (or historically necessary condition) of our drawing the self/not-self distinction, in the way in which we do in fact draw it, that *as persons* we desires. Care is needed about the deployment of the concept of a person here. If being a person entails being self-conscious, we are not entitled to talk of persons at this stage of the dialectic. Desire is the emotion which establishes a *discrepancy* between our own needs, or perceived needs, and the way the world of our experience is. The suggestion is that if there was never any disharmony between the way we are and the way the world is then we would have no reason, no occasion to draw a self/rest of world distinction. This is a historical claim. It is a story about how self-consciousness *came about*.

To be self-conscious we must draw a self/other distinction. It looks like a contingent fact that it is made through *desire*. Desire and the self - certainty obtained in its gratification are conditioned by the object.

On the other, metaphysical, reading the subject of desire is not this or that individual conscious being but the universal mind, *Geist*, who finds fulfilment through the individual's aspiration, action and expression. This reading is consistent

with the first. In the context of Hegel's system, it includes or entails that account. The cosmic process of *Geist's* coming to self-realisation may then be correctly called 'desire' or viewed as essentially entailing desire. This gives the concept a teleological sense. As *Geist* is infinite, and conscious beings finite, and the second is part of the first, so *Geist's* desire is infinite and the individual's finite and the second is part of the first. On this metaphysical view, the individual's coming to self-consciousness is a passive development, the expression of *Geist's* cosmic purpose.

It is part of Hegel's philosophical style to make the preconditions for something (metaphysical, causal, logical) part of the fully explicated concept of that thing. Thus he can say about self-consciousness

[...] es ist Begierde überhaupt (121)

'[self-consciousness] is desire in general'. (M 105)

The Possibility of the Herrschaft/Knechtschaft relation.

The precondition for being self-conscious that Hegel devotes most attention to is this: A being can only be fully self-conscious through a struggle for recognition by another consciousness. Self-consciousness is necessarily social and a private or solipsistic self-consciousness is impossible. These for Hegel are dialectical claims. The argument for them is mainly in the famous Herrschaft und Knechtschaft section but they are anticipated in the demonstration of the inadequacies of self-certainty.

Before examining this social precondition for self-consciousness one qualification must be made. I have been using language like 'someone is conscious' which implies that there is a subject of consciousness; a person or being of some sort which is conscious. Now Hegel sometimes speaks like this but, much more often, he makes *consciousness itself* into a subject and agent. This is peculiar at first sight. It would seem to be a simple category mistake to hold that what is conscious is consciousness, what is self-conscious is self-consciousness etc. To account for this usage we have to refer again to Hegel's system. There *Geist* is not a subject over and above consciousness. *Geist* is embodied in nature but there is no Cartesian soul, or transcendental ego, which could do the work of being *that which is* conscious, at least, not any particular soul or ego.

If there is a subject of consciousness it is ultimately *Geist*, the world- spirit itself. In this very fundamental sense, the subject of every individual consciousness 'is' this universal spirit. The explication of this 'is' is the progress through states of consciousness described in the *Phenomenology of Spirit*.

The transition from Self-Certainty to the more fully developed social self-consciousness of the Master and Slave chapter takes place, as usually in the *Phenomenology*, by the earlier more rudimentary shape of consciousness being exhibited as in some way deficient, or inadequate, and its being 'superseded' and

shown to be part of the structure of a more complete form of consciousness. Self-Certainty's inadequacy is as follows;

Diese Selbstaendigkeit der Gestalt erscheint als ein Bestimmtes, fuer anderes, denn sie ist ein Entzweites; und das Aufheben der Entzweigung geschieht insofern durch ein anderes. (123)

'This independence of the shape appears as something determinate, for an other, for the shape is divided within itself; and the supersession of this dividedness accordingly takes place through an other.' (M107)

What does this 'dividedness within itself' (107) consist in? It consists at least in being the subject-object structure that, with an underlying identity or unity, is the structure of self-consciousness. But Hegel is saying more than this. The putatively private self-consciousness is divided from itself. It does not know it is itself. It can not identify what is conscious with what self-consciousness is consciousness of. In other words, to use a Hegelian turn of phrase, self-consciousness does not know itself *as such*. In individual terms, a person may perceive or think of an aspect of the person who they in fact are but not know that they are in fact identical with that person. This is precisely the condition of the *deficient* self-consciousness. The condition that needs to be fulfilled is: for x to be self-conscious x must know that x is related to x by the relation 'is numerically identical with'. But, it is not enough for x to just know 'x is x'.

x must know the further fact, 'I am x '. It is in this context that this passage is to be understood:

[...] das Selbstbewusstsein, ist sich zunaechst nur als dieses einfache Wesen, und hat sich als reines Ich zum Gegenstande. (125)

'[...] self-consciousness, exists in the first instance for self-consciousness only as this simple essence, and has itself as pure I as object'. (M109)

The stress must be put on 'pure I', not on 'object'. Self-certainty knows itself as the object of its consciousness but not as subject and not as identity of subject and object.

Hegel says that the 'I' must 'enrich' itself. The 'I' is really the 'simple universal for which the differences are not differences'. (M109)

Towards the end of the Self-Certainty section, Hegel is quite explicit that self-consciousness is necessarily social:

Das Selbstbewusstsein reicht seine Befriedigung nur in einem andern Selbstbewusstsein. (126)

'Self-consciousness achieves its satisfaction only in another self-consciousness.'

(M110)

'Achieves its satisfaction' means 'becomes more fully what it is'. Being what it is and knowing what it is are not independent processes for self-consciousness. Each is a necessary condition for the other and so both, in a dialectical sense, are 'identical'. Coming to know that I am self-conscious is part of what becoming fully self-conscious consists in.

'Otherness' in this passage does not refer to what self-consciousness is not, but to what self-consciousness is of:

Es ist ein Selbstbewusstsein fuer ein Selbstbewusstsein. Erst hiedurch ist es in der tat; denn erst hierin wird fuer es die Einheit seiner selbst in seinem Anderssein. (126)

'A self-consciousness exists for a self-consciousness. Only so is it in fact self-consciousness; for only in this way does the unity of itself in its otherness become explicit for it.' (M110)

Only when the subject of self-consciousness *recognises itself in* what it is conscious of, in self-consciousness, will it be fully self-conscious and thus 'undivided'.

The critique of self-certainty ends with an introduction to *Geist*. Here Hegel is mentioning 'what still lies ahead' (M110) in the explication of self-consciousness and so giving us a short preview of a later and more developed understanding of self-consciousness. *Geist* coming to self-realisation cannot be achieved before the further stages of consciousness are enacted. *Geist* is described at this point:

Was fuer das Bewusstsein weiter wird, ist die Erfahrung, was der Geist ist, diese absolute Substanz, welche in der vollkommenen Freiheit und Selbstaendigkeit ihres Gegensatzes, naemlich verschiedener fuer sich seiender Selbstbewusstsein, die Einheit derselben ist; Ich, das Wir, und Wir, das Ich ist. (127)

'What still lies ahead for consciousness is the experience of what Spirit is - this absolute substance which is the unity of the different independent self-consciousnesses which, in their opposition, enjoy perfect freedom and independence: "I" that is "we" and "we" that is "I".'. (M 110)

This statement of what *Geist* is is consistent with the interpretation of the subject of self-consciousness and with the metaphysical outline given so far. The concept of *Geist* is invoked at this point in the argument for a very appropriate reason. There is a strong analogy here between, on the one hand, the limitations in the structure of a particular self-consciousness and, on the other hand, looking ahead, in the limitations inherent in the inter-subjective structure of a plurality of individual self-consciousnesses. The analogy is this: What the solitary or solipsistic self-

consciousness cannot do is know itself as undivided; as a unity. This is precisely what a plurality of self-consciousnesses cannot do either. They conceive of themselves as unique individual units even though they are essentially interrelated in complex ways. Just as one self-consciousness finds its unity and true structure by encounter with another so, in a parallel way, disparate self-consciousnesses come to realise their underlying unity as so many expressions or manifestations of One universal consciousness which is *Geist*, the universal spirit.

The role of *Geist* in displaying the identity-in-difference of a plurality of self-consciousnesses is structurally identical to the role of one self-consciousness in generating the recognition of unity. Common to both processes is the notion of independence and dependence (*Selbstandigkeit und Unselbstandigkeit*). The very concept of dependence implies a degree of independence: *a* can only be independent of *b* if *a* is not *b*. Independence is at least a two term relation so there must be two or more, to some degree independent, elements which can be dependent on one another for dependence or independence to be possible. The title Hegel gives to the master and servant dialectic is: 'Independence and Dependence of self-consciousness Lordship and Bondage'. (M111) The logical relationship exhibited here is: if *x* is dependent on *y* then *x* is in some sense independent of *y*. Note that the relation is not a reciprocal or symmetrical one. If *x* is independent of *y* it does not follow that *x* is in any sense dependent on *y*.

The conclusion of the Master and Slave dialectic is its first sentence;

Das Selbstbewusstsein ist an und fuer sich, indem, und dadurch, dass es fuer ein anderes an und fuer sich ist; d.h. es ist nur als ein Anerkanntes. (127)

'Self-consciousness exists in and for itself when and by the fact that, it so exists for another; that is it exists only in being acknowledged.' (M111)

The biggest single question of this chapter is: Why? Why is the encounter with an other consciousness a necessary condition for each self-consciousness to be fully possible?

The answer is this: One consciousness *models itself* on another. There are many different and complex reasons why self-consciousness is necessarily social but this is the most fundamental. It is not possible for a solitary self-consciousness to have a complete picture of itself as a self-consciousness *only from its own case*. The very structure of self consciousness in a sense belies this. In order to understand this dependence more fully it helps to examine the process of the acquisition of the concept of a self on a more commonsensical level. Each of us does in fact conceive of himself or herself as an individual person in the world. From a Hegelian perspective, this self-image is a composite one. It is partly made up of *reflection on oneself* but also partly based on *observation of other persons*. Each of us must be able to think: "I am one of those". If there were no exposure to other conscious beings in this way then each conscious being would be restricted to the partial and peculiar

phenomenological perspective they have on themselves. Unless I conceive of myself as an individual, as a particular which can in fact be discriminated from the rest of the world, from its environment, then the concept of myself as a person would not be possible. If 'I' and 'myself' could find the slightest grip here, I would perhaps think of myself as coextensive with my sensory field.

On this view, the concept of the self, at a profound level, is a construct. We are not speaking at the level of personality but at a level which is a precondition for personality. The way each of us thinks of him or herself is modelled partly on our observation of others, partly on our experience of ourselves.

Hegel in this chapter is less concerned with how we come to conceive of ourselves as persons than how self-consciousness is possible. Although Hegel does think that being self-conscious is in a sense criterial of personhood. The level of the discussion is not just one of self-image but of self-consciousness. It is precisely the encounter with another self-consciousness which overcomes self-consciousness' otherwise divided nature; which forces the recognition of its own unity. We have already seen that this cannot be done solipsistically. It is perceiving another, and recognising itself in another, that supplies the missing component in the otherwise solipsistic self-consciousnesses' self consciousness. What it learns from the other, it realises applies to itself. It perceives the other as a unity, as undivided, realises it is essentially like the other, and no longer perceives itself as divided but as a unity.

In order to see *how* self-consciousness finds its unity through the encounter with an other we have to follow Hegel's account in detail. Hegel says:

Es ist fuer das Selbstbewusstsein ein anderes Selbstbewusstsein; es ist ausser sich gekommen. Dies hat die gedoppelte Bedeutung; erstlich, es hat sich selbst verloren, denn es findet sich als ein anderes Wesen; zweitens, es hat damit das Andere aufgehoben, denn es sieht auch nicht das andere als Wesen, sondern sich selbst im andern. (128)

'Self-consciousness is faced by another self consciousness; it has come out of itself. This has a twofold significance; first, it has lost itself, for it finds itself as an other being; secondly, in doing so it has superseded the other, for it does not see the other as an essential being, but in the other sees its own self.' (M111)

In order for one self-consciousness to encounter a second it must 'come out of itself'. This means it must logically cease to be a self-consciousness *qua* consciousness of something other than itself. In this sense it has 'lost itself'. It finds itself as an other being. Its attention is absorbed by that other so in a sense it 'is' it. This is a familiar and recurrent theme in idealist and in radical empiricist philosophy where the self is identified with its experiences. Berkeley holds that what experience is of is states of oneself and Hume holds that the self is its experiences. One is reminded of Heidegger's 'I am what I am concerned with'. Similarly, Hegel is saying that consciousness of the other self-consciousness partly constitutes what that

consciousness is.

The second implication Hegel draws is that, in this consciousness of the other self-consciousness, the consciousness of the other recognises itself as the other. This in turn can be understood in at least two ways. Hegel is saying that the other is recognised as a fellow self-consciousness, as a being just like the one perceiving it, by the one perceiving it. It is also being maintained that the first consciousness finds itself in the second in the idealist and empiricist sense mentioned above. The perceiving self-consciousness recognises that it requires the other to be what it is; not just to return to being a self-consciousness, and fully become self-conscious, but also just to be the present structure of consciousness directed at the other self-consciousness.

In these passages there is an ambiguity in Hegel's use of the term 'other'. Sometimes it is used to mean 'the other self-consciousness', that is, the one encountered by the first. Sometimes it means 'the other of the first self-consciousness', that is, what that self-consciousness is consciousness of. It means the object pole of the subject - object relation which structures self-consciousness. In this second sense, the use is parallel to *Geist's* 'other', which is nature. The other of *Geist* is its embodiment the physical world.

Prima facie, and paradigmatically, 'the other of' the individual's self-consciousness is the human body. This is the embodiment of the person (understood on Aristotelian not on Cartesian lines). Interpreting 'other' in the second sense, the following sentence shows self-consciousness coming to be possible:

[...] es sieht auch nicht das andere als Wesen, sondern sich selbst im andern.
(128)

'[...] it does not see the other as an essential being, but in the other sees its own self.'
(M 111)

Here the subject of self-consciousness realises it is its own object. Before it was divided, it did not know that what its consciousness is of is itself. Modelling itself on the other, unified as self-consciousness, it 'supersedes this otherness of itself'.

There is a further line of thought which supports this interpretation. How can *x*, the subject of consciousness, be conscious of *y* a distinct subject of consciousness? How is the encounter possible? Is this only possible through sense experience, perceiving the other self-consciousness as embodied? Of course *x* can think of *y* but to perceive *y* entails sensing *y*. It is then as an embodied consciousness that *x* perceives *y*. Not only is *x's* concept of *x* partly modelled on *x* and partly on *y* but also, *x's* concept of *y* must be partly modelled on *x* and partly on *y*. By perceiving *y* as embodied *x* realises that it is that itself embodied. *x* realises that its other in self-consciousness is its own embodiment - the embodiment of itself as subject of consciousness and not just an other - as object of consciousness.

The ambiguity in the concept of 'other' is deliberate. Hegel speaks of 'This

ambiguous suppression of its ambiguous otherness' and 'an equally ambiguous return to itself'. The two processes, or the double process, is this

[...] erstlich erhelt es durch das Aufheben sich selbst zurueck; denn es wird sich wieder gleich durch das Aufheben seines Anderseins. (128)

'[...] first, through the supersession, it receives back its own self because by superseding its otherness it again becomes equal to itself' (M 111)

and

zweitens aber gibt es das andere Selbstbewusstsein ihm wieder ebenso zurueck, denn es war sich im andern, es hebt dies sein Sein im andern auf, entlaesst also das andere wieder frei. (128)

'but secondly the other self-consciousness equally gives it back again to itself for it saw itself in the other, but supersedes this being of itself in the other and thus lets the other again go free.' (M 111)

'Other' in (1) is the embodiment of the first self consciousness. 'Other' in (2) is the second self-consciousness.

There is a point that needs to be emphasised next which is often missed but crucially affects the way in which the Master and Slave dialectic is understood. Hegel is describing the relationship between two self-consciousness from the point of view of one of the only. It should not at all be concluded from this that what he says about one self-consciousness does not apply, equally, simultaneously, and in every respect, to the 'other' self-consciousness. Master is slave and slave is master. The relation between the two self-consciousnesses is a symmetrical and reciprocal. Hegel puts it this way:

Diese Bewegung des Selbstbewusstseins in der Beziehung auf ein anderes Selbstbewusstsein ist aber auf diese Weise vorgestellt worden, als das Tun des Einen; aber auf dieses Tun des Einen hat selbst die gedoppelte Bedeutung, ebensowohl sein Tun als das Tun des Andern zu sein; denn das andere ist ebenso selbstaendig, in sich beschlossen, und es ist nichts in ihm, was nicht durch es selbst ist. (128)

' [...] this movement of self-consciousness in relation to another self consciousness has in this way been represented as the action of one self-consciousness, but this action of the one has itself the double significance of being both its own action and the action of the other as well. For the other is equally independent and self-contained, and there is nothing in it of which it is not itself the origin.' (M112)

By an action being both one self-consciousness' own action and that of the other, Hegel does not mean numerically the same action. He means there is *a mirror action*, a numerically distinct but qualitatively identical action performed by the other self-consciousness.

This observation has an important consequence for understanding the Master and Slave dialectic. It means that it is extremely misleading to say that there are two qualitatively distinct forms of self-consciousness; one master and one slave. This is only true from an internal perspective that is, from the point of view of one of the two self-consciousnesses. Objectively speaking, or from the view from nowhere, the structure of individual self-consciousnesses is identical. Hegel at this point describes the relationship symmetrically:

Die Bewegung ist also schlechthin die gedoppelte beider Selbstbewusstsein. Jedes sieht das andre dasselbe tun, was es tut; jedes tut selbst, was es an das andere fodert; und tut darum, was es tut, auch nur insofern, als das andre dasselbe tut; das einseitige Tun waere unnuetz, weil, was geschehen soll, nur durch beide zu Stande kommen kann. (129)

'Thus the movement is simply the double movement of the two self-consciousnesses. Each sees the other do the same as it does; each does itself what it demands of the other, and therefore also does what it does only in so far as the other does the same. Action by one side only would be useless because what is to happen can only be brought about by both.' (M112)

Hegel says:

Jedes ist dem anderen die Mitte, durch welche jedes sich mit sich selbst vermittelt und zusammenschliesst, und jedes sich und dem andern unmittelbares fuer sich seiendes Wesen...

'Each is for the other the middle term, through which each mediates itself with itself and unites with itself' (M112)

'They recognise themselves as mutually recognising each other' (M112)

In understanding the Master and Slave dialectic, we must bear constantly in mind that Hegel is describing a symmetrical or reciprocal relation from one side only.

Hegel says we have to now see how 'the duplicating of self-consciousness in its oneness appears to self-consciousness' (112). We are to adopt the standpoints of the self-consciousnesses themselves. This is yet another example of Hegel's heuristic tendency to oscillate between an 'internal' and an 'external' perspective. It is clear that here he is abdicating rational reflection for hermeneutical reconstruction: 'hermeneutic' in the sense of *understanding understanding* through a kind of empathy,

or identification with the subject of study. Hence, in describing how the inter-action 'appears to self-consciousness' Hegel speaks of:

'At first [...] one only recognised, the other only recognising'.

This is a heuristic device on Hegel's part to show the interaction from one point of view only. In this way the argument continues.

A self not self distinction is necessary for the form of self-consciousness called 'self-certainty' to be possible but in addition to this, for full self-consciousness to be possible a 'self-other' distinction is necessary.

The distinction between the self/not-self distinction and the self/other distinction is between, on the one hand, oneself and everything that is not oneself and, on the other hand, oneself and another individual self consciousness. It is the first distinction that Hegel is employing when he says that

'self-consciousness is to begin with, simple being for self, self-equal though the exclusion from itself of everything else'. (M113)

but the second when he says that:

'the other is also a self-consciousness' (M113)

Although self consciousnesses are in all essential respects similar when considered objectively, it should be emphasised that there are two numerically distinct self consciousnesses which interact in the Master-Slave dialectic. Although Hegel is trying to show what self-consciousness in general consists in, he has to thereby consider two self consciousness and not just the structure of one. Hegel says:

'one individual is confronted by another individual'. (M113)

The precise role of the other self-consciousness in rendering the first self-consciousness fully what it is, is to supply it with an objective understanding of itself. The first self-consciousness cannot experience itself as an individual object which can be discriminated from the rest of the items in its experience. In this sense it is solipsistic. The other provides the model on which it understands itself as itself, discriminates itself from what it is not. The other stands proxy for the alleged impossibility of viewing oneself as another views one. Hegel says:

'Each is indeed certain of its own self, but not of the other, and therefore its own self-certainty still has no truth. For it would have truth only if its own being-for self had confronted it as an independent object'. (M113)

Hegel says this essential condition for self-consciousness is fulfilled 'only when each

is for the other what the other is for it' (M113).

Hegel calls the 'presentation of itself' at this stage a 'pure abstraction'. This means that the concept of self is partly modelled on the other and partly on the self. It is a synthesis, a construct.

Hegel emphasises that the concept of self-consciousness is not empirical or, at least, cannot be entirely acquired empirically. There is no one particular object one could have experience of and thus acquire the concept of self-consciousness, even though certain experiences are necessary for the synthesis of the concept and its self-application. Indeed, the concept requires something like the denial of its restricted employment in sense experience. It must be seen to be applicable to make oneself intelligible and this involves negating its role as just part of the other:

'The presentation of itself [...] as the pure abstraction of self-consciousness consists in showing itself as the pure negation its objective mode, or in showing that it is not attached to any specific existence as such, that it is not attached to life. This presentation is a twofold action; action on the part of the other and action on its own part'. (M113)

The thesis that concepts of the self (here specifically self-consciousness) in some way do not have application within experience is a recurrent one in modern philosophy. It is present in Hume as the inability to find any one impression to give rise to the idea of self, in Kant as the transcendental unity of apperception; a logical precondition of experience and not any part of experience, in the early Wittgenstein as a 'limit' of the world and not as part of it, in Heidegger as a reluctance to employ terms like 'self' and 'subject' as 'completely miss(ing) the phenomenal content of Dasein' (BT 150), in Ryle as the 'systematic elusiveness of I'. Hegel is concerned with self-consciousness in particular, but making a parallel point by saying a concept has no application to any 'particular existence' and is 'not attached to life'.

The question arises here of what it is about the concept of self-consciousness which prevents it from finding application within experience. That the concept of self is part of the concept of self-consciousness provides good Humean grounds to be sceptical about any straightforward empirical employment of it. Also, the self-consciousness of the other is hidden and that rules out any possibility of acquiring the concept through some process of abstraction from perceptual exposure to multiple cases. Perhaps the 'conscious' component in the meaning of 'self-conscious' also prevents the concept from finding straightforward empirical use. For example, if I am conscious it is true of me that I can know that I am conscious. What is not true is that I can perceive, or notice, my consciousness in the ways I can perceive or notice the objects I am conscious of. (Kant was in this respect wrong to talk about 'inner sense'.) My consciousness of an object is not another object I can be conscious of. Rather it is a precondition of my being confronted with any objects at all. The concept 'conscious' seems to give rise to similar sorts of empirical scepticism as the concept of self, and seems equally to be a precondition of experience. Perhaps 'I was

having experiences but not conscious' is not self contradictory. Perhaps I was dreaming, and so not clearly conscious or unconscious, or can give sense up to 'unconscious experience' but that we are conscious is a precondition of having the experience we do have including unconscious experiences.

Hegel does not think this consideration about consciousness is a reason for denying that the concept of self-consciousness is partly acquired through sense experience.

We have now covered the parts of the Phenomenology which show how self-consciousness is constructed out of experience of self and other. The Master and Slave dialectic continues with the emphasis switching from metaphysics to social theory, and a detailed investigation of how self-consciousness is changed by interaction with other self-consciousnesses

Notes

(1) G.W.F. Hegel *Phaenomenologie des Geistes* Neu Herausgegeben von Hans-Friedrich Wessels und Heinrich Clairmont, Mit einer Einleitung von Wolfgang Bonsiepen, Felix Meiner Verlag, Hamburg, 1988, p. 120, Hegel's *Phenomenology of Spirit* translated by A.V. Miller with analysis of the text and forward by J.N. Findlay (Oxford University Press, Oxford, 1979) p. 104. Page references are to this text and this translation in this paper.

(2) The contemporary North American philosopher Donald Davidson argues that we employ a 'principle of charity' in interpreting the utterances (inscriptions etc.) of others. We assume *prima facie* they are largely mutually consistent and true. Davidson fails to extend the principle of charity to the first person singular case. Arguably each person assumes their own beliefs to be largely mutually consistent and true because this assumption is constitutive of their being beliefs. Each person's own beliefs about themselves provide no exception. See Donald Davidson

BIBLIOGRAPHY

Early Writings (Berne, Frankfurt)

Hegels theologische Jugendschriften 1793-1800 (1907)

Hegel's Early Theological Writings

Jena Period (1801-1807)

Differenz des Fichteschen und Schellingschen Systems (1801)
The Difference Between the Philosophical Systems of Fichte and Schelling

Critical Journal of Philosophy (ed.) 1802-3

Phänomenologie des Geistes (1807)
The Phenomenology of Spirit

Nuremberg Period

Wissenschaft der Logik (1812-16)
The Science of Logic

BEYOND HEGEL ?

The philosophy of Hegel is introduced and some existentialist and post-structuralist responses to it are examined.

Lecture list and reading:

(1) HEGEL: AN INTRODUCTION

Paul Edwards (ed.) *The Encyclopaedia of Philosophy*, entry under 'Hegel'.
Stephen Priest *Theories of the Mind* Chpt.3. pp 80-97.
Frederick Copleston *A History of Philosophy* Vol. 7, Part 1, Chpts. 9, 10, 11 'Hegel'.

(2) HEGEL'S PHENOMENOLOGY

G.W.F. Hegel *The Phenomenology of Mind* trans. T.N. Baillie
or
Hegel's Phenomenology of Spirit trans. A.V. Miller
esp. A I 'Sense-Certainty', B IV 'The Truth of Self-Certainty' and DD 'Absolute Knowing'.
Richard Norman *Hegel's Phenomenology: A Philosophical Introduction*.

(3) HEGEL'S DIALECTIC

G.W.F. Hegel *The Science of Logic*.
Hegel's Logic trans. William Wallace.
J. McTaggart *A Commentary on Hegel's Logic*.

(4) DERRIDA AND DECONSTRUCTION

Jacques Derrida Positions

Jacques Derrida Margins of Philosophy

Jonathan Culler 'Derrida' in John Sturrock (ed.) Structuralism and Since

(5) DERRIDA: FROM RESTRICTED TO GENERAL ECONOMY

Jacques Derrida 'From Restricted to General Economy: A Hegelianism Without Reserve' in *Writing and Difference*.

(6) DERRIDA: THE PIT AND THE PYRAMID

Jacques Derrida The Pit and the Pyramid: Introduction to Hegel's Semiology' in *Margins of Philosophy*.

(7) MERLEAU-PONTY'S EXISTENTIAL PHENOMENOLOGY

Merleau-Ponty *The Phenomenology of Perception*.

Mary Warnock *Existentialism* Chpt. 4.

(8) MERLEAU-PONTY ON HEGEL'S 'EXISTENTIALISM'

Merleau-Ponty 'Hegel's Existentialism' in *Sense and Nonsense*.

(9) KIERKEGAARD: AN INTRODUCTION

H.J. Blackham *Six Existentialist Thinkers* Chpt. 1.

Patrick Gardiner *Kierkegaard*

Frederick Copleston *A History of Philosophy* Vol.7. Part II, Chpt. 17.

Kierkegaard *Fear and Trembling*

(10) KIERKEGAARD: EITHER/OR

Kierkegaard *Concluding Unscientific Postscript*

Kierkegaard *Either/Or*

1. HEGEL'S SYSTEM

Stephen Priest *Theories of the Mind* (Penguin) Chpt. 3 pp. 80-97

Frederick Copleston *A History of Philosophy* (Doubleday) Vol. VII, Part 1 'Fichte to Hegel' pp. 194-294

Michael Inwood *Hegel* (Routledge) *passim*.

Michael Inwood (ed.) *Hegel* (OUP)

Michael Inwood *A Hegel Dictionary* (Blackwell)

Alexander Kojeve *Introduction to the Reading of Hegel*

Richard Norman *Hegel's Phenomenology* (Sussex)

J. McT. E. McTaggart *Studies in Hegelian Cosmology* (CUP)

Jean Hyppolite *Genesis and Structure of Hegel's Phenomenology of Spirit* (Northwestern)

2. HEGEL'S PHENOMENOLOGY

Hegel's Phenomenology of Spirit trans. Arnold Miller (OUP):

I Sense Certainty

II Perception

III Power and Understanding

IV The Truth of Self Certainty

A The Dependence and Independence of Self-Consciousness:
Master and Slave

B The Freedom of Self-Consciousness; Stoicism, Scepticism
and the Unhappy Consciousness

VIII Absolute Knowing

3. HEGEL'S DIALECTIC

Hegel's Logic trans. W. Wallace (OUP) 'The Doctrine of Being'

Hegel's The Science of Logic trans. W.H. Johnston & L.G. Struthers
'The Doctrine of Being' (George Allen & Unwin)

J. McT. E. McTaggart *Studies in the Hegelian Dialectic* (CUP)

J. McT. E. McTaggart *A Commentary on Hegel's Logic* (CUP)

Further Reading on Hegel

Lucio Colletti *Marxism and Hegel* (Verso)

Jacques Derrida 'The Pit and the Pyramid' in *Margins of Philosophy* (Sussex)

Jacques Derrida 'From Restricted to General Economy: A Hegelianism Without Reserve' in *Writing and Difference* (Routledge)

Jacques Derrida *Glas: What remains of absolute knowledge?*

Martin Heidegger *Hegel's Concept of Experience* (Harper and Row)

Georg Lukacs *The Young Hegel* (London)

Herbert Marcuse *Reason and Revolution: Hegel and the Rise of Social Theory* (Routledge)

Karl Marx *Early Writings* ed. Lucio Colletti (Penguin)

Maurice Merleau-Ponty 'Hegel's existentialism' in *Sense and Non-Sense* (Northwestern)
Stephen Priest (ed.) *Hegel's Critique of Kant* (OUP/Gower)
Stephen Priest *Merleau-Ponty* (Routledge) 'Hegel's Existentialism'

Questions

- (1) Has Hegel exposed the limitations of sense-certainty?
- (2) Has Hegel shown that self-consciousness is essentially social?
- (3) What does Hegel mean by 'Geist'? Does Geist exist?
- (4) Is Hegel's dialectic a genuine method of problem solving in philosophy?
- (5) Is Hegel (a) an existentialist, (b) a proto Marxist, (c) a kind of theologian ?
- (6) Is Hegel right to hold that there is absolute knowing?
- (7) What is the relation between being and nothingness according to Hegel? Is he right?
- (8) Has Hegel refuted any of the claims Kant made in *The Critique of Pure Reason* ?

UNUSED

HEGEL'S PHENOMENOLOGY OF SPIRIT

A. Consciousness

I. Sense Certainty: or the 'This' and 'Meaning'

(PGT 58)

What kind of knowledge is sense-certainty?

'knowing which is at the start or is immediately our object'

'immediate knowledge'

'a knowledge of the immediate or of what simply is'

'Our approach to the object must also be immediate or receptive; we must alter nothing in the object as it presents itself. In apprehending it, we must refrain from trying to comprehend it.'

Prima facie appearance of sense-certainty:

'Because of its concrete content, sense certainty immediately appears as the richest kind of knowledge'

'sense-certainty appears to be the truest knowledge'

because

'[it] has the object before it in its perfect entirety'

however

'this very certainty proves itself to be the most abstract and poorest truth'

Sense-certainty and being

'All that it says about what it knows is just that it is; and its truth contains nothing but the sheer being of the thing'

'this pure being, or this simple immediacy, constitutes its truth' (PGT 58-9)

'pure being (which) constitutes the essence of this certainty' (PGT 59)

(PGT 58 cont.)

I and This

'Consciousness, for its part, is in this certainty only as a pure 'I'

'I am in it only as a pure "This" and the object similarly only as a pure "This".

'I, this particular I, am certain of this particular thing'

'the thing is, and it is, merely because it is'

(PGT 59)

Sense-certainty = a relation between 2 relata:

'an immediate pure connection'

between

'a pure "I" (PGT 58)

and

'a pure "This" (PGT 59)

Sense-certainty as object of knowledge

(PGT 59)

'essence and instance':

'An actual sense certainty is not merely this pure immediacy, but an instance of it'
I and Object:

'in sense certainty, pure being at once splits up into what we have called the two
'Thises', one 'This' as 'I', and the other 'This' as object'

Mediation:

'each is at the same time mediated' [by the other]

'One of the terms' =

'a simple immediate being'

'the other' =

'unessential and mediated' 'not in itself'

'The object':

'is'

'it is what is true'

'it is the essence'

'It is regardless of whether it is known or not'

Sense-Certainty (2)

(PGT 59)

'What is the "This"?'

(PGT 60)

'twofold shape of its being' =

(1) 'Now'

(2) 'Here'

Q: 'What is Now?'

A: 'Now is Night'

Test = 'We write down this truth'

Result = 'If now, this noon, we look again at the written truth we shall have to say that
it has become stale'

[NB similar analysis of indexical 'Here' (PGT 60-1)

""here" is eg, the tree. If I turn round, this truth has vanished and is converted into its
opposite: "No tree is here but a house instead". "Here" itself does not vanish']

(PGT 60 cont.)

'So it is in fact the universal that is the true [content] of sense-certainty'

Language:

'We utter the universal'

'we do not strictly say in this what in this sense-certainty we mean to say'

'But language, as we see, is the more truthful; in it we ourselves directly refute what we mean to say, and since the universal is the true [content] of sense certainty and language expresses this true content alone, it is just not possible for us ever to say, or express in words, a sensuous being that we mean'

[cf (PGT 66)

'the sensuous This that is meant cannot be reached by language, which belongs to consciousness, ie to that which is inherently universal']

'the "This" shows itself to be a mediated simplicity, or a universality'

'Its truth is in the object as my object, or in its being mine'

(PGT 62)

'What does not disappear in all this is the "I" as universal'

'I can no more say what I mean in the case of "I" than I can in the case of "Now" and "Here"'

'Sense-certainty thus comes to know by experience that its essence is neither in the object nor in the "I"'

'Thus we reach the stage where we have to posit the whole of sense-certainty itself as its essence, and no longer one of its

Appendix

SENSE CERTAINTY AND THE NATURAL ATTITUDE

Part A Section I of Hegel's *Phenomenologie des Geistes* and Part II Chapter One of Husserl's *Ideen II* () have similar philosophical roles. Both are, in a sense 'starting points' of philosophies which claim to operate without unacknowledged presuppositions. Both in some measure are descriptions of what appears to pre-philosophical thinking. Both are fundamentally Cartesian in their first-person vocabulary and, finally, both are, in very different senses, 'suspended' for the conduct

of philosophy: 'Sense-Certainty' is 'aufgehoben' within self-consciousness; the 'Natural Attitude' is subject to the phenomenological ????? They each raise the question: how can we and how should we start philosophy?

My first question is 'What is it that has the Natural Attitude and what is it that has Sense Certainty?', Hegel's answer is 'we' and, derivatively, 'I'. Husserl's answer is 'I' and, derivatively, 'we'. The voice of Hegel's text is first person plural, he says 'Our approach ...', 'we much ...', 'we take ...' (58). But is this voice part of 'sense certainty'? There are reasons for and against this. Hegel says explicitly; 'Consciousness for its part, is in this certainty only as a pure 'I' (58) On the other hand who are the 'we' of the text? If he means 'I, Hegel' then 'we' is not part of sense certainty. 'Sense certainty' is being written from the standpoint of philosophy. If he means 'You the reader and me Hegel' then there is a problem. If 'sense certainty' does at least partly correspond to pre-philosophical attitudes then even though Hegel does not share the attitude the reader might well. Even if the reader is a 'philosopher he might be a 'natural attitude philosopher'. I mean by this a philosopher who does not call into question or otherwise transcend, criticise or bracket the ontological commitments of common sense here 'common-sense' means pre-philosophical attitudes, accurately characterised. There is a unanswered question then for Hegel: who are the 'we' and are 'we' in sense-certainty?

Husserl has a parallel problem but he is aware he has it and he solves it. He too begins his chapter in the second person plural; 'We begin ...', 'We shall ...' (51). Here I think 'we' does mean 'I Husserl and you the reader'. He says quite unambiguously that 'we' are within the natural attitude. He then says 'meditations best be carried out in the first person singular' (51). Why should this be? He doesn't explain the ground for 'singular' rather than plural' but I think this is to be just explained as part of the true descriptions of what the Natural Attitude is (like). Husserl's treatment of this problem of persons has then two advantages over Hegel's. 'We' are unambiguously within the pre-philosophical attitude and the transition to the first person is made deliberately and, I shall argue, is justified.

There is then an 'I' given to sense-certainty and to the natural attitude. Is this a good candidate for that which has sense certainty or the natural attitude? In Hegel's case he says that is consciousness that is as pure 'I' 'in this certainty'(58) 'or' ... 'I am in it as a pure 'This'. If 'or' here implies 'this is another way of saying the same thing' then Hegel's position is: a) In sense-certainty consciousness is and 'is as' (ist als) 'a pure 'I' (ein reines Ich), b) 'I' am in it only as a pure 'this' () c) the relation between consciousness, 'I' and 'this' is identity. There are not three entities here but three ways of specifying one entity. As he puts it: 'Consciousness is 'I', nothing more, a pure 'This' (59).

For Hegel consciousness is the subject of sense certainty but only knows itself as 'I';

as a 'this' that is - as a particular given 'I'. There is then not a numerical or ontological distinction between what is sense-certain and the self which appears to sense certainty. It is just that consciousness's self knowledge is extremely limited under this sort of knowing. That a sort of self-consciousness is possible at all in sense-certainty effect helps the transition to 'self-consciousness' but that will be trat????? later.

Who or what is the 'I' of the Natural Attitude? Husserl's answer is 'the cogito'. Is this the same answer as Hegel's 'consciousness'? It is true at least that for both authors the subject is consciousness. Husserl says 'The complexities of my manifoldly changing spontaneities of consciousness then relate to this world ...' (53) where consciousness is something I have. For Hegel is something I am, or to put it more accurately, 'I am something that consciousness is. Husserl's talk of the manifold again breaks with Hegel. Hegel says: 'The I' does not have the significance of a manifold imagining or thinking' (58). This I think, casts doubt on interpretations of Hegel's sense certainty chapter which make it correspond to closely to the positions of classical empiricism. Although Husserl and Hegel use the Kantian 'manifoldheit' here both are characterising a traditional empiricist self (Hume), Husserl to make it part of the Natural Attitude, Hegel to exclude it from sense-certainty. 'I' is an unqualified given in sense certainty, just as a 'thing' is not a substance with a bundle of qualities. For Hegel there are things and myself; not substances with properties. 'I am in it (SC) as a pure This and the object similarly as a pure This' (58). For Husserl the self of the Natural Attitude is Humean:

'I always find myself as someone who is perceiving, objectivating in memory or in fantasy, thinking feeling, desiring etc' (54)

What then is the relation of 'I' to 'consciousness' for Hegel and Husserl in these chapters? The question breaks down into two; the relation within sense-certainty and the natural attitude; the relation as conceived by the author philosophically. The second is easier than the first. For Husserl the Transcendental Ego is the subject of consciousness and consciousness constitutes objects. For Hegel Geist is the subject of consciousness and consciousness constitutes nature, where 'constitutes' means 'makes it be what it is for us'. Both authors think these are the correct accounts while locating pre-philosophical thought within them.

Despite their disagreement over the empirical self Hegel and Husserl both think it arises within and for consciousness: as an object for consciousness. Husserl lists the mental acts constituting the 'I' and says: 'All of them, including the simple Ego acts in which I, in spontaneous advertance and seizing am conscious of the world as immediately present - are embraced by the one Cartesian expression cogito. Husserl doesn't just mean it would be semantically legitimate to preface the description of

each mental act that appears to him with 'cogito', although this would be legitimate. He is drawing attention to what makes this possible. Unless seemingly date or at least distinguishable mental acts were acts constituted within one and the same consciousness then they could not be 'embraced' () by 'cogito' (of Descartes It think .. Kant). The unity of consciousness is a condition for the cogito's semantic content.

Immediacy

Both sense-certainty and the natural attitude are naive realisms. The subject of both does not know that objects are constituted by consciousness. Husserl: 'I am conscious of it: that signifies above all, that intuitively I find it immediately, that I experience it' (51). Hegel; 'Our approach to the object must ... be immediate or receptive; we must alter nothing in the object as it present itself' (58) and 'neither I nor the thing has the significance of a complex process of mediation' (58). Husserl speaks of 'a knowing of them which involves no conceptual thinking' (52)

There is such a thing as letting appear what appears as it appears (Heidegger ...) but Husserl and Hegel think this taken for granted world is constituted even though it appears immediate. This is one reason for Hegel why sense certainty 'appears as the richest kind of knowledge' (58) but is 'really' the most abstract and poorest truth' (58) It is abstract and poor because it can be explained philosophically. Similarly, Husserl's epoche has the result and knowledge within it is not as 'fundamental' as phenomenological description.

Objects for Consciousness

In Sense certainty and the natural attitude these seem to be given immediately (as does the 'I'). But Hegel and Husserl divide their account of objects into a spatial and a temporal aspect. Hegel treats these under 'here' and 'the now' respectively, Husserl under time and space.

For Husserl 'I am conscious of a world endlessly spread out in space' (51). Here 'world' means 'what it is' for the natural attitude. For Hegel: 'the this' is denoted spatially as 'here'. Both Hegel and Husserl are describing what is for pre-philosophical consciousness; a spatially extended present. Both too allow that this sort of consciousness is not exhausted by presence. Both admit absence; 'here' is For example; a tree. if I turn around, this truth has vanished and converted into its opposite; 'No tree is here, but a house inserted'. (PG 61) but despite this change of content 'Here itself does not vanish; on the contrary, it abides constant in the vanishing of the house; the tree etc and is indifferently house or tree' (61pg). Hegel's claim can be understood on a semantic or a phenomenological level and is intended on both. Semantically the claim 'Here is the tree' is falsified by Hegel's

turning away from the tree (of Hegel .. Heidegger) but the meaning of 'here' is unaffected by this change. Phenomenologically Hegel is describing the facts reported in these uses of 'here'. There is a 'here' denoted by 'here that remains constant through changes???' of content. I think that this here is 'what is given to consciousness spatially, whatever is given to consciousness spatially' and this corresponds to 'the field of consciousness' where objects are 'objects for consciousness'. 'Here' is part of 'the this' so Hegel can say 'the This' is a simplicity and a universality. 'Simple' because undivided, undifferentiated, universal because constant between shifting contents. (610)

There is a closely corresponding passage in Husserl. 'I find myself actively related ... for the most part to the actuality continually surrounding me' (54) 'the world in the usual sense of the word is and has been there for me continuously as long as I go on living naturally'. For Husserl world and I co-exist as do I and 'the This' in Hegel. Sense certainty and the Natural Attitude are paradigmatically consciousness of present objects. Just as Hegel allows himself to turn from the tree to the house. Husserl 'can let my attention wander away from the writing table which was just now seen and noticed, out through the unseen parts of the room which are behind my back, to the veranda, into the garden' (52). There is a problem about spatial metaphors here (is it clear how the imagination 'wanders') but that will have to be left aside here. For Husserl 'I' and world co-exist, he doesn't prejudice the issue by making one determined solely on the other. A position like this is in Hegel's turning away from the tree. If was his-consciousness' - action that falsified the 'here' claim. The world - self mutual dependence is very close in both authors. I and what is for I are mutually dependent for existence and essence.

Future and past to not form part of the natural attitude or sense-certainty. The reasons for this are more apparent in Hegel than Husserl. Hegel is after all describing sense certainty and in an important sense the past and future, are not perceived by the senses. In a senses no dimension of time is but present objects are perceived; this is captured in the two sense of 'present' - x may be present in time, and so present tp consciousness, present to the senses. If x is not present in time then x cannot be present to the senses, in the 'present to' sense. (present and present to). There is time in sense-certainty but only a series of punctual 'thises'.

Bibliography