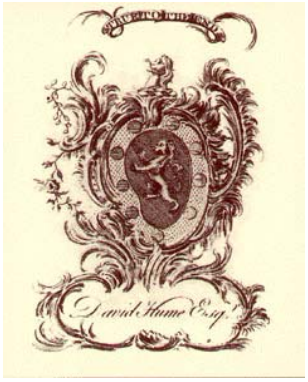

Hume's Bundles, Self-Consciousness and Kant

S. C. Patten

Hume Studies Volume 2, Number 2 (November, 1976), 59-75.

HUME'S BUNDLES, SELF-CONSCIOUSNESS AND KANT

Even if we are inclined to view Hume's attempt to explain ascriptions of personal identity as an abysmal failure, we might still be sympathetic toward his proposal to replace the going substance theory of the nature of mind with his bundle account. Thus we might fault Hume for erecting an unachievably high standard for personal identity,[1] or round on him for excluding bodily criteria from his solution,[2] but nonetheless applaud his spiriting away of the Cartesian soul substance leaving his "true idea of the human mind" (261)[3] as a collection of perceptions in its stead.

However, Hume's concern to give an account of the idea of the unity of a human mind[4] in terms of related perceptions has run into notoriously rough sledding, not the least of which is the fact that Hume himself in his Appendix to "Of Personal Identity" seems to abandon his bundle theory in despair. In this paper I shall first be concerned to examine the criticism of his bundle theory which Hume brings forth in the Appendix and explain why I think the core difficulty remains insoluble for Hume. I shall then go on to suggest that the unworkability of Hume's bundle theory is directly tied to his failure to appreciate certain features of the peculiarly first person nature of his question about the make-up of minds.[5] In thus emphasizing the place of the first person voice in my interpretation I am stealing a leaf, if not a chapter, from Kant. But this is as it should be since one of the aims of this essay is to provide an interpretation of Hume's difficulties with his account of the nature of mind which explains part of the reason why 'self-consciousness' (or 'apperception') looms so very large in the Critique of Pure Reason attack upon Hume's type of bundle theory.[6]

1. Hume's Second Thoughts

In a burst of self criticism at the Appendix to the section on personal identity, Hume brings forth this objection to his own efforts:

> *...Having thus loosen'd all our particular perceptions, when I proceed to explain the principle of connexion, which binds them together, and makes us attribute to them a real simplicity and identity; I am sensible, that my account is very defective, and that nothing but the seeming evidence of the precedent reasonings cou'd have induc'd me to receive it (635).*

After additional reflections Hume encapsulates his objection thus:

> *...All my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness (635-6).*

What is the defect Hume calls to our attention? Notice first that the footnote on page 635 explicitly refers us back to that part of his account Hume considers defective; and when we look back to the text of the Treatise in question, page 260, we find it is precisely there that Hume uses the analogy of a republic to sketch his bundle theory of the unity of mind in terms of the two associative principles, resemblance and causation. So it must somehow be that this account in terms of the associative principles is mistaken. But what then is the mistake? According to Hume it is that he cannot 'explain' the principles which unite the discrete percepts. Now it seems to me there are only two possible things Hume could mean in saying this. He could mean either (1) that the principles which connect the percepts of one mind are instances of the associative principles, but he cannot explain them; or (2) there must be other, non-associative, principles which connect the percepts of a single mind and these he cannot explain.

I think we can safely conclude (1) cannot represent Hume's misgivings when he says he is unable to explain the uniting principles of the mental bundles. There is first the fact that he does not call into question his general programme of accounting for the nature of the unity of complex things by the principles of causation, resemblance and contiguity -- for example, ships, plants, animals, houses and rivers (e.g. 257f) -- but only his attempt to apply this programme to minds. Second, and conclusively, if -- as we are now supposing -- the connecting principles are indeed instances of the associative principles, then he can only mean that he cannot explain his theory of association of ideas. Yet he tells us in separate places in the Treatise that our failure to explain the associative principles is not a defect at all. The associative principles are first principles, ultimate principles in the science of human nature, and

> *...If this impossibility of explaining ultimate principles should be esteemed a defect in the science of man, I will venture to affirm that 'tis a defect common to it with all the sciences, and all the arts...* (xxii. See also 13).

Thus, unless Hume is to be judged flatly inconsistent, it must be (2) which represents Hume's mood: the principles which connect discrete percepts cannot be instances of the principles of association, and these new principles cannot be explained by him. But why should he say that these new and non-associative principles cannot be explained? I think there can be only one answer here: it is just because the relations between distinct perceptions cannot be associative ones that it is impossible for him to explain them. For it is a basic dictum of the Treatise that to explain the relation between what are called distinct existences is always to establish an associative link - for Hume that is just what

it is to explain the relations between matters of fact. Hence, since perceptions are distinct existences, to complain that the relations between them cannot be explained by him can only mean that the relation is not an associative one.

Why then should Hume have second thoughts? How does his attempt to explain the unity of a human mind as bundles of perceptions in causal and resembling relations break down? For Hume an account of the unity of a human mind is to proceed in a fashion similar to the phenomenalistic rendering given for entities like plants and animals in the Treatise: to say distinct percepts are states of one mind can only mean that they are found to stand in relations of resemblance or causation.[7] But how can this possibly do? On Hume's account, my assertion that this is one plant is to be accounted for as a claim that various plant states exhibit the relations of causation, resemblance and spatial contiguity: for example, what I see now (these flowers) is a causal effect of what I saw previously (those budding branches). But there seems no similar way in which the perception of the coffee cup I see before me now can be said to be the effect of the perception of my pen which immediately preceded it. It is not that some sort of non-Humean notion of causation is being used here against the analysis -- let invariable succession be the mark. Then it is just a matter of fact that my perception of this pen is often followed by many different percepts -- of this book, that writing pad, those scribbled notes. Shall we then say that one of the two perceptions in question resembles the other? But this is even more absurd. For Hume, to say one perception resembles the other is to say that it is an image of the same or a similar thing as the other, but if my perception of a coffee cup is to be an image of anything at all it must be 'of' just that, a coffee cup,

and not a pen.

It might be thought there is a way in which any two of my perceptions can be said to be causally related: it is just that it is always the case that one of my perceptions is invariably followed by another of my perceptions. However, it is quite clear that Hume does not mean anything so empty as this. When giving us instances of the causal relations which are supposed to obtain between the perceptions of a single mind Hume cites as an example the fact that "our impressions give rise to their correspondent ideas" (261). And if we are to let the rest of the Treatise act as our guide in interpreting what this means, it can only be that, for example, the impression of a tree gives rise to its corresponding idea -- the idea of a tree. In other words, the terms of the causal relations in question are to be individuated - at least in part - by the objects or propositional elements of the percepts. But textual matters aside, it is quite clear that if one proposes to explain the idea of a single mind or mental history, and if this problem is supposed to arise, as it does with Hume, because one's various mental states differ, then it is simply to refuse an explanation to say one of my mental states is always followed by another of my mental states.

Another way of putting Hume's self criticism in the Appendix is that he perhaps comes to see that the analogy he strikes between the unity of a republic and that of a human mind misfires. For it is interesting to note on this score that Hume does not make use of the examples he has discussed at length and which are naturally explained as causal wholes -- plants and animals -- when the unity of a mind is in question. Rather he turns to a different kind of entity, a republic, where the changes of this 'variable object' are the changes of members and changes of laws (261). But here, surely, "the reciprocal

ties of government and subordination" (261) insofar as they do explain the unity of a republic, are not causal relations. Of course the citizens may bring on a new government and the leaders may allocate positions for the citizens within the privilege system. But the republic is not one republic because these causal relations obtain; instead it is that the causal relations can come about because there are already reciprocal ties -- substantive laws and procedural rules -- which define the rights and duties and the means of effecting change within a single republic. Thus, if there is any point at all to this new analogy in helping us understand the unity of a human mind, it would seem to be one that is antagonistic to Hume's programme: that the various mental states are somehow non-associatively related and it is the presence of these other relations which accounts for the unity of a single mind.

Someone may want to respond that even though it may be true that the Humean relations of resemblance and causation do not hold between every pair of the percepts of one's mind, nonetheless such relations do hold between many subsets of the perceptions of a single mind, and taken together these resembling and causally related subsets constitute one mind for Hume. In response to this suggestion I wish to maintain that given the skeptical or critical strain of the Treatise, this problem of the unity or the oneness of the mind must still remain insoluble for Hume. Here I begin to echo a Kantian theme which centers on the fact that it is often Hume's concern to emphasize the radical contingency of uniformities in the world.

For, as we all know, Hume will say that even though it has always happened that a flame gives heat, it could come about even in the most ideal circumstances that I perceive a flame and receive a feeling of cold from

it. But to say it could come about that I perceive a flame and then experience cold is just to say my perception of a flame need not always be followed by a perception of heat. And this is to suppose that it does make sense to speak of the perceptions of one mind even where the causal and resembling relations do not obtain between the percepts. Indeed, Hume generalizes this point about the radical contingency or accidentalness of the order of one's percepts to apply to all the associative relations. Toward the beginning of the Treatise where Hume explains his notion of association of ideas, he explicitly recognizes that in the absence of associative connections we could still speak of one mind. Thus

> *This uniting principle among ideas is not to be consider'd as an inseparable connexion; ... nor yet are we to conclude, that without it the mind cannot join two ideas; for nothing is more free than that faculty: but we are only to regard it as a gentle force, which commonly prevails...(10).*

Generally speaking human minds operate via the gentle force of association. However it is essential to the Humean picture that there is no necessity in this. For even where there is no uniting principle of association there can still be a mind, viz., one which joins ideas in an entirely 'fanciful' way. So again, Hume's critical thoughts of the Appendix are perfectly understandable: if we can speak of a mind where association is not present, we cannot expect to explain the idea of a single mind in terms of a bundle of the perceptions connected by the associating principles of causation and resemblance.

Thus the depth of his despair compels him to toy with doctrines he previously rejected with confidence:

> *Did our perceptions either inhere in something simple and individual, or did the mind perceive some real connexion among them, there wou'd be no difficulty in the case. For my part, I must plead the privilege of a sceptic, and confess, that this difficulty is too hard for my understanding (636).*

But of course: if he could say that his perceptions were states of a mental substance "then there would be no difficulty in the case", i.e. he could thereby explain the unity of his mental states. Likewise, if he could discover some "real connexion" between perceptions -- minimally, some connection such that in its absence it could not be said that the perceptions made up one mind -- he could thereby explain the idea that he has, viz., that his perceptions constitute a unity.

Hume's problem seems peculiarly poignant when we recall that much of his discussion of unified objects in the Treatise seems intended to show us that when we speak of various qualitative states of one thing we mean to refer to more than a heap of parts. A mere collection of limbs and trunks and roots and leaves is a future bonfire, not a living tree. By analogy - and this is the kind of analogy Hume wants to stress - it would seem that the perceptual states of a single mind would have to exhibit some connection; by analogy it would seem that just any collection of perceptions could not account for the idea of the unity of one's mind. Yet as we have seen by the way he treats his connecting principles (his associative ones) as merely contingent links, he is forced to admit that the perceptions of a mind need not be connected in any way whatsoever.

This seems the proper place to consider two intriguing counter interpretations of Hume's misgivings in the Appendix. Although it is probably the case that our agreements in interpretation outweigh our disagreements, Stephen Nathanson voices a rendering of Hume's despair that, at least in part, runs counter to mine.[8] According to Nathanson Hume's first thought as enunciated in "Of Personal Identity," is that a mind is nothing but a bundle of perceptions; but then, belatedly (in the Appendix), he comes to realize that he has tucked something extra into his bundles, viz., the principle of association.

Thus, for Nathanson, the conflict: Hume wants to say that the idea of a human mind is simply that of a collection of perceptions, but he finds himself saying as well that this idea involves a set of perceptions and the principle of association.

Now, as Nathanson notes, one will be quick to respond that there need not be any real conflict if, as seems natural, "...the principle of association is reducible to the principles of resemblance, contiguity and causality",[9] i.e., features of the percepts. But Nathanson meets this objection head on. According to him there is strong textual evidence that for Hume the principle of association is not reducible to relations between perceptions, but involves mental propensities (of the imagination and memory) as well.[10] So now the conflict that Nathanson finds can take its final form: Hume first proposes to account for the idea of a mind solely in terms of percepts, and yet he finds that he must also invoke mental propensities which he himself grants cannot be reduced to perception talk.

It seems to me that Nathanson does open up new lines of inquiry and nicely suggests one way in which Hume's bundle theory might be defective as a complete account of mind, but nonetheless I think the crucial question remains: is this a fair treatment of Hume's intentions, does it properly diagnose Hume's despair in the Appendix? I have my doubts. For Nathanson, in citing the same passages I quote at the beginning of this paper,[11] must thereby acknowledge that Hume says that he cannot explain the principles which unite his percepts; and on Nathanson's account this must mean that he cannot explain those principles captured by the label of 'association' and understood as propensities of mind. But whether or not we view the principle of association as referring to mental propensities, the fact remains -- as I have argued earlier on in this paper -- that Hume sees the principle of association as a first principle, acknowledges that it

cannot be explained, and insists by analogy with other sciences that this is not a defect. Thus, if my reasoning here is sound, Nathanson's reading of Hume's intent in the Appendix entails a curious inconsistency in Hume's attitude toward the principle of association.

The other counter interpretation I wish to consider could cause difficulties for both Nathanson and myself.[12] On this account it is noted that in the Treatise Hume voices what is sometimes referred to - unhappily, I think - as the 'no ownership' theory of mind. Thus Hume: "...there is no absurdity in separating any particular perception from the mind;..."(207). What a remark like this seems to suggest is that Hume, mistakenly without doubt, considers it logically possible for perceptions to exist unowned. Now the thrust of this counter comes right here: Isn't it simpler, and perhaps more natural as well, to view Hume's second thoughts in the Appendix as a recognition that the logical independence of perceptions precludes the possibility of even the looser and tentative sort of connection that Hume attempts to elucidate by his analogy with a republic?

On first view this interpretation does seem tempting, but reflection prompts doubts. My reasons for saying this are both textual and logical. First, even though Hume's description of perceptions as distinct existences does seem to commit him to the view that perceptions can (logically) exist unowned, it is a fact that he argues that there are "a few...experiments, which convince us, that our perceptions are not possest of any independent existence" (210).[13] In other words, even though Hume seems wedded to the position that a proposition like 'All S's perceptions are S's' is not logically necessary, he nonetheless constructs a case for claiming that that same proposition represents a compelling contingent belief (it is a belief which is confirmed by

empirical experiments). Second, and connected, there need be nothing inconsistent for Hume to continue on and give a psycho-genetic account of this supposedly factual belief - just as he seems to do in the republic passage. Indeed, if we let Hume's treatment of causation guide us as a model - where he first argues that causal beliefs are not logically necessary, then acknowledges their compellingness, and then explains their genesis -- the heart of the programme of the Treatise would recommend just this treatment of the idea of mental unity. Thus for these reasons, I think the counter interpretation in question, if accepted as an explanation of Hume's misgivings in the Appendix, leaves us with a serious puzzle. Why should Hume find his account of the putatively contingent belief in mental unity in the republic passage inconsistent with the fact that that belief is not logically necessary?

2. Reflecting on Hume's Question: The Kantian Turn

How does one go about answering Hume's question about the unity of his mind? Better still, does Hume's question deserve an answer? I should like to make one suggestion as to how Hume's dilemma about the unity of the mind is to be resolved: that Hume's question does not need a restating or reformulating so much as an underscoring of one feature of the question. If I am successful in establishing this point, part of the anti-Humean importance of 'self-consciousness' in Kant's transcendental deduction should be obvious.

Consider one important difference between asking a question about the nature of one's own mind and asking a similar question about the mind of another. When, as an heuristic device, Hume proposes that we imagine looking into the breast of another, he suggests that we would see a connected concatenation of perceptions, some in causal relations with others, some resembling others (260). But now note that when Hume supposes that we look into another's

breast, say S, the whole question of S's rationality or S's sanity is left entirely undetermined. Thus, rather than the orderly mind of Hume's acquaintance I might choose to picture myself perusing a peculiarly untidy mind, the mind of a lunatic say, where the mental life is one of total disarray. Yet, to draw out the significant difference in the first person case, when I look within my breast, when I make my own perceptions objects of introspection, my rationality or sanity is assumed, at least in a minimal sense. The minimal sense of rationality I am referring to here is just that of being able to distinguish between the objects of awareness, to be able to say things like 'this is a ϕ' and 'this is a ψ'. In the case where the objects of awareness are elements of one's own mental life this facility for distinguishing between one's perceptions comes down to an ability to describe one's perceptions, an ability to say (as does the author of the Treatise), 'this is a perception of a house and here is one of a chimney'.

The point of the previous remarks is that Hume's first person question about the unity of his mind is a significant one only if it is taken in a certain way. I am suggesting it should be spelled out thus:

> What sort of connection must there be between the perceptions of my mind if I am to know that they are the perceptions of my mind, rather than not to have any knowledge at all about my perceptual states?

In underlining the second half of this question I mean to reiterate the point made at the beginning of this section to the effect that the solution to Hume's problem lies in appreciating where the emphasis should be placed. For Hume's question cannot significantly be

> What sort of order must there be between my perceptions if I am to know that they are the perceptions of my mind, rather than to have some other bit of knowledge about the ownership of these perceptions?

for there just is not any other bit of knowledge I could have about the ownership of my perceptions (e.g. I could not know they were yours, nor could I know they belonged to nobody).[14] On the other hand, we can make sense of the idea that one's perceptual life might be in such a disarray that one can have no consciousness at all that his perceptions are his. It is this feature, the fact that one's question must be about the connection required for being able to introspect one's experiences, that Hume seems to miss in his reflections.

Let me put my point a bit differently. Hume has been credited for excising the notion of mental substance from deliberations about the nature of mind. But in executing this bit of philosophical surgery Hume is forced to explain the knowledge he has of his own mind in terms of connections between his perceptions. So how does Hume explain the unity of his own mind? What I am suggesting is that Hume cannot even begin to answer his question until he realizes that his first person question calls for a different treatment than questions of similar form which are requests for explanations of the unity of a thing.

That is to say when I ask, as Hume does (257), about the unity of a complex thing like a ship I could put my question as

(1) What sort of connection must there be between the parts of a ship for me to know that they are all parts of a single thing, a ship?

But I could just as well ask

(2) What sort of connection must there be between the parts of a ship for one to know that they are all parts of a single thing, a ship?

Yet I could press essentially the same question with

(3) What sort of connection must there be between the parts of a ship for it to be a single thing, a ship?

My point in showing the natural reduction of the question asked in (1) to that posed in (3) is just the rather

obvious one that usually when we ask about the unity or oneness of a complex entity the reference to the knowing subject is otiose -- a discussion of the knowing subject is not essential to an account of the idea of a complex thing like a ship.

On the other hand, when I ask

(1') What sort of connection must there be between my perceptions for me to know that they are all perceptions of one thing, my mind?

I cannot exchange this question for

(2') What sort of connection must there be between my perceptions for one to know that they are all perceptions of one thing, my mind?

Nor can I be asking

(3') What sort of connection must there be between my perceptions for them to be perceptions of a single thing, my mind?

For even though (1') is a perfectly sensible question about the unity of my mind, it cannot, by analogy with other unity questions, be reduced to the questions posed in (2') and (3'). The reason this is so is that whereas (1') does make sense, (2') and (3') do not.[15] For, with regard to (2'), there could not be any order which explains how one (i.e. 'anyone') knows my perceptions are mine, just because logically no one else could be acquainted with my perceptions. Likewise with (3'). Unless one takes it as a truncated or abbreviated form of (1'), (3') is as well a non-question. For there can be no connection among my perceptions which explains how they are states of my mind rather than those of another, or simply free-floating. With regard to (2') and (3') one is perfectly justified in responding with an analytic club: 'Your perceptions are yours, period'. On the other hand (1') does make sense because the contrast implicit in asking it is just that between one who is able to know or be aware of his perceptions and one who can have no such ability. And the very fact that (1') cannot be reduced to (2') and (3') in a way analogous to the reduction we can make in questions

about the unity of other compex things, shows us that reference to the knowing subject must be an essential element in any account of what sort of connection there must be between my perceptions if I am to know that they are all mine, those of my mind.[16]

Conclusion:

The moral of these deliberations is that I want to follow an insight of Kant's and fault Hume for not perceiving clearly the nature of his first person question. That is to say, contemporary philosophers are right, I think, to applaud Hume's attack on the idea of mental substance. Equally Hume is to be commended for seeing that any account of the belief that all my perceptions have a unity must be given in terms of a certain connection between my perceptions. But what he does not see - and this seems to be the source of his perplexity in the Appendix - is that a first person question about the unity of my mind has sense only if it is a question about what sort of perceptual connection is required if I am to know or be aware that I have them, rather than not to be aware of my perceptions at all.

Thus I am suggesting that if Hume had seen the full force of his first person query about the unity of his mind he would have discovered at least the beginnings of an answer to his question. It would have been an answer which anticipated Kant in stressing the importance of the concept of self-consciousness.[17]

S. C. Patten
University of Lethbridge

1. See Terence Penelhum, "Hume on Personal Identity," reprinted in Human Understanding, ed. by A. Sesonske and N. Fleming (Belmont: Wadsworth, 1965), 99-114.

2. See David Pears, "Hume on Personal Identity," David Hume: A Symposium (London: Macmillan, 1966), 43-54.

3. Here, as throughout, page references in the text refer to A Treatise of Human Nature, ed. by L. A. Selby-Bigge (Oxford: Oxford University Press, 1888; reprinted, 1965).

4. In this paper I use the locution, 'an account of the idea of the unity of mind', and similar expressions with the same sense, in such a way that one could view Hume as offering either a meaning account or a psycho-genetic one. I adopt this ambiguity so that the soundness of the points of this essay can be considered independently of whether or not one is inclined to say that Hume is engaged in meaning analysis or psychology. However, for compelling arguments that Hume's discussion of mental unity is predominately psychological see Penelhum's recent paper, "Hume's Theory of the Self Revisited," Dialogue, vol. XIV, No.3, 389-409, esp. 403-6.

5. In saying that Hume misses certain features of his first person question I do not mean to suggest that he fails to see that his concern with the unity of mind is a first person enterprise. Although brevity precludes discussion of this point, I agree completely with the tradition which holds that Hume is not dealing with the problem of individuating minds or persons in "Of Personal Identity", but rather aims to account for the unity one ascribes to one's own perceptual life. On this see especially Penelhum, ibid., 399f. and 401-3.

6. It is true that Kant never mentions Hume's bundle theory by just that title, but the Humean separability of perceptions is a clear target of the Critique of Pure Reason.

7. Spatial contiguity has, of course, no place in Hume's discussion of minds since perceptions do not take spatial predicates (see 260).

8. "Hume's Second Thoughts on the Self," Hume Studies, 1-12. See especially sections 4. and 5.

9. Ibid., 10.

10. Ibid., 10f.

11. See above, 2, and Nathanson, ibid., 5.

12. I am grateful to the anonymous referee for Hume Studies for making clear the force of the objection which follows.

13. Or again: "...When we compare experiments, and reason a little upon them, we quickly perceive, that the doctrine of the independent existence of our sensible perceptions is contrary to the plainest experience" (210). For a lucid critique of these 'experiments' see Jonathan Bennett, Locke, Berkeley, Hume (Oxford: Oxford University Press, 1971), 346f.

14. This point has been made as a charge against Hume. See e.g. Pears, op cit, 52.

15. There are abundant 'psychological' analogues to (1') which do not reduce to their corresponding (2') and (3') forms. Thus, e.g., the myopic individual may ask what he needs to see the road or the chair or his book, without suggesting in the least that others may require the same therapy, or that the information he desires bears on the ontology of roads, chairs and books.

16. At this point I have only been arguing that the Kantian variant of Hume's mental unity question represented by (1') makes sense. Whether or not Kant can actually establish that a particular connection among perceptions is needed for self-consciousness is a complicated matter that goes well beyond the purposes of this paper. The relation between self-consciousness and perceptual connection in Kant is discussed in part in my "An Anti-Skeptical Argument at the Deduction," Kant-Studien, forthcoming.

17. I am grateful to several people for helpful criticism: my colleagues at Lethbridge, especially Ron Yoshida, Mike Kubara and Peter Preuss; Professor Graham Bird at Stirling; and the anonymous referee for Hume Studies. This paper was completed while assisted by a Canada Council research fellowship.