



## Is there a good argument against the incorrigibility thesis?

Frank Jackson

To cite this article: Frank Jackson (1973) Is there a good argument against the incorrigibility thesis?, *Australasian Journal of Philosophy*, 51:1, 51-62, DOI: [10.1080/00048407312341031](https://doi.org/10.1080/00048407312341031)

To link to this article: <http://dx.doi.org/10.1080/00048407312341031>



Published online: 18 Sep 2006.



Submit your article to this journal [↗](#)



Article views: 242



View related articles [↗](#)



Citing articles: 10 View citing articles [↗](#)

FRANK JACKSON

IS THERE A GOOD ARGUMENT AGAINST  
THE INCORRIGIBILITY THESIS?

By the incorrigibility thesis (ICT), I mean the traditional empiricist thesis that it is logically impossible to be mistaken about certain of one's current mental states; in particular, that it is logically impossible to be mistaken whether one is in pain, has an itch, is having a red after-image, and the like. For example, ICT holds that '*S* believes at *t* that he is in pain at *t'*' logically entails '*S* is in pain at *t'*'.

There are a number of theses about the logical status of a person's beliefs about his own mental states which are closely related to ICT: for example, that '*S* is in pain at *t'*' entails '*S* believes at *t* that he is in pain at *t'*' (self intimation), and that a person's beliefs about his mental states may be false but cannot be shown false by anyone else (privileged access).<sup>1</sup> We will, however, be concerned just with whether there is a good argument against ICT as defined. My conclusion will be that no argument so far offered could reasonably be thought decisive, and so, that the possibility of false belief about whether one is in pain, seeming to see red, and so on, has still to be established.

My reason for embarking on this essentially negative task is to counter the widespread belief that ICT is not worth serious consideration.

The first two arguments I consider concern the fallibility of memory.

1. *The first memory argument*

In 'Is Introspective Knowledge Incorrigible?', D. M. Armstrong argues thus:

. . . consider such a report as 'I was in pain a few seconds ago', is it not a paradigm of an empirically indubitable statement? . . . mistake may be logically possible, but it is empirically impossible. Now what the upholder of the logical indubitability of current introspective reports has to maintain is that the logical character of our certainty *changes* as we move from the past to the present. Remember here that the experience does not even need to have occurred

<sup>1</sup> For careful discussions of how to define incorrigibility and related notions see W. Alston: 'Varieties of Privileged Access', *American Philosophical Quarterly* 8 (1971) pp. 223-41 and G. Nakhnikian: 'Incorrigibility', *Philosophical Quarterly* 18 (1968) pp. 207-15.

some seconds ago for error to be (logically) possible. Place the experience the merest fraction of a second in the past, and it is intelligible to say that error has occurred. Are we prepared to say that this fraction of a second changes the nature of our certainty at that error, from being empirically impossible, becomes logically impossible?<sup>2</sup>

It is hard to see the force of this rhetorical question. There are many cases where a fraction of a second changes the whole nature of one's certainty. For example, the table I am writing on is, and has for some time been brown; and I am absolutely certain of this fact. Nevertheless my certainty that it is false that this table is brown now but was brown a millionth of a second ago, is *empirical*; while my certainty that it is false that this table is brown now and not brown now, is *logical*.

It appears that Armstrong is trading on an equivocation in 'the nature of our certainty'. If this means 'the *psychological* nature of our certainty' then it is quite true that the nature of our certainty will not change in a fraction of a second. But this fact will have no bearing on the logical issues at stake. On the other hand, if 'the nature of our certainty' means 'the *logical* nature of our certainty', then, as the example of the preceding paragraph shows, there is no difficulty at all in supposing that the nature of our certainty changes in a fraction of a second.

## 2. *The second memory argument*

Armstrong has a second argument turning on the fallibility of memory.

Suppose I report 'I am in pain now'. If we take the view that the latter reports a piece of indubitable knowledge, to what period of time does the word 'now' refer? Not to the time before I started speaking, for then I am depending on memory, which can be challenged. Not to the time after I finish speaking, for then I depend on knowledge of the future, which can be challenged too. The time in question must therefore be the time during which the report is being made. But then it must be remembered that anything we say takes time to say. Suppose, then, that I am at the beginning of my report. My indubitable knowledge that I am in pain can surely embrace only the current instant: it cannot be logically indubitable that I will still be in pain by the time the sentence is finished. Suppose, again, that I am just finishing my sentence. Can I do better than *remember* what my state was when I began my sentence? So to what period of time does the 'now' refer?

. . . so what becomes of the alleged indubitability of the statement 'I am in pain now' when I speak at normal speed?<sup>3</sup>

<sup>2</sup> D. M. Armstrong: 'Is Introspective Knowledge Incorrigeable?', *Philosophical Review*, LXII (1963) pp. 417-32, p. 420; the argument also appears in his *A Materialist Theory of the Mind* (1968) p. 104.

<sup>3</sup> *A Materialist Theory of the Mind*, pp. 104-5; the argument also appears in 'Is Introspective Knowledge Incorrigeable?', pp. 420-1.

This argument involves a misunderstanding of the sense of 'statement' for which the incorrigibility thesis asserts that statements about one's present state of mind are incorrigible. The incorrigibility thesis is not a thesis about sentence tokens of certain kinds: it is not a thesis about particular utterances or inscriptions. It is a thesis about statements considered as divorced from any particular language and from any particular manifestation in that language; it is, that is, a thesis about certain propositions.

Thus, questions about how long it takes to utter 'I am in pain now' or, for that matter, to write 'I am in pain now') are irrelevant, because propositions (and sentence types, if it comes to that) do not have temporal duration.

It is, of course, still relevant to ask what time 'now' in 'I am in pain now' refers to. But answering this question does not pose the incorrigibilist's particular problem. He holds that 'I believe at  $t$  that I am in pain at  $t$ ' entails 'I am in pain at  $t$ '; thus he holds, substituting 'now' for ' $t$ ', that 'I believe now that I am in pain now' entails 'I am in pain now'. The requirement on 'now' in 'I am in pain now' is, therefore, that it refers to the same time as the 'now' in 'I believe now that I am in pain now'.

The next kind of argument against ICT that I consider turns on various doctrines about the meaning of predicates.

#### *Classification arguments*

The meaning doctrine involved in the first such argument is that predication necessarily involves classification: to say that  $x$  is  $F$  is to place  $x$  in the class of  $F$ s. R. A. Imlay puts the argument thus:

To believe that a particular sensation . . . has a certain quality is to believe that it belongs to the class of things having that quality. But that there is a class of things having that quality is something I remember. It is, therefore, something that I can misremember. It may be that there is no such class or that, although there is such a class, the quality I attribute to my present sensation is not one had by its members. In either case the belief I have about the nature of my present sensation will be a false one.<sup>4</sup>

How is Imlay using the word 'class' here? The obvious hypothesis is as a synonym for 'set' in set theory. In this case, Imlay's doctrine that 'to believe a particular sensation has a certain quality is to believe that it belongs to the class of things having that quality' seems reasonable enough. If the doctrine that memory is necessarily involved in my belief that there is such a class will be false. True, I will need memory to know that certain past sensations *belong* to the class; but the existence of the class, in the set-theoretic sense, does not depend on the existence of past memories.

R. A. Imlay: 'Immediate Awareness', *Dialogue* 8 (1969) pp. 228-42; see pp. 235-6.

It appears Imlay means by 'the class of things', not the set, but the things (members) themselves. So his doctrine about what it is to believe that a sensation has a certain quality is that it is to believe the sensation is like, in the relevant respect, other sensations having the quality.

But precisely how are we to read this doctrine, call it *D*, that believing a sensation is *F* is to believe it is like other sensations that are *F*? If we read *D* as that to believe a sensation is *F* is to believe it is like *certain given* other sensations that are *F*, then *D* certainly entails the falsity of incorrigibility. And if the certain given sensations are past, then Imlay will be right to bring in memory. However *D*, so read, is obviously false. It is obvious that one can believe a sensation is *F* without believing any other *particular* sensation is *F*. Moreover *D*, so read, will imply that it is impossible to believe a sensation is unique in being *F*. On the other hand, if we read *D* as: to believe a sensation is *F* is to believe it is like *any* other (if there be such) sensations that are *F*, *D* is reasonable but does not entail the falsity of ICT. *D*, so read, does not require me to believe some other given, perhaps past, sensation is *F*, about which I can be wrong; it only requires me to believe any other sensation which is alike in being *F*, is alike in being *F*; and I cannot be wrong about that, for it is a tautology.

In this connection it is worth considering Don Locke's version of the classification argument. On the question of whether judgements of immediate perception are incorrigible (can be mistaken), he argues:

To say something is e.g. blue is to classify it, to class it along with other things that happen to be blue, such as the deep sea and the summer sky. Clearly it must be possible to make mistakes, to classify it along with the wrong things.<sup>5</sup>

In reply to A. J. Ayer's objection that to say something is blue is not to say that anything else is blue,<sup>6</sup> he argues:

But we are not saying that 'blue' means 'like the deep sea, the summer sky', etc. We are saying, merely that 'blue' means 'like such things as are blue'. And, as it happens . . . the deep sea and the summer sky are blue. So to call something blue is, as it happens, to classify it along with the deep sea and the summer sky.<sup>7</sup>

What Locke overlooks here is that the (true) contention that if *x* is blue, then, as it happens, *x* is like the deep sea (in colour), conjoined with the possibility of mistake about *x* being like the deep sea, is not sufficient to enable him to infer that mistake about *x* being blue is possible. The fact that, taking *A* to be a current after-image of mine, it is clear that I can be mistaken about *A* being like the deep sea, shows that I may believe *A* is like the deep sea when it is not; and so, that I may believe that *A* is like the deep sea when it is not blue. But nothing follows from

<sup>5</sup> D. Locke: *Perception and Our Knowledge of the External World* (1967). See p. 191.

<sup>6</sup> A. J. Ayer: 'Basic Propositions', *Philosophical Essays* (1954). See pp. 63-4.

<sup>7</sup> *Op. cit.*, p. 191, n.2.

this about the possibility of believing that *A* is *blue* when it is not; for 'I believe that *A* is like the deep sea' does not entail 'I believe that *A* is blue'.

Perhaps the argument Locke has in mind is that if it is possible that I believe that *A* is blue when it is not like the deep sea, then, as *A* is *blue* only if it is like the deep sea, it is possible that I believe *A* to be blue when it is not blue. However, the incorrigibilist has a clear reply to this. He holds that 'I believe that *A* is blue' entails '*A* is blue', and thus, that the possible world where I believe *A* is blue when *A* is not like the deep sea, must be the possible world where the deep sea is not blue; and in such a world it is false that something is blue only if it is like the deep sea.

A further modification of the classification argument, which stresses the fact that knowing the established meaning of a predicate involves memory, and which is not exposed to the objections above, is put by J. H. Chandler thus:

An incorrigible description must not venture beyond one's present experience, or rely on memory. Therefore predicates with previously established meanings will not do.

This entails however, that to be used incorrigibly, every successive application of a predicate will be the use of a new and distinct predicate, whose meaning is ostensibly specified by reference to the thing it is currently being applied to. . . . The only alternative to this kind of futility is that on many, indeed most occasions one must use words according to rules established or specified independently of the particular instance the word is being currently applied to. This of course involves the possibility of error.<sup>8</sup>

In short, Chandler's argument is that, to take the familiar pain case, to know that 'is a pain' is a predicate which applies to a current mental state of mine, if it is to know something substantial, involves knowing the (or an) established meaning of 'is a pain'. And this involves memory, and so, the possibility of error.

This argument rests on a false premise. It shows that 'I believe that the predicate "is a pain" applies to a current mental state of mine' does not entail 'I am in pain'. But it does not show that 'I believe that I am in pain' does not entail 'I am in pain'; that is, it does not show that it is logically possible to be mistaken about whether one is currently in pain. It does not show the latter because 'I believe that the predicate "is a pain" applies to a current mental state of mine' is not logically equivalent to 'I believe that I am in pain'. A Frenchman who knows no English may believe that he is in pain, but he will not believe that the predicate 'is a pain' applies to a current mental state of his. (This point also applies to the common objection to ICT that applying predicates is a skilled, learnt activity. The objection confuses believing that one is in a certain mental state with believing that a certain predicate applies to one's mental state.)

<sup>8</sup> J. H. Chandler: 'Incorrigibility and Classification', *Australasian Journal of Philosophy* 48 (1970) pp. 101-6, see p. 104.

Chandler might reply that, although to believe that one is in pain is not to believe something about a predicate in English, it nevertheless is to believe something about some predicate synonymous with 'is a pain'. Thus he might suggest that 'I believe that I am in pain' is logically equivalent to 'There is a predicate  $P$  such that  $P$  is synonymous with "is a pain" and I believe that  $P$  is true of a current mental state of mine'. This analysis, however, is not without its difficulties. Dogs presumably believe that they are in pain without having any beliefs about predicates, and, further, the purported analysis does, while the analysandum does not, entail the existence of the predicate 'is a pain'.

Moreover, waiving these difficulties, it is hard to see how this analysis could help Chandler's argument. For  $P$  does *not* need to have an established meaning: the fact that  $P$  is synonymous with 'is a pain', which has an established meaning, ensures that to believe  $P$  true of a current mental state is to believe something substantial.

#### 4. *If you can't be wrong then you can't be right*

In 'Is Introspective Knowledge Incorrigeible?', Armstrong derives from the *Philosophical Investigations* the following argument:

If introspective mistake is ruled out by logical necessity, then what sense can we attach to the notion of gaining knowledge by introspection? We can speak of gaining knowledge only in cases where it makes sense to speak of thinking wrongly that we have gained knowledge. In the words of the slogan: 'If you can't be wrong, you can't be right either.' If failure is logically impossible, then talk of success is meaningless.<sup>9</sup>

What is the principle of this argument? It seems to be most clearly stated in the slogan. Now, to be wrong is to believe  $P$  when not  $P$ . To be right is to believe  $P$  when  $P$ . Therefore, if we read the 'can't' in the slogan as indicating logical impossibility, the principle of the argument comes out as:

If Nec  $\sim(\sim P \ \& \ S \text{ believes } P)$  then Nec  $\sim (P \ \& \ S \text{ believes } P)$ .

This principle is invalid. If we substitute 'I' for 'S', and 'I exist' for 'P', we get a false substitution instance; the antecedent being true and the consequent false.

In view of the talk of meaninglessness in the last sentence in the above quote, it might be suggested that the slogan should be read as: If you can't be wrong, then it is *meaningless* to suppose one is right. This reading, however, does not help matters. It gives as the principle of the argument:

If Nec  $\sim (\sim P \ \& \ S \text{ believes } P)$  then  $(P \ \& \ S \text{ believes } P)$  is meaningless.

and this principle may be shown invalid by the same substitution instance as before.

<sup>9</sup> *Op.cit.*, p. 422.

### 5. The distinct existences argument

The most widely canvassed argument against ICT turns on the claim that belief in (or awareness of, or apprehension of) a mental state is distinct from that mental state.<sup>10</sup> We will consider the version put by Armstrong. He argues:

The acquiring of introspective knowledge must consist of the making of (sincere) reports of current mental occurrences, or else a non-verbal apprehension of these occurrences. In both cases the apprehension of the occurrence will have to be *distinct* from the occurrence that is apprehended. But if this is granted, then we can apply Hume's argument about 'distinct existences'. Wherever we have two distinct things, Hume points out, there we can always conceive of the one existing in the absence of the other. It follows that it is logically possible to have a sincere report of a current inner experience, or a non-verbal apprehension of that experience, without the experience existing . . . But this state of affairs would be the state in which we would be mistaken about our current inner mental state.<sup>11</sup>

In support of his claim that mental state and non-verbal apprehension are distinct, Armstrong argues:

. . . the apprehension must be distinct from the thing apprehended. For if not, we are faced with a flagrant circularity. Having a pain logically involves apprehension of—what? The pain itself! This is as bad as saying that to be a cat logically involves being the offspring of cats. It seems therefore, that there must always be a distinction between *being* in a certain mental state and *being aware* that we are in that state.<sup>12</sup>

Armstrong talks here of apprehension and awareness; but it is, as Armstrong himself notes, impossible to apprehend or be aware of a non-existent state, because 'apprehend' and 'aware' are 'success verbs'. So we will stick to our approach to incorrigibility which regards the central issue as whether mistaken *belief* is possible.

The distinct existences argument, as presented above, has two steps. The first is an argument to the conclusion that, for example, my belief that I am in pain is distinct from my pain. The second is that, if belief and pain are distinct, they can occur apart; that is, I may believe that I am in pain when I am not (and conversely, for that matter).

The main problem that faces any attempt to assess these two steps is that Armstrong never really makes clear the sense he is giving 'distinct'. There are two obvious alternatives: on the first, to say, for example, that

<sup>10</sup> See e.g., Armstrong, *op. cit.*; B. Medlin: Critical notice of Ayer's *The Concept of a Person*, *Australasian Journal of Philosophy* 42 (1964) pp. 412-31; J. J. C. Smart: *Philosophy and Scientific Realism* (1963) p. 100; J. L. Mackie: 'Are There Any Incorrigible Empirical Statements?', *Australasian Journal of Philosophy* 41 (1963) pp. 12-28.

<sup>11</sup> 'Is Introspective Knowledge Incorrigible?', *op. cit.*, see pp. 422-3.

<sup>12</sup> *Ibid.*, p. 423.

my pain and my belief that I am in pain are distinct, is to say that they are *not numerically identical*. On the second, to say that my pain and my belief that I am in pain are distinct, is to say that 'I am in pain' and 'I believe that I am in pain' are not *synonymous*.

If we give 'distinct' the first sense (of numerical distinctness) the circularity argument for distinctness is invalid. Circularity as a defect pertains to analysis or to definition, not to numerical identity. Thus the circularity argument could only be relevant to whether 'I am in pain' and 'I believe that I am in pain' are synonymous. Nevertheless, though on the reading of 'distinct' in question the circularity argument is invalid, it has a true conclusion: my belief that I am in pain and my pain *are* numerically distinct. The former is a belief, and so correct or incorrect, but neither throbbing nor in the hand; while, the latter is a sensation, and so may be throbbing and in the hand, but is neither correct nor incorrect.

This, however, is of no help to the distinct existences argument as a whole. If 'distinct' means 'numerically distinct', the second step, from the distinctness of my belief that I am in pain from my pain to the possibility that they occur apart, is invalid. A husband is numerically distinct from his wife, but 'I am a husband' entails 'I have a wife'; an object's colour is distinct from its extension, but '*A* is coloured' entails '*A* is extended'; and so on.

The situation is similar if we give 'distinct' the second sense (of meaning distinctness). On this sense, the circularity argument is again invalid. It would be circular to define 'bachelor' as 'man who is a bachelor', but the two are, nevertheless, synonymous. However, though invalid, the circularity argument again has a true conclusion. 'I am in pain' and 'I believe that I am in pain' are not synonymous. If they were, 'I believe that' would be a vacuous prefix in 'I believe that I am in pain', which it is not: witness the difference between 'I will believe that I am in pain' and 'I believe that I will be in pain', for example.

But, as was the case with the first sense of 'distinct', the truth of this conclusion is no help to the distinct existences argument as a whole. The second step of this argument will now be from 'I believe that I am in pain' and 'I am in pain' are not synonymous, to 'I believe that I am in pain' does not entail, nor is it entailed by, 'I am in pain'; and this step is invalid. '*A* is red' and '*A* is not green' are not synonymous (and, for that matter, the former cannot be analysed as a conjunction containing the latter as a conjunct), but the former entails the latter.

In *A Materialist Theory of the Mind* Armstrong supplements his distinct existences argument (in response to points similar to, but distinct from, those just raised) as follows:

But now let us consider the mechanical analogue of awareness of our own mental states. It is clear here that the operation of scanning and the situation scanned must be 'distinct existences' . . . there must be an absolute distinction between the scanner and the scanned . . . [and] . . . Why should the substitution of spiritual for material sub-

stance abolish the need for a distinction between object and subject?<sup>13</sup>

Here it seems that Armstrong is using 'distinct' to mean numerically distinct, because his argument is designed to show that subject and object cannot be one and the same thing. But this does not seem to advance matters. For, as we have already noted, two things may be both numerically distinct and logically dependent.

### 6. *The empirical argument*

This argument against ICT is that it is possible to describe empirically plausible cases where someone is mistaken about his current mental state. Don Locke puts this argument thus:

. . . surely it is possible to be mistaken about our bodily sensations? Take the initiation trick where the unfortunate subject is blindfolded and told that he will be branded with the red hot poker he can hear, and smell, being heated for the purpose. A piece of ice is pushed into his bare stomach and, naturally enough, he screams. There seems a reasonable case for saying he mistakes a sensation of cold for one of warmth. To be sure he will probably realise the deception almost immediately, but if he dies of fright on the spot it seems he will die thinking that he feels a sensation of intense heat, or pain, when in fact he does not.<sup>14</sup>

The first thing to note about this argument is that it does not follow from the fact that the sensation is produced by an iceblock, that the sensation is one of cold. Cold things (for example, dry ice) can produce burning sensations, and it may be that in the situation Locke describes, the psychological stage setting is such that the ice block does, just for a moment, produce a sensation of heat. Nevertheless, I think it must be admitted that, though this might be what happens, it does not *have* to be what happens. So let us consider Locke's argument on the assumption that the only sensation the subject experiences is one of cold.

My reply to the empirical argument as presented by Locke turns on the invalidity of:

If (*S* believes *P*) & (*P* entails *Q*), then (*S* believes *Q*).

It is obvious, and universally acknowledged, that this principle is invalid.

I believe that, in the case described by Locke, the subject believes propositions which entail that he is experiencing a painful sensation of heat, but does not believe that he is experiencing a painful sensation of heat. (Though he does, of course, believe, just prior to being touched, that he is *about* to have a painful sensation of heat). The subject believes that he is being touched by a red hot poker and that red hot pokers cause painful sensations of intense heat, but, as the above principle is invalid,

<sup>13</sup> *Op. cit.*, p. 106-7.

<sup>14</sup> *Perception and Our Knowledge of the External World, op. cit.*, p. 86,

it does not follow that he believes he is suffering a painful sensation of intense heat. Therefore, an incorrigibilist can admit the essentials of the case described by Locke without abandoning his thesis.

A variant on the empirical argument against incorrigibility has been put forward by Ayer. This variant rests on the claim that empirically plausible cases, not of error, but of *doubt*, can be described. Thus Ayer argues, on the question as to whether one can be mistaken about how something looks to one now:

Suppose that two lines of approximately the same length are drawn so that they both come within my field of vision and I am then asked to say whether either of them looks to me to be the longer, and if so, which. I think I might very well be uncertain how to answer. . . . But if I can be in doubt about this matter of fact, I can presumably come to the wrong decision. I can judge that this line looks to me to be longer than that one, when in fact it does not.<sup>15</sup>

This argument contains a suppressed premise to the effect that one or other of the two lines looks longer to me, or else they look the same length. But this premise simply overlooks the possibility that the two lines neither look equal in length, nor does one look longer than the other; that is, it overlooks the possibility that the appearance presented is *indeterminate*. Therefore, it is open to the incorrigibilist to attribute the doubt in question to the indeterminacy of the experience rather than to the possibility of error.

The familiar speckled hen case may make this reply to Ayer clearer. If I look at a speckled hen, no doubt it will appear to have more than ten speckles; but there will be a number, depending of course on the particular hen, which will be such that I hesitate, indeed am unable, to say whether the hen appears to have more or less than this number of speckles. The obvious explanation for this is not that the hen looks to have a definite number of speckles which I am unable to specify, but that the hen does not look to have a definite number of speckles at all.

A quite different argument from doubt is put forward by J. L. Austin in 'Other Minds':

I'm not sure it *is* the taste of pineapple: isn't there perhaps just *something* about it, a tang, a lack of bite, a cloying sensation, which isn't *quite* right for pineapple? Isn't there perhaps just a peculiar hint of green, which would rule out mauve and would hardly do for heliotrope?<sup>16</sup>

As far as the taste case goes, it seems clear that the possibility of error Austin establishes is not about the taste, but about its *relation* to certain other tastes, namely those usually caused by pineapples. But let us look more carefully at the colour case.

<sup>15</sup> *The Problem of Knowledge* (1956) Ch.2. sec.VI.

<sup>16</sup> The quotation is from the reprinted version in *Logic and Language*, second series, ed. A. G. N. Flew (1961), see p. 138.

Suppose I am having a red after-image. Can I be mistaken about its *precise* shade of red, about whether it is, say, red 6; where red 6 is specified on a chart which divides each colour into barely discriminable numbered shades. I think we must distinguish two questions here: What (precise) colour is my after-image? and what panel on the chart exactly matches the colour of my after-image? Clearly I can be mistaken about the latter, but this does not show that I can be mistaken about the former. The question as to the colour of something is quite distinct from the question as to the colour relationship of that thing to other things.

Perhaps the simplest way of showing this is by noting that it is possible that just *one* thing have a certain colour (the number of things having any given colour is contingent).

And so, being a certain colour cannot be analysed in terms of bearing the relation of colour-identity to other things. It might be argued that this does not rule out the doctrine that our use of colour *words* can be analysed in terms of colour identities between what we are describing and standard reference objects. But this would be to return to the kind of confusion discussed in §3, between believing that my after-image is red, and believing that 'red' applies to my after-image.

### 7. *The argument from science*

It is commonly suggested in discussion that ICT is somehow incompatible with a proper respect for the possibility of scientific investigation of the mind (or brain). For example, it is often pointed out that neurophysiologists might develop a well-founded theory correlating brain states and beliefs, and brain states and sensations, such that we could confidently assert that subject  $S$  has brain state  $B_1$  if and only if  $S$  believes that he is in pain, and that  $S$  has brain state  $B_2$  if and only if  $S$  is in pain. It is then pointed out that it is at least logically possible that  $B_1$  occurs without  $B_2$  occurring, and so, that it is at least logically possible that we have strong evidence for  $S$  believing that he is in pain, and strong evidence against  $S$  being in pain.

All this is true, but no objection to ICT. The logical possibility of having strong evidence for  $P$ , and against  $Q$ , is not inconsistent with  $P$  entailing  $Q$ . Suppose we have established over the years that a brilliant, but very cautious mathematician only asserts that a thesis is a theorem if it is, and that a thesis is not a theorem if it is not. Further, suppose that  $T_1$  and  $T_2$  are theorems such that  $T_1$  entails  $T_2$ . Clearly, despite all his brilliance and caution, our mathematician might (logically) assert that  $T_1$  is a theorem, and that  $T_2$  is not; then, as theses in mathematics which are not theorems are (necessarily) false, we would have strong evidence for  $T_1$ , and against  $T_2$ , despite the fact that  $T_1$  entails  $T_2$ .

Another argument sometimes offered is that ICT would make scientifically significant laws (as opposed to 'accidental' generalisations) relating

sensations to bodily states impossible.<sup>17</sup> For the incorrigibilist must concede that any such law could be overthrown by (honest) testimony at any time. No matter how well confirmed a correlation between being in pain and being in brain state *B*, say, had been established, and no matter how well this correlation fitted into a comprehensive and successful scientific theory, the incorrigibilist would always have to allow a subject's honest testimony that he is not in pain (when he is in state *B*) to be decisive, and so, to establish that the correlation has broken down.

The reply to this argument is just that any empirical generalisation or law can be overthrown by *something*, so why not by honest testimony. Of course, if the correlation between brain state *B* and pain is well confirmed, the incorrigibilist will have to hold that the probability of getting conflicting honest testimony is very low. But this is not a difficulty in itself.

Arguments other than those discussed above have been offered against ICT. I hope, however, enough has been said to place the onus of proof on those who believe there is a decisive objection to the incorrigibility thesis.

*La Trobe University*

Received November 1972

<sup>17</sup> I think this is the argument being offered by Kathryn Pyne Parsons in 'Mistaking Sensations', *Philosophical Review* LXXIX (1970), see pp. 201-2.