Personal identity and mental content

JAMES BAILLIE

ABSTRACT In this paper, I attempt to map out the 'logical geography' of the territory in which issues of mental content and of personal identity meet. In particular, I investigate the possibility of combining a psychological criterion of personal identity with an externalist theory of content. I argue that this can be done, but only by accepting an assumption that has been widely accepted but barely argued for, namely that when someone switches linguistic communities, the contents of their thoughts do not change immediately, but only after the person becomes integrated within the new linguistic community. I also suggest that recent work on personal identity, notably by Derek Parfit, has tacitly assumed internalism regarding mental content. I do not intend to argue for either externalism or a psychological criterion. My aim is merely to explicate the issues involved in making them compatible.

In this paper, I will attempt to map out the 'logical geography' (to use Ryle's phrase) of the territory in which issues of personal identity confront matters of mental content. In particular, I will investigate the possibility of combining a psychological criterion of personal identity with an externalist theory of content. I will argue that this can be done, but only by accepting an assumption that has been widely accepted but barely argued for, namely that when someone switches linguistic communities, the contents of his or her thoughts do not change immediately, but only after the person becomes integrated within the new linguistic community.

I

Suppose that I enter a teletransporter, equipped with a new scanner (Parfit, 1984, p. 199) that can produce a molecule-for-molecule duplicate of me without destroying my body. Thus, one person enters, and two emerge. In what follows, I will refer to the person who steps into the teletransporter as 'J'. In order to avoid begging any questions, I will refer to the emerging person that is physically continuous with J as 'O', and the newly created person as 'D'. To be precise, J, O, and D should be considered not as persons, but as temporal 'person-stages', i.e., temporal slices of some continual person. In the following pages, all mention of J, O, or D should be taken as referring to person-stages, rather than to persons per se. Such a case of 'branching' forces the question as to which, if either, of O and D is copersonal.

James Baillie, Department of Philosophy, University of Portland, 5000 N. Willamette Boulevard, Portland, OR 97203-5798, USA. Email: baillie@uofport.edu.

0951-5089/97/030323-11 © 1997 Carfax Publishing Ltd
with J, i.e., which constitutes a later stage of the same temporally continuous person as J? When restated in terms of persons, the question is "Do I survive?—and, if so, where am I?" So, when I ask whether J is copersonal with O, I want to know whether O is me.

One’s answer to these questions will depend on one’s criterion of personal identity. As is well known, a major disagreement here is between those who see identity as reducible to, or at least supervenient upon (a) physical continuity or on (b) psychological continuity. These positions are usually referred to as the physical and psychological criteria, respectively. The idea behind the physical criterion is that x (at time t₁) is copersonal with y (at t₂) if and only if (1) y constitutes a later stage of the same physically continuous living body as x, and (2) there is no other z (at t₃) who is likewise related to x. Most modern variants modify this to the claim that one physically continuous functioning brain is what determines identity. So I go where my brain goes. The psychological criterion takes identity to rest on psychological continuity. Thus, x is copersonal with y if and only if (1) x and y are related by overlapping chains of psychological connections, and (2) no other z exists who has equal or stronger psychological continuity with x.

In the most influential formulation, by Derek Parfit [1], this amounts to the claim that for any two contiguous stages, a day apart, at least half of one’s direct psychological connections hold until the following day [2]. (The paradigm case of a direct psychological connection is that between an experience and its recall.) Parfit acknowledges that this criterion cannot be taken as a strict rule, and that our notion of personhood affords a certain degree of vagueness. It would clearly be artificial and unbelievable to grant that someone survived if 50.001% of his psychological connections held until the following day, but not if only 49.999% had done so. To take the criterion as a strict rule would be to admit some possible case like this, in which one’s personal survival hung on the retention of a single memory.

But back to the story. Is J copersonal with O or with D? A physical criterion would obviously select O, since only he is physically continuous with J. I will now set this option aside. Since the physical criterion is defined without reference to mental states (or, more precisely, to mental states qua mental states), it has no implications regarding theories of mental content. I will also set aside any theory in which sameness of personhood is taken to consist in the continued existence of some nonmaterial substance. Apart from the sheer obscurity of such a view, it has, again, no implications regarding mental content. The rest of this paper will focus on the psychological criterion, where the most interesting connections are to be found. Parfit distinguishes narrow and wide versions of this criterion. A narrow criterion requires that psychological continuity be realized by its ‘normal’ cause, namely the continuous functioning of a single brain. By contrast, a wide criterion admits any cause. A narrow psychological criterion is thus extensionally equivalent to a physical criterion, and would therefore judge that J is copersonal with O. By what I might call a pure wide psychological criterion (i.e., by which only psychological continuity determines identity) we would have a ‘tie’, since O and D, ex hypothesi, have exactly similar psychologies. That is, not only is ‘what it is like’ to be O or D exactly similar, but D inherits J’s contentful mental states, since D was copied from him. Thus, D
is different from some Davidsonian swampman who might emerge by accident, causally unconnected to J. But such ‘purity’ is implausible. It is clearly more reasonable to hold that while psychological continuity is the overriding determinant of identity, it is not the only one. Thus, in such a case, our intuitions would surely be that J was copersonal with O, since O’s physical continuity makes him J’s ‘closest continuer’ given the ‘tie’ between J’s psychological connectedness to O and to D.

II

The contemporary debate over mental content has never, to my knowledge, been discussed in the personal identity literature, even in the most recent work. It is therefore impossible to irrefutably prove that particular theorists of personal identity were internalists or externalists about content. However, I believe it to be a plausible inference that this absence of discussion indicates a lack of recognition that recent externalist arguments have any implications for personal identity. In other words, discussion of personal identity has gone on under a tacitly internalist assumption. Present debates about the relative merits of ‘reductionist’ theories like the psychological and physical criteria get their agenda from John Locke, who was working from a Cartesian theory of mind. It has become increasingly recognized in recent years that the deepest root of this Cartesian conception lies not in substance dualism, but in what McCullough (1995) has recently called the ‘self-containedness’ of the mind, that is, that the nature of minds and the individuation conditions of their contentful states, are fixed wholly independently of any facts concerning physical objects or properties. This self-containedness thesis is precisely the picture of mind challenged by externalism. My claim that present work on personal identity takes place within a (tacitly) internalist model of mind relies on the following, very plausible, premise: to enter into a debate, the very nature of which was defined by an internalist theory of mind, and to not even acknowledge, let alone critique, that underlying conception is to be oneself working under these same internalist assumptions.

For example, the ubiquitous teletransporter thought-experiments are employed to present us with cases of psychological continuity in the absence of physical continuity, in which the newly created psychological duplicate may even reside on a different planet. In holding that someone and her duplicate are psychologically continuous, such thought experiments clearly assume that this continuity holds wherever the duplicate happens to be. In other words, they assume that one’s identity is not affected by such extrinsic factors as one’s location or features of one’s physical or social environment. I have devised an example which will make all this clear.

J steps into the teletransporter, as before. However, something goes wrong. There is an explosion at the base of the machine, which projects the traveler at great speed down into the bowels of the earth. There O discovers a whole society of people who have been living there for eons, without any contact with the earth’s surface. This society (which I’ll call ‘Middle Earth’) looks just like any regular town, but has innumerable as-yet-undetected microstructural differences to any ‘surface-
town" regarding its natural kinds. That is, it is crammed with what we might call 'Putnam cases' [3] (see Putnam, 1975). Furthermore, the inhabitants of this Middle Earth speak a language phonetically and syntactically identical to English (or, to be more realistic, let us assume that the inhabitants of O’s new town speak a multitude of languages, each one a phonetic and syntactic copy of those of a surface language). However, their words are used according to a subtly different set of social conventions, such that a large number of their terms will have different contents than their counterpart expressions on the earth’s surface. For example, the Middle Earth word 'arthritis' will refer to any rheumatic or inflammatory condition of the joints or limbs. I assume that the general scenario is becoming clear: this place is host to a multitude of what I will informally label 'Burge cases' [4] (see Burge, 1979) as well as 'Putnam cases'. As for the newly created D, he walks out of the teletransporter back in the original room, dazed but unhurt.

What should we say about this case? By the psychological criterion, as its adherents normally apply it (i.e., without recognizing the relevance of externalism), J is unproblematically copersonal with O. It doesn’t matter where O is, as long as he remembers enough of what happened before [5]. However, if you take an externalist view of mental content, it might seem that a massive psychological discontinuity has taken place between J and O as a result of the trip. To adapt familiar examples: if a competent Middle English speaker asks for a glass of water, the satisfaction-condition of her request will involve XYZ, not H₂O; such a person can have a true belief that she would describe as being that her arthritis has spread to her thigh (although this belief will not include the concept ‘arthritis’, but a different one that we can call ‘arthritis’—applying to all rheumatic conditions of the joints and limbs). Bear in mind that my thought-experiment is stipulating that Middle Earth has a vast number of such cases—so it is not like Putnam’s Twin Earth in this regard. So we can say that if O is speaking and thinking in this Middle English language, then the contents of a vast number of his thoughts are different to those that J would have had in phenomenally identical situations here on the earth’s surface before the accident. It would then follow that if most of their thoughts differ in content, then J would not be copersonal with O, because they would not be psychologically continuous. Of course, even if this line of thought were correct, the subjective view, the 'what it is like' to be J, would be unaffected by these changes in content. Thus, it might seem, in the same way that externalism raises serious epistemological problems concerning first-person authority regarding what one is thinking, parallel problems occur over knowledge of who one is.

III

An externalist would probably reply that when one changes environments, one’s mental contents do not all change instantaneously and completely, right there and then. Rather, this 'semantic changeover' would only occur gradually, after one had been within that linguistic community for some time. This is Burge’s view: “The thoughts would not switch as one is switched from one actual situation to another twin actual situation. The thoughts would switch only if one remained long enough
in the other situation to establish environmental relations necessary for new thoughts" (Burge, 1988, p. 652). Thus, on my descent into Middle Earth, my mental contents would continue to conform to the norms of my linguistic community on the Earth surface. Were my residency to be long-term, I would presumably go through a 'gray' stage in which there is no fact of the matter about the content of my thoughts. After this has been worked through, I would have succeeded in having 'gone native' as a Middle Earth English speaker and thinker. When this variant of externalism is adopted, then, so long as the learning period governing the semantic changeover is gradual enough to ensure that psychological continuity is maintained, then problems of personal identity would be averted. I will return to this theory, which I will call the slow-switching theory, in a moment.

I suggested above that if you take a Parfitian wide psychological criterion, and hold that the change in the mental content occurs either immediately on arrival, or at least very quickly (I'll call this the 'quick-switching' theory), then psychological continuity is destroyed, and J is not copersonal with O. However, suppose that someone objected that no matter what happens to O in Middle Earth, psychological continuity with J is ensured by O's memories of J's experiences back on the earth's surface.

But is this an option that an externalist can take? As Ludlow (1995) has recently argued, such a move would be to assume that the semantic content of a memory is fixed, once and for all, by the physical and/or social environment in which the remembered event took place. It would be to say that if O, now a competent Middle English speaker, remembers J's nearly drowning as a child back on Earth, then he is thinking about H₂O, not XYZ. Likewise, if J had suffered from arthritis in the past, then O's recollection of the pain would concern arthritis, not 'arthritis'. But, as Ludlow convincingly points out, an externalist ought to say that a memory's content is determined by environmental factors pertaining at the time of recollection, not those at the time of the original experience. "After all, what would it mean to say that the contents of our memories are fixed by our social environment if in fact those contents, once fixed, are totally inert to all environmental changes?" (Ludlow, 1995, p. 158).

If I am correct in ascribing a tacit internalism to Parfit, then he is assuming, in his teletransporter thought-experiments and elsewhere, that in remembering what I did last week, the content of my memory state is the same as that of the original experience. For example, if I believe that p at some time t₁, and recall that act of believing at t₂, the original belief and its recollection have a common content. It follows, by contraposition, that if a particular mental state and a purported recollection of it do not share a common content, then this cannot be a genuine case of memory. So my conclusion would seem to hold, namely that acceptance of a Parfitian wide psychological criterion, plus a quick-switching model of content change, leads to the conclusion that psychological continuity is destroyed, and J is not copersonal with O.

By contrast, if you take this wide psychological criterion while following Burge in maintaining that one's mental contents would change only very gradually (the slow-switching theory), then psychological continuity might hold. It would depend
on the length of the learning period one undergoes in acclimatizing to the new linguistic community. But how long is that? Are some concepts acquired quicker than others? Is the quick-switching model appropriate for any concepts? If so, how many, and which ones? As Ludlow points out, the slow-switching theory does not take our thought contents to be totally inert to environmental changes. However, they are not immediately responsive either. Rather, the theory suggests that one’s past membership in a particular linguistic community imparts some psychological power, some form of *semantic impetus*, to the mind, allowing one to withstand the absence of the environment itself, and permit words to continue to keep their old meanings, at least for a time. The nature of this process is not obvious, at least not to me. Burge himself has offered no detailed accounts of the process. This is understandable, since problems of personal identity are not his immediate concern. Still, if we are to have a definitive answer to whether the psychological criterion is compatible with a slow-switching model of externalism, then we need a more detailed account of precisely what is required to establish the relations necessary for new thoughts. I see this as the most basic issue to be resolved, and regret that I can merely identify the problem at present, not solve it.

However, here are a few methodological suggestions towards its solution. Take some criterion of personal identity C, and some particular case P, where C determines that identity fails; to hold in P. Here we have two options open to us. Firstly, we can argue that since P fails C, identity does not hold in this case. Secondly, we can argue that since C yields such a result, then C is an inadequate criterion of identity. However, there is no formal procedure to tell you which option to take in any given case. Clearly, we cannot take all our initial beliefs about identity as sacrosanct, since they may be based on factual error, or on the influence of an untenable metaphysical theory, and so on. On the other hand, it is equally obvious that there is no *a priori* way of determining the truth of any criterion, in advance of its application to any particular case.

This is particularly important in thought-experimental situations far from the everyday world in which the concepts are at home. It would be preposterous to think that our common usage of concepts of personhood determined, in advance, whether or not personal survival took place in all logically possible situations. In these cases, employed as they are to chart the outer limits of the application of concepts of personal identity, our judgments are as much a matter of decision as of discovery. Similar remarks apply equally to matters of content.

The only reliable method of coming to a plausible criterion of personal identity is to engage in a process of *reflective equilibrium*, where, after some mutual adjustments, we might come up with a criterion that accords with our considered judgments about particular cases—i.e., not the opinions with which we entered the discussion, but those that we are left with after philosophical reflection. It would follow from such a method that if any purported criterion clashed with significant numbers of such considered judgments, then we would have a strong reason to reject that criterion. It would also follow that if any theory concerning some other domain implied such strongly counterintuitive results about personal
survival, then we have, again, prima-facie evidence against this other theory. In order to overturn this judgment, such a theory would need strong independent support.

As regards our central question of whether some form of the psychological criterion of personal identity is compatible with some type of externalist theory of mental content, much will depend on where one stands on the issue of holism versus atomism regarding the individuation of thoughts. There are, of course, many intermediate positions corresponding to different degrees of semantic localism. I choose to describe the geography of the issues in terms of the limiting cases of atomism and holism purely for expository reasons. Let us start by considering a fully fledged atomism, where each concept is grasped singly and serially. Each would have its own ‘learning period’, taking place either with direct contact with standard cases (Putnam cases), or in social interactions with competent speakers using that concept (Burge cases). For example, I would learn the correct usage of ‘sofa’ by being exposed to usage of this term by competent Middle English speakers. However, my grasping the content of ‘sofa’ would have no effect on my grasp of ‘arthritis’, ‘football’, or any other Middle Earth concept. On this atomistic picture, the learning period required for acquiring the new language, and new thoughts would be very scattered and diffuse, in fact probably never ending, and would certainly ensure psychological continuity.

The position of O in Middle Earth would then be strongly analogous to Parfit’s ‘psychological spectrum’ thought-experiment (cf. Parfit, 1984, p. 231) where one has a few memories erased every day, each one being replaced by a qualitatively different quasi-memory that originated in an experience of someone else. Even if repetitions of this procedure finally led to a point in which no original memories remained, psychological continuity would be maintained, since extremely strong connectedness would hold between the sets of mental states present on any two contiguous days.

Let us now consider a full-blooded holism, in which the meaning of every word, and its corresponding mental content, is determined by the meanings of every other word or concept. It would follow that if any one of your concepts (from Earth) is replaced by a Middle Earth correlate—e.g., if you acquire the concept ‘XYZ’ in place of ‘H₂O’, or ‘marthritis’ rather than ‘arthritis’—then a chain reaction is set off in which all your concepts are transformed. Now in this case, assuming that this process is fairly dynamic, then psychological continuity is breached.

So a wide psychological criterion plus a slow-switching version of externalism might preserve psychological continuity; by contrast, when a wide psychological criterion is combined with a quick-switching model of externalism, the learning period is so rapid (i.e., not much longer than the time it takes to acquire a single concept), that the conceptual rate of change is far too severe to support psychological continuity. As I mentioned above, if one is uncomfortable with a strict dichotomy between the quick- and slow-switching models (or between atomism and holism), and prefer to see an entire spectrum of intermediate positions, then my conclusion can be easily restated: the closer you are to the slow-switching (or atomist) limit, the more can psychological continuity be ensured.

Let me now briefly turn to another set of reflections. I commented above that
there was a close parallel between the slow-switching process and Parfit’s ‘psychological spectrum’ thought-experiment. I now want to bring out another close parallel with a well-known thought-experimental device from the personal identity literature, namely the ‘future pain test’, as most famously depicted by Williams (1970). So assume a psychological criterion, plus a quick-switching externalism. Now take J, here on the earth’s surface. Suppose that he is told that he is to be sent to Middle Earth. Suppose J to be a philosopher, who grasps the metaphysical issues involved. Thus, J believes that his thought contents will change on arrival in Middle Earth. It would follow from our assumptions that this Middle English speaker (i.e., O) is not identical with J. It would then follow that J ought not to have any self-interested concern for that person. For example, were J to be told that O was to be tortured, then J might feel sympathy, but he ought not to feel overwhelming fear and dread. However, since J most definitely would have this latter response, then surely we have a reductio of the conjunction of our assumptions. In other words, one cannot, save at a cost of a radical transformation of one’s views on personal identity, hold both a psychological criterion and a quick-switching theory of mental content.

IV

In this section, I will illustrate a couple more cases in which considerations from the externalist literature have interesting consequences for personal identity. So let us return to J, O, and D. Let us grant a psychological criterion accompanied by a quick-switching model of conceptual change. When we consider O’s relationship to J, it seems that we have a variant of a Lockean case in which J is the same man as O (that is, they are stages of the same animal organism) but is not copersonal with O. Thus, if we accept Wiggins’ (1980) thesis of the sortal dependency of identity, whereby every identity statement must be covered by some sortal concept (that is, if x is identical to y, they must be identical qua some sortal concept: so x is the same F as y), then we would have a case of identity being ‘sortal-relative’: J is the same man as O (i.e., they are identified under the sortal concept ‘man’) but not the same person [6]. When we are talking strictly about personal identity, a peculiar and strongly counterintuitive result emerges, namely that J survives as D!

The strangeness of this result can be brought out by considering a variant on this thought-experiment, in which D is not produced. In fact, teletransportation is a dispensable factor as well. Just suppose J were to somehow fall into Middle Earth. Since, on our assumptions stated above, the truth and satisfaction conditions of the majority of the ensuing person’s thoughts would change, massive and sudden psychological discontinuity would take place, and this would forbid a psychological criterion from saying that J survives as O. Thus, in such a case, when we consider personal identity (rather than identity as a human animal), we must conclude that J has no future copersonal stages. In other words (switching to the vocabulary of persons), I do not survive. By contrast, I would have survived had O relocated to any environment that afforded continuity of psychological content between J and O. As before, the most compelling response to such a thought-experiment is to say that such a result merely constitutes a reductio ad absurdum, either of a quick-
switching externalism or of a wide psychological criterion. However, let us push the present line of inquiry further.

Grant, with the previous paragraph, that J does not survive if O goes to Middle Earth. Recall that we are assuming a psychological criterion with a quick-switching model of conceptual change. Now suppose that O escapes from Middle Earth and makes it back to the Earth’s surface. Since we have granted that J is not copersonal with O, it would follow, for the same reason, that O is not copersonal with the physically continuous person-stage who returns (call him ‘R’). It would follow, by transitivity, that J is not copersonal with R either. For what it is worth, this is intuitively indefensible. Let us consider what this conclusion’s implications would be for the attribution of responsibility, and therefore of praise and blame. Suppose that J had committed a murder just prior to his fall into Middle Earth; suppose that O lies low down there before risking a return to the surface, but R is recaptured on arrival. Surely even Johnny Cochran would have too much shame to attempt his client’s acquittal on the grounds of nonidentity!

V

On the other hand, taking the line that J was copersonal with R would force us to grant some form of discontinuous existence, (at least qua person, if not qua man), and drop the claim, central to the psychological criterion that we have been assuming, that personal identity requires unbroken psychological continuity. Either way, the orthodox construal of the psychological criterion comes under stress, and some part of it has to give.

To sum up: I suggested above that discussions of personal identity have taken place without a recognition of the relevance of recent debates about mental content. If I am correct, then now that this fact has been made explicit, someone who supports an externalist theory of mental content faces the task of accommodating it within a credible criterion of personal identity. This task may require a process of mutual adjustment between the theoretical domains. The desired result of such a process would be a combined picture of personal identity and mental content which accords with our considered and theoretically informed judgments regarding a wide range of individual cases, both real and hypothetical.

So a criterion of personal identity is obliged to provide answers which are in line with our considered judgments about standard cases. This requirement is compatible with there being cases (usually hypothetical) in which there is no fact of the matter regarding personal survival. We need to acknowledge some vagueness, but not so much that it makes the criterion empty and meaningless. We want to say that a credible psychological criterion requires enough psychological connectedness to sustain personhood, bearing in mind that we cannot specify precisely how much that is—any more than we can say precisely how much hair one can lose without becoming bald. A criterion of personal identity can be, at best, a rule of thumb. Likewise, we want an account of the rate of conceptual change which equally accords with our considered judgments on these matters, again allowing for some
degree of vagueness. This latter account awaits development. However, one conclusion can be drawn while we wait.

I have stated that if a philosophical theory produces results that are in contradiction with our most well-considered judgments about personal identity, then there is a prima-facie reason to reject the theory. It should now be clear that for those who accept both a psychological criterion of personal identity, and an externalist theory of mental content, we have an argument in favor of the slow-switching theory of semantic change: namely that its competitor, the quick-switching theory, leads to preposterous results that our considered judgments about personal identity cannot tolerate. I have been gathering such results throughout this article, and surely have not exhausted the list.

Acknowledgments

I am grateful to Norah Martin for much helpful discussion, and to two anonymous referees for suggestions that greatly improved this paper.

Notes

[1] Parfit is only describing the psychological criterion. As is well known, he rejects any criterion of personal identity, and advocates the replacement of the identity relation with ‘relation R’, i.e., psychological continuity without identity.

[2] It has been suggested to me (by an anonymous referee) that personal survival requires only moment-to-moment psychological continuity, and that the rate of change of mental contents is irrelevant. So, suppose that someone were to live a lifetime of experience in a week. Call him ‘the fast-forward kid’, ‘FF’ for short. FF’s lifetime is so accelerated that a decade’s worth of factual learning, emotional maturation, etc., takes place in a matter of days, so that he lives what to us would be hundreds of lifetimes of experience. So, since the minute-by-minute stages were sufficiently similar (i.e., sharing the majority of their contents), personal survival would be guaranteed despite no memory of Monday remaining by Thursday. I disagree.

Personhood surely requires that what Dennett (1978) calls ‘the intentional stance’ can be successfully applied to you, i.e., others can generally explain and predict your behavior successfully, via the positing of beliefs, desires, etc. To be seen as a rational agent, these mental states must be regarded as a fairly well-integrated set, so that your conduct today makes sense in terms of your behavior of the week or month before. This explanatory/predictive enterprise would fail in FF’s case.

The normative dimension of personhood includes rationality as well as morality. We ascribe personhood to those who can reliably interact with others. But how much weight could we place on FF’s promises on Monday if his entire worldview could be utterly transformed by Thursday? What sense could we make of his behavior on Thursday if we had no news of him since Monday? How much could we predict? Given that FF’s mind on Monday would be as qualitatively different from his mind on Thursday as to that of someone else then, it would be no guide to his views or actions on Thursday. This breakdown of the intentional stance not only threatens to undermine FF’s personal identity, but his personhood itself.

However, in a world where everyone had these highly accelerated mental processes, FF would be regarded as a paradigm case of personhood and of personal survival. Of course, were you or I to visit, we would be unable to see them as persons.

[3] For example, the metal employed for currency and decoration that the Middle Earthers call ‘gold’ is in fact iron pyrite. Likewise, the clear, odorless liquid flowing from their taps that they call ‘water’ is in fact some sensually indistinguishable but chemically different substance XYZ.
These cases are very common—we don’t need to resort to science fiction. For example, take the term ‘professor’. In the USA, the extension of this expression is the set of all university teachers. In the UK, the term is restricted to those of the highest academic rank (i.e., full professors). Given Burge’s plausible assumption that an otherwise competent user of a language can nevertheless have an incomplete grasp of a concept, it is possible for such errors to go undetected by the user or by others, whether the user is a recent immigrant or even a native.

The same conclusion would be drawn by supporters of narrow content. They would grant that the ‘wide’ contents of one’s thoughts would change, but insist that the type of content relevant to psychology in general, and a fortiori to personal identity, is an internalistic narrow content that would survive one’s relocation in Middle Earth.

See Bailie (1990) for further reflections on how teletransporter cases lead to the sortal-relativity of identity.

References


