

## His Burning Pants

THOMAS MCKAY\*

*I* John Perry has helped us to see more clearly the special nature of some beliefs about ourselves.<sup>1</sup> When I look over and see a man with his pants on fire, I am alarmed. I shout, “Drop and roll!” But my alarm becomes panic, shock, or dropping and rolling when I realize that I am looking in a mirror.

When I first looked in the mirror, I formed a belief about someone—in fact, about me. But I acted as I might toward almost anyone but me. My subsequent discovery of the mirror led to a belief about myself that motivated a particular kind of action—action that is appropriate for me to preserve myself rather than action of the sort that I might do in the attempt to help someone else.

We shall call the second more intimate type of belief about myself “*self-attributive*” belief. This is not a particularly good name, because before my big realization I saw myself in the mirror and attributed the property of *having burning pants* to me. I just didn’t “realize” that it was me. Getting clearer about the nature of this *realization* will be the point of this paper. But meanwhile we need a name for the special beliefs about the self that result from such realizations, and although “self-attributive” is not very good, neither is anything else.

Perry’s discussion has made it clear that no purely descriptive belief, of the sort

(B) The man with property *F* has burning pants,

can be identified with self-attributive belief. Unless *F* is filled out by expressions involving “I”, “here”, or “now”, there is always the possibility that one can accept (B) and yet not accept

(C) I have burning pants

(at least for a short time). My self-attributive belief is not reducible to any descriptively characterizable belief lacking “I”, “here”, or “now”.

---

\*The writing of this paper has been supported by a grant from the Syracuse University Senate Research Fund.

In [1] Barwise and Perry have developed a semantical apparatus that they use for characterizing belief-states (pp. 83–85). In particular, they define an individual role  $i$ , intended to be used in characterizing self-attributive beliefs (pp. 85–87, 259–260).

In this paper I will go a few steps beyond what is said in [1] to explore what is involved in self-attributive belief. It seems that Barwise and Perry have the material necessary for giving a sound and systematic account of self-attributive belief, but they never bring it together to do the job.

2 Many aspects of the phenomenon of self-attributive belief also occur when there is no self-reference.<sup>2</sup> Barwise and Perry give a fundamentally sound and simple account here.

Consider the familiar case of Pierre, who has beliefs about a place he calls “Londres” and about a place he calls “London”. In particular, he sincerely asserts

- (1) Londres is pretty.
- (2) London is not pretty.

But since his uses of “Londres” and “London” refer to the same city, it appears that he has contradictory beliefs about that city. Yet (the story goes on) he has made no logical or conceptual error in having these contradictory beliefs.

Barwise and Perry have introduced the systematics for representing Pierre’s state of mind and saying clearly what goes wrong and what is all right in Pierre’s belief.

The relation  $B_r$  relates a believer to a belief-schema. A belief-schema (an event-type or “disjunction” of event-types) is a construct of *individual roles* and properties. Thus Pierre’s belief-state can be characterized in the following way ([1], p. 252):

in  $e_0$  at  $l$ :  $B_r$ , Pierre,  $E_1(a)$ ; yes  
                    $B_r$ , Pierre,  $E_2(b)$ ; yes  
                   of,  $a$ , London; yes  
                   of,  $b$ , London, yes

where

in  $E_1$ : at  $l_1$ : pretty,  $a$ ; yes  
 in  $E_2$ : at  $l_2$ : pretty,  $b$ ; no.

Since event-types  $E_1$  and  $E_2$  are jointly possible, this is a *cognitively coherent* situation. Pierre could have been so situated that events of types  $E_1$  and  $E_2$  both occurred.

But the relation  $B_r$  and the event-types it relates Pierre to are only the internal (cognitive) part of the belief-story. There is also an *external* component. For the belief to be correct, there must be an anchoring for the roles that makes it true that events of types  $E_1$  and  $E_2$  occur. But, in the case of Pierre’s beliefs, the two roles  $a$  and  $b$  are anchored in the same individual, London. Events of types  $E_1$  and  $E_2$  *with that anchoring* cannot co-occur, thus the beliefs are externally incoherent even though they are cognitively coherent. The *doxastic conditions* (the  $B_r$  relations and associated event-types) are coherent, but they are

not realizable in that particular *setting* (i.e., with those *anchorings* or *of-relations*).

On the basis of different encounters (or different sets of encounters) with an individual, we might develop two different concepts, two different sets of beliefs, about that individual. If we come to realize that these concepts are concepts of the same individual they become *merged*, usually giving us a richer stock of beliefs about it (as long as consistency considerations do not force too much weeding out). This merging is the discovery that two individual roles (*a* and *b*) are anchored in the same individual.

This idea has sometimes been introduced under a rubric like Frege's, saying that there are two modes of presentation that are ways of presenting a single individual.<sup>3</sup> While that is on the right track, it is not quite correct, and the Pierre case is a good one for illustrating this. There is not (or at least need not be) a simple *mode of presentation* associated with Pierre's "London"-thoughts. There is rather a network of associated thoughts, beliefs, and attitudes that are cognitively linked. The appropriate sort of linkage is what makes it true that Pierre has unified them as representations of a single place. (Pierre's "Londres"-thoughts also stand in a network of associations, but they are not unified with Pierre's "London"-thoughts.) There need be no privileged "mode of presentation" associated with such a network, if a mode of presentation is a picture or property (or small set of them) that I can contemplate at a single time. Pierre might have a "London"-thought centered around visits to Madame Tussaud's, or around long waits at Victoria Station, or around bitters and beef-and-kidney pie. There need be no common descriptive element that constitutes a single "mode of presentation" linking these thoughts. Rather, it is their being parts of the same network (their being readily linkable by Pierre) that constitutes their being "London"-thoughts. The individual roles that play a part in belief are best handled as rich networks, not necessarily giving any privileged position to a particular propositional node of the network. On any particular occasion, I think of London in one way or another. All of the ways in which I can think of London are parts of a network of representations involving the individual role *London*, though at any particular moment some nodes of the network may be more available than others.

This is a step beyond Barwise and Perry's explicit treatment, for they define each role by a single event-type. To adopt my approach we must see this as being merely one node of a network associated with (constituting?) an individual role.

Roles are anchored by causal relationships to events of the type in the cognitive network. Roughly, a role *a* is initially anchored in a particular object *o* if the agent's perceptual state  $E(a)$ , initiating that role, carries information about an event of type  $E(o)$ . This initial anchoring event of type  $E$  does not define the role *a*. The role *a* might cognitively persist even after the initial anchoring events no longer play a part in the representational network (having been long forgotten). Such anchoring events can initiate a cognitive role without defining it.

Barwise and Perry speak of *recognizing* the subject of a belief "through a mode of recognition contained in the belief" ([1], p. 250). They also presuppose that a concept can contain a mode of recognition ([1], p. 250). Their view about the nature of these containments is unclear, but even as far as it goes *contain-*

*ment* seems like an inappropriate metaphor. The Barwise and Perry concept of an individual role is better seen as a marker of the fact that the doxastic state characterized is linked to a particular network. Being linked to that network is what constitutes a belief's being properly characterized as involving that role. A property is a mode of recognition of an individual by being linked to such a network. The metaphor of containment can only conceal the mutuality of these relationships among roles, beliefs, and modes of recognition.

Cases (such as Pierre's) of unmerged roles anchored in a single individual make the account of actions more complicated than it might otherwise be. One might ordinarily be inclined to generalizations something like this one connecting belief and action:

If  
 (i) individual  $S$  wants to  $V$  some  $F$ ,  
 (ii)  $S$  believes  $x$  to be  $F$ ,  
 then  
 $\bar{S}$  will try to  $V x$  (if  $S$  has nothing better to do).<sup>4</sup>

But clearly this will not work. If Pierre is near London and he wants to be in a pretty city, this will not motivate him to act even though he believes it to be pretty (etc.). Pierre must connect his ability to act (his proximity to the city) with his belief that the city is pretty. But if his ability is recognized by him only in connection with his concept *London* and his belief that the city is pretty is associated only with his concept *Londres*, then Pierre will fail to act. This failure occurs even though Pierre believes it to be pretty and also believes of it that he is near.

If Pierre merges his concepts he will not have these problems. But merging is more than we really need. We need for him to *apply* the relevant concept to the action at hand ([1], pp. 248, 251). If his motivating desire is to be in a pretty city, then he must: (i) have a belief of some city that it is pretty, and (ii) he must be able to *apply* the concept that plays a role in the belief to the individual with respect to which he has the desire. Some single concept must play a role in both his belief and his desire.<sup>5</sup> This is less than merger. (For example, Pierre might apply his *Londres* concept if he has a French-language map of England and is near London, even if he fails to merge the concepts *London* and *Londres*.)

3 In applying this to the first-person case, we can say, as a first approximation, that the same ideas should work. When John sees the man with his pants on fire, he first has a belief the character of which is formulated by the event-type

in  $E$ : at  $l$ : pants on fire,  $a$ ; yes.

When he "finds out that it is his own pants", he applies the belief. In this case, he merges  $a$  with his "self-concept"  $i$ . Having done so, he *applies*  $E$  by connecting it with a desiderative state  $E'$

in  $E'$ : at  $l$ : avoids burning,  $i$ ; yes

and a doxastic state

in  $E''$ : at  $l$ : best avoids burning (when pants are on fire) by dropping and rolling,  $i$ ; yes.

This way of dealing with self-attribution has one immediate and obvious virtue. The fact that the individual role *i* figures in a belief does not require that there be some privileged *mode of presentation* associated with *i* that is a component of the belief representation. Like the individual roles *London* and *Londres*, the role *i* can be associated with a network of representations, requiring no privileged mode of presentation for any particular thought involving that role.<sup>6</sup>

But we should still be a bit puzzled about the cognitive role *i*. What makes beliefs involving *i* especially motivating? Note that the answer just given at the opening of this section will not really do. Even if John had not merged *a* with *i* in this case, he might have had a desiderative state  $E^*$  corresponding to  $E'$  and a doxastic state  $E^{**}$  corresponding to  $E''$

in  $E^*$ : at *l*: avoids burning *a*: yes

in  $E^{**}$ : at *l*: best avoids burning by dropping and rolling (when pants are on fire), *a*; yes

But we have not answered the real demand to say precisely when  $E'$  and  $E''$ , involving *i*, motivate John in a special way that  $E^*$  at  $E^{**}$  do not.

To meet that challenge, we will do best to consider some simpler desires and actions. Suppose that I see a cookie on the table, and I want to eat it. I pick it up and eat it (if I have no special reason not to). I might do this because I am in a particular doxastic state:

in  $E$ : able to eat a cookie by picking it up and putting it in the mouth, etc., *i*; yes.

But such a doxastic state explicitly entertained would be rare. And this doxastic state need not play even an implicit role in motivating my action. Much of what I do is an immediate response to beliefs about what is around me and desires concerning those individuals. No mediating *belief* about my abilities is needed. Rather, I just respond by picking up the cookie. It seems clear that at some time in our lives we are capable of such responses without the corresponding beliefs about our abilities. I suggest that this is a continuing feature of our abilities to act. The belief that I have the ability does not need to play a role in my exercise of the ability.

Of course standards of belief attribution are notoriously slippery. One might be tempted to go as far as to say that the fact that I acted in the way I did in picking up the cookie is enough to show that I have the belief that I am able to do that. Young infants who reach out to pick up cookies simply have such beliefs "wired in".

But even if we were to take this more liberal position about belief attribution, we would still have to recognize an important distinction between self-attribution and attribution to others. Whether we say that I have an immediately available ability to pick up appropriately nearby objects when I want them (without requiring a belief that I have such an ability to satisfy the desire), or whether we say that I have a wired-in belief that I have such abilities to fulfill my desires, my relationship to my ability to fulfill my desires is different from my relationship to other individuals' abilities to fulfill their desires. My beliefs about what is around me are immediately available to serve my desires. I don't need to

explicitly entertain a further belief that I have the desire or a further belief about what abilities I have to fulfill the desire. Once I see the cookie I might gobble it down without explicitly entertaining any further beliefs about myself. To motivate giving the cookie to Jones to satisfy his desire, I would need beliefs about his desire (and, of course, desires of my own regarding the satisfaction of Jones's desire).

Beliefs that are at the immediate service of my desires are among the self-attributive beliefs (characterizable using '*i*'). But my special knowledge and/or abilities extend beyond abilities specifically related to fulfilling particular desires. As Barwise and Perry point out, for example, I have special knowledge (or a special ability to know) whether I am sitting or standing ([1], pp. 86–87). Self-attributions are beliefs about myself that connect with this network of special knowledge and/or abilities. Individual roles in belief-state characterizations represent the connection of a belief with a network of other beliefs. The individual role *i* represents the connection with the specially available network of knowledge and/or abilities that is available in satisfying my desires.

When I initiate dropping and rolling, I probably do not have to think about how to do it. Once I attain the appropriate desiderative condition

drop and roll, *i*; yes!

I just do it. In such a situation I would not ordinarily need to entertain a belief about how to drop and roll. This is the level at which I am just built with the ability to initiate the action.

When I have these beliefs:

pants on fire, *a*: yes

when pants are on fire, best avoids burning by dropping and rolling, *a*; yes

they might join with a desiderative state

avoids burning, *a*; yes

to produce a further desiderative state

drops and rolls, *a*; yes

Yet to motivate my action, I require some further belief about how I can get the anchor of *a* to drop and roll. In contrast, the corresponding doxastic and desiderative states involving *i* require no further beliefs on my part. To say that they involve *i* is to say that they connect directly with my desires and with my native knowledge and/or abilities to act.

Barwise and Perry's own treatment of self-attribution is not adequate. They tie the role *i* to a single event-type

*LI*: = at *l*: present, *a*; yes.

They simply define *i* as  $\langle a, LI \rangle$ , rather than explicitly recognizing it as connecting beliefs into a network of associated beliefs and desires. In addition, the event-type they employ in attempting to define the role *i* does nothing to guarantee that the role *i* will be a marker of self-attribution. Whenever I have a belief about anyone, the event-type *LI* might characterize my belief about that person just so long as I believe him or her to be somewhere. It is not at all clear how this was supposed to work in defining a role with a special relationship to me.

Nonetheless, it seems that Barwise and Perry's general apparatus of individual roles can be employed to mark the way in which beliefs get applied by being linked to other beliefs. And we can extend that apparatus to recognize a special individual role *i* that is distinguished by its links with immediate abilities to act in response to desires.

We should not think, however, that the special doxastic role of *i* corresponds to a special semantical role for "I" and "my". I come to an important realization in accepting as true "My pants are on fire" and not just "His pants are on fire". There is a major difference in my doxastic condition, because the property of *having burning pants* is tied in to a substantial network characterized by the individual role *i* and directly linked with my ability to drop and roll. But this property is still playing a role in a belief about a particular person—the belief that the property *having burning pants* applies to him. If bystander Smith sees me and notices what he would express by

His pants are on fire

and bystander Jones notices and warns

Your pants are on fire

their beliefs agree with mine both before and after my big realization.<sup>6</sup>

It has sometimes been thought that "My pants are on fire" expresses a very different belief from "His pants are on fire" when I come to accept it. Given that both express a belief that agrees with a belief held in common by Smith and Jones, I see no reason to say that they express distinct beliefs. The English words "I" and "my" refer to me by a different mechanism than do the English words "he" and "his", but they simply refer and bring to the proposition expressed only the referent. Nevertheless, if I am a normal speaker of English, then the word "I" (or "my") will be used in expressing that belief only after it is tied in with the network of beliefs and special abilities constituting the individual role *i*. When I accept "My pants are on fire" I will come to have many new beliefs (e.g., that for the first time in my life my pants are on fire, that the pants I bought just yesterday are on fire), but those beliefs are not expressed by "My pants are on fire".

## NOTES

1. Most notably in [6]. In [1] the special character of such beliefs is considered in several places, e.g., pp. 86–87, 248–249, 259–260.
2. Here I follow [2] and the parallel drawn there by Boër and Lycan between attributions of properties to other things and self-attributions. Barwise and Perry draw a similar parallel in [1], p. 249.
3. Peacocke, for example, introduces '[self]' for the mode of presentation that he assumes to be part of self-attributive beliefs. See [5], p. 108.
4. The clause 'if *S* has nothing better to do' is obviously a feeble attempt to take account of the fact that *S* has to choose from among the things he wants to do. The problems in clarifying this are irrelevant to the points under discussion here.

5. This is also suggested by Sosa and Pastin in [4]. Their suggestion is refined by me in [3].
6. We are not forced into the Fregean view that sentences like “My pants are on fire” express a proposition that only I can entertain, i.e., a proposition containing my unique mode of self-presentation as an element.

## REFERENCES

- [1] Barwise, Jon and John Perry, *Situations and Attitudes*, MIT Press, Cambridge, Massachusetts, 1983.
- [2] Boër, Steven and William Lycan, “Who, me?,” *Philosophical Review*, vol. 89 (1980), pp. 427–466.
- [3] McKay, Thomas, “Actions and de re beliefs,” *Canadian Journal of Philosophy*, vol. 14 (1984), pp. 631–635.
- [4] Pastin, Mark and Ernest Sosa, “A rejoinder on actions and de re beliefs,” *Canadian Journal of Philosophy*, vol. 11 (1981), pp. 735–739.
- [5] Peacocke, Christopher, *Sense and Content*, Oxford University Press, New York, 1983.
- [6] Perry, John, “The problem of the essential indexical,” *Noûs*, vol. 13 (1979), pp. 3–21.

*Department of Philosophy*  
*Syracuse University*  
*Syracuse, New York 13210*