



# LUND UNIVERSITY

## Barcan Marcus on Belief and Rationality

Olsson, Erik J

*Published in:*

Modalities, Identity, and Moral Dilemmas: Themes from Barcan Marcus

2015

[Link to publication](#)

*Citation for published version (APA):*

Olsson, E. J. (2015). Barcan Marcus on Belief and Rationality. In W. K. Essler, & M. Frauchiger (Eds.), *Modalities, Identity, and Moral Dilemmas: Themes from Barcan Marcus* (Vol. Lauener Library of Analytic Philosophy, vol. 3). De Gruyter.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00



# **Barcan Marcus on Belief and Rationality**

Erik J. Olsson

Olsson, E. J. (in press). Barcan Marcus on Belief and Rationality. In Essler, W. K. & Frauchiger, M. (Eds.)  
Modalities, Identity, and Moral Dilemmas: Themes from Barcan Marcus, Lauener Library of Analytic  
Philosophy, vol. 3. Walter de Gruyter.

## **1. Introduction**

Ruth Barcan Marcus is famous for her influential work in modal logic, the logic of necessity and possibility. Her perhaps most distinctive contribution to the subject, the so-called Barcan formula, stating that if it is possible that something has a certain property, then there is something that possibly has that property, gave rise to an extensive debate concerning the interpretation of systems of modal logic. While most of Marcus's work on the logic of necessity and possibility dates back some 50 years, she has been writing, more recently, on issues of belief and rationality, rejecting received accounts of propositional attitudes and attempts to systematize epistemic logic. This paper discusses this lesser known part of Marcus's contribution to philosophy and logic, which in a recent publication she describes as still being "in a formative stage" (Marcus, 2005). More narrowly still, I will focus on her reply to Kripke on a puzzle of belief for the theory of direct reference.

## **2. The theory of direct reference**

What is the meaning of a proper name, such as the name George W. Bush?

According to description theories of proper names, each speaker associates with a given name a description that picks out, as it were, the bearer of that name. It is here required that the description uniquely determines the name's referent. Thus, when a speaker uses the name *George W. Bush* and in doing so succeeds in referring to a particular object or individual *x*, he manages to do so because he thinks of George W. Bush as the unique

person having a certain property, e.g. that of being the current president of the United States, and x in fact does have this property uniquely. By contrast, a direct reference theory of the kind advocated by J. S. Mill (1867) states that the meaning of a proper name is simply its referent or bearer. This view was revived and further developed by Ruth Barcan Marcus in a paper in *Synthese* from 1961, and later adopted by Kripke and others.<sup>1</sup>

The difference between the two theories stands out clearly when one considers how they deal with identity statements of the type “a = b” where a and b are different names. Take for instance the case of “Cicero is identical with Tully”. On the direct reference theory, the name Cicero and the name Tully have the same meaning in virtue of referring to the same individual. So the identity statement simply says that that object is identical to itself, a necessary truth. On the description theory, by means of contrast, the statement “Cicero = Tully” is a mere contingent truth. To see this, suppose that, for a given speaker, the name Cicero means “the most famous ancient orator” and that the name Tully means “the person called ‘Tully’ by the English”. The identity statement in question then expresses that the most famous ancient orator is identical to the person called Tully by the English, which is only contingently true. It is contingent because the most famous ancient author could have been called something else by the English.

Both the description and direct reference theory have their well known advantages and disadvantages. The description theory has the advantage of giving an immediate answer to the question of how a name can refer to a given object. The answer is that it does so via the associated definite description. While the direct reference theory does not give an answer to that question by itself, it can be supplemented with a causal theory of reference that explains reference in terms of a causal chain connecting the user of a name with a previous process of baptizing. A problem with description theories is what definite description to associate with a given name. Famous advocates like Frege and Russell allow for different people to associate different descriptions with a given name, something which may appear intolerably subjective because it makes linguistic meaning relative to a speaker. The direct reference theory is clearly less vulnerable to charges of

---

<sup>1</sup> See Burgess (1996) for an extensive discussion of the relation between Marcus’s and Kripke’s work on proper names and related matters.

subjectivism. It has been suggested, e.g. by Susan Haack (1978, pp. 64-65), that the two theories are not necessarily rivals but that they may give complementary accounts of the meaning of proper names. Be that as it may. I shall proceed to discuss a puzzle that arises for the direct reference theory in the context of rational belief.

### **3. A puzzle about belief**

In his paper “A puzzle about belief” from 1979, Saul Kripke raised the following problem for the theory of direct reference. Suppose that all assent is sincere and reflective so that speakers are not conceptually or linguistically confused. Then the following disquotational principle seems highly plausible:

(DP) If (a) a normal English speaker assents to ‘p’, and (b) ‘p’ is a sentence of English, then he believes that p.

Now a speaker may assent to

(i) Cicero was bald

and seemingly coherently assent to

(ii) Tully was not bald

And hence also to the conjunction:

(iii) Cicero was bald and Tully was not bald.

By the disquotational principle, the speaker since he assents to (iii) must be taken to believe that sentence to be true. Yet, on the theory of direct reference, the names Cicero and Tully have the same meaning which means that the conjunction (iii) is a plain

contradiction. It says of something that it both is and is not bald. The example suggests that, on the theory of direct reference, a fully rational person may believe a contradiction without having committed any logical blunder or being subject to linguistic or other confusion.

Here is another example, this time of a bilingual nature. From what he has heard or read about the attractiveness of Londres, Pierre a native Frenchman assents to

(iv) Londres est jolie.

Pierre now emigrates to England where he takes up residence in a neighborhood in London which he finds ugly. There he learns English by exposure, without recourse to translation manuals, dictionaries and the like. He now assents to:

(v) London is not pretty.

Hence, he would also assent to:

(vi) Londres is pretty and London is not pretty.

By a similar (bilingual) disquotational principle, Pierre believes the conjunction (vi) which, by the theory of direct reference, is a plain contradiction saying of London that it is both pretty and not pretty. Pierre exemplifies the predicament of a person who has fallen into contradiction through no fault of his own.

#### **4. Marcus's proposed solution**

Kripke views the situation as deeply unsatisfactory. "We may suppose that Pierre ... is a leading philosopher and logician. He would never let contradictory beliefs pass ... He lacks information, not logical acumen. He cannot be convicted of inconsistency."

Marcus, too, in her 1981 paper, insists that we cannot rest content with this state of

affairs. One way of solving the puzzle would be to give up the theory of direct reference in favor of a descriptive theory according to which, as we saw, proper names have a meaning distinct from their reference. On the descriptive theory, the names *Londres* and *London* may mean different things for Pierre in virtue of being associated with different definite descriptions. For Pierre, “*Londres*” could mean “the town I read about in my geography book” and “*London*” could mean “The town in which I live”. If so, the statement “*Londres* is pretty and *London* is not pretty” will be false but not contradictory. It will be false in virtue of the contingent fact that the town Pierre read about happens to be the town in which he now lives. That a fully rational person should entertain some false but non-contradictory beliefs is unfortunate but hardly surprising.

Marcus is not inclined to give up the theory of direct reference in response to Kripke’s puzzle. Instead, her proposal amounts to arguing that it is the disquotational principle that is ultimately to be blamed for the untoward result. As a preliminary, she rejects the common view that believing is a “propositional attitude”, in the sense of an attitude the believer has to a proposition, in favor of the proposal that the object of belief is a *state of affairs*. She goes on to suggest the following condition, which I shall refer to as the *belief-possibility thesis*:

(BP) If X believes that p, then possibly p.

According to this principle, belief is not an entirely internal matter, but what one believes, and can believe, depends on what is, in an external mind-independent sense, possible. As Marcus points out, this part of her proposal consists in “strengthening the connection between belief and ‘reality’, i.e. possible worlds or structures” (Marcus, 1981, p. 507).<sup>2</sup>

The following modified disquotational principle now suggests itself:

(DP\*) If (a) a normal English speaker assents to ‘p’, (b) ‘p’ is a sentence of English, and (c) p is *possible*, then the speaker believes that p.

---

<sup>2</sup> The reference to “structures” is explained on p. 139 in Marcus’s 1990 paper. There she writes “[w]e may think of states of affairs as ordered structures of actual objects which includes individuals as well as properties and relations.”

The effect of this modification is a weakening of the connection between assertion and belief. Assent carries over into belief only if the state of affairs in question is possible.

How does this solve the puzzle about belief? Consider again the case of Pierre, who would assent to “Londres is pretty but London is not”, which on the theory of direct reference is a contradiction. On the old disquotational principle, bilingually adjusted, we would be forced to ascribe to Pierre a contradictory belief. This is not so, of course, on the modified principle. Since the sentence in question is contradictory it expresses a state of affairs that is impossible, and so we are not obliged to attribute belief in this case.

The bottom line is that the theory of direct reference can be saved if we are willing to say that belief is not a wholly subjective affair, but something that is governed by Marcus’s belief-possibility thesis according to which belief in *p* requires that *p* express a state of affairs that is possible. But what independent reason is there for taking this thesis to be true? This is the question to which I now turn.

## **5. The original motivation for the belief-possibility principle**

In her original 1981 paper on Kripke’s puzzle, Marcus gives two main reasons why her belief-possibility thesis should be adopted. The first is an argument from linguistic intuition, the second one from analogy. Let us consider the linguistic argument first. Marcus writes (505):

Suppose that someone were to claim that he believes Hesperus is not identical with Phosphorus or that Tully is not identical with Cicero, or that London is not the same as Londres ... It is my (non post-hoc) intuition that on *discovery* that those identities hold, and consequently that the associated name pairs name the same thing, I would *not* say that I had changed my belief or acquired a new belief to replace the old, but that I was mistaken in claiming that I *had* those beliefs to begin with.

But many readers would no doubt find that they themselves lack Marcus's reported intuition. From a pre-systematic standpoint, it seems more natural to say simply that we used to believe, falsely, that Cicero is not identical with Tully but that we have now corrected this belief.

Let us instead turn to the other argument, which makes use of what Marcus thinks of as an "analogy" between knowledge and belief (1981, p. 505), the idea being that there is a close parallel in the following sense:

Just as a condition for knowing that *p* is that *p* obtains, so a condition for believing that *p* is: if *X* believes that *p*, then possible *p*.

Perhaps we can reformulate this proposal more succinctly in this way:

(KB) Just as knowing that *p* requires that *p* *actually* obtains, so believing that *p* requires that *p* *possibly* obtains.

Now this is an interesting thought. What (KB) expresses is a formal relationship between knowing and believing, indeed one of considerable appeal. Still, I believe that Marcus is making too much of her parallel when she refers to it as an "analogy". It is not that the *very same* property holds for both knowing and believing. What is indicated is rather a weaker kind of structural relationship which cannot strictly be subsumed under the more demanding concept of an analogy. It must also be said that the appeal of this structural connection derives from its simplicity and elegance, features that are not obviously indicative of truth. Hence, while Marcus's appeal to (KB), or something similar, does constitute an independent argument for her solution to the belief puzzle, the strength of that argument should not be overestimated.

## **6. Later motivations for the belief-possibility thesis**

In a later paper called "Rationality and Believing the Impossible", which appeared in *The Journal of Philosophy* in 1983, Marcus returns to the task of giving independent

motivation for her belief-possibility thesis, her first point being that Berkeley apparently took a similar position in his *Principles of Human Knowledge*. Having argued that existence of matter involves a “contradiction”, Berkeley goes on to say (in paragraph 54),

Strictly speaking, to believe that which involves a contradiction ... is impossible ... . In one sense, indeed, men may be said to believe that matter exists; that is, they act *as if* the immediate cause of their sensations, which affects them every moment ... were some senseless unthinking being. But that they ... should form thereof a settled speculative opinion is what I am unable to conceive.

Berkeley concludes:

This is not the *only* instance where in men impose upon themselves, by imagining they believe those proposition they have often heard, though at bottom they (those propositions) have no meaning in them.

As Marcus notes, Berkeley seems to be claiming that sentences that describe impossible states of affairs are *meaningless*, whereas Marcus herself is only making what is plausibly seen as a weaker claim, namely, that “where a state of affairs is impossible, there is a sense of ‘belief’ such that an agent is mistaken in claiming that he is in the belief relation to that state of affairs” (1983, pp. 324-5).

In the same paper, Marcus refers, for indirect support, to an experiment described by Donald Davidson on pp. 235-36 in his 1980 book (quoted from Marcus, 1983, footnote 15 on p. 330):

After spending several years testing variants of Ramsey’s theory [of belief] on human subjects, I tried the following experiment (with Merrill Carlsmith). Subjects made all possible pairwise choices within a small field of alternatives, and in a series of subsequent sessions, were offered the same set of options over and over. The alternatives were complex enough to mask the fact of repetition, so that subjects could not remember their previous choices, and pay-offs were

deferred to the end of the experiment so that there was no normal learning or conditioning. The choices for each session and each subject were then examined for inconsistencies – cases where someone had chosen  $a$  over  $b$ ,  $b$  over  $c$ , and  $c$  over  $a$ . It was found that as time went on, people became steadily more consistent; intransitivities were gradually eliminated; after six sessions, all subjects were close to being perfectly consistent ... apparently, from the start there were underlying and consistent values which were better and better realized in choice. I found it impossible to construct a formal theory that could explain this, and gave up my career as an experimental psychologist.

Marcus remarks that, “[i]n Davidson’s experiment it is as if the subjects who assented to sentences that are inconsistent declined to carry them over into belief. Such a subject, if my intuition is shared, would not say, ‘I once believed I preferred  $a$  to  $b$ , and  $b$  to  $c$ , and  $c$  to  $a$  but now I don’t’. He would disclaim having had such a belief.” (1983, footnote 15, p. 330).

There are a number of reasons for being dissatisfied with this interpretation of the experiment. First, while we may agree with Marcus that a subject would not say “I once believed I preferred  $a$  to  $b$ , and  $b$  to  $c$ , and  $c$  to  $a$  but now I don’t”, this is so for the trivial reason that a subject would not be in a position to say anything about what she previously preferred, or believed she preferred. Recall that as Davidson describes the experiment, it was set up in such a way that “subjects could not remember their previous choices”. Second, even if we assume that the subject did recall their previous preferences, which would require another experimental setup altogether, Marcus’s interpretation of the experiment makes use of the very “intuition” that that experiment was intended to support. For these reasons, the belief-possibility thesis is not even “supported indirectly” by Davidson’s experiment.

## **6. A problem for Marcus’s thesis**

It should be noted that, while Marcus seems to think of her account of the object of belief being a state of affairs rather than something propositional as playing a central role in her

proposed solution to Kripke's puzzle, these are seen on closer scrutiny to be two independent views. There are, after all, impossible states of affairs just as there are impossible propositions, and neither Marcus nor anyone else has, to my knowledge, produced any convincing argument to the effect that it should be more tempting to exclude belief in the impossible on one account rather than the other, as it would have been if, for instance, there were impossible propositions but no impossible states of affairs. For this reason, I see no point in upholding the distinction between propositions and states of affairs in the present context, although I believe that this distinction can do useful work in other argumentative settings.<sup>3</sup>

But the real problem that I would like to raise lies elsewhere, for we may legitimately ask whether Marcus's proposal really solves the problem at hand. To be sure, her move does prevent us from ascribing belief in a contradictory sentence to a perfectly rational person. It is still true, though, that our Pierre believes two things that are jointly incompatible, even if he does not believe a contradictory sentence. Pierre assents to each of "Londres is pretty" and "London is not pretty", and so on the original as well as the revised disquotational principle, he comes out believing two sentences that, given the direct reference theory, together form an inconsistent set. That seems almost as bad as believing in a contradictory sentence. For how can a perfectly rational person who is not in any way confused or logically or linguistically misguided believe two things which, for reasons of logic alone, cannot both be true?

In other words: it is still true that the totality of a perfectly rational person's beliefs can be inconsistent, so that there is no possible world where all the person's beliefs are true. My point is that this may seem almost as baffling as the thought that such an agent could believe a contradictory sentence. Marcus's proposal does not address that remaining difficulty.<sup>4 5</sup>

---

<sup>3</sup> See Marcus (1995) for an extended argument to the effect that the distinction is needed to avoid a questionable exclusion of non-linguals from the domain of creatures having beliefs and desires.

<sup>4</sup> Brown (1991), p. 358, argues that Marcus's endorsement of the belief-possibility thesis in fact also commits her to what I have called the strong belief-possibility thesis: "Marcus's argument, if successful, would require us to deny not only that one can believe the impossible but, also, contrary to her intentions, that one [can] have contradictory beliefs." See also Altricher (1985) for a related point.

<sup>5</sup> Furthermore, it is still true that a perfectly rational person may fail to believe in the logical consequences of what he believes. A person may assent to "Cicero was bald" without assenting to "Tully was bald", although, on the theory of direct reference, these two sentences mean the same thing and hence should be logically equivalent, and so the one should follow from the other. It is not just the case that logical closure

## 7. The strong belief-possibility thesis

Fortunately, there is a quick fix of Marcus's proposal that would block the more general inconsistency problem to which I just drew attention. We recall that Marcus proposes the following belief-possibility thesis:

(BP) If X believes that p, then the state of affairs that p is possible.

The best motivation for that principle is, again, the following proposal for a structural relationship between knowing and believing:

(KB) Just as knowing that p requires that p *actually* obtains, believing that p requires that p *possibly* obtains.

But it seems just as plausible to suggest that

(KB\*) Just as there is a possible world (namely the actual world) where everything *known* by a person is true, so there is a possible world where everything *believed* by a person is true.

Hence,

(SBP) There is a possible world where everything believed by a person is true.

---

fails when conjoining two beliefs results in contradiction. That logical closure fails in this way is a consequence of Marcus's proposal. But it is also the case that it fails when there is no apparent inconsistency. To make this quite clear: suppose that the person assents to "Cicero was bald" but assents to neither "Tully was bald" nor "Tully was not bald", perhaps because he has no idea who Tully was. In that case, logical closure fails without any inconsistency being inflicted. This problem, too, is not addressed by Marcus's approach, but one may hold it to be less severe than the inconsistency problem, which is why I shall not here pursue this line of thought any further.

This strong belief-possibility thesis seems no less reasonable than the weaker thesis proposed by Marcus. In both cases, we may draw on an attractive formal relationship between knowing and believing.<sup>6</sup>

The point of course is that the stronger belief-possibility principle blocks the general inconsistency problem. On the stronger principle, for the person to believe separately that Cicero is bald but Tully is not there has to be a possible world where both these sentences are true. But this is prohibited by the theory of direct reference, according to which the names Cicero and Tully have the same meaning in virtue of denoting the same individual in all possible worlds. The same individual cannot be both bald and non-bald.

While the original belief-possibility thesis suggested a natural modification of Kripke's disquotational principle, this is not so for the stronger version. One might think that something like this would work:

(DP\*\*) If (a) a normal English speaker assents to 'p', (b) 'p' is a sentence of English and (c) *p is possible given the speaker's other beliefs*, then the speaker believes that p.

---

<sup>6</sup> Here is a slightly different argument. The following is true of knowing:

(K) If X knows that p, then there is a possible world w (the actual world) such that p is true in w.

On the basis of this property, we might postulate, with Marcus, that something similar should be true of believing, namely

(BP) If X believes that p, then there is a possible world w such that p is true in w.

But what we have observed, in effect, is that knowing actually satisfies a principle stronger than (K):

(K\*) There is a possible world w (the actual world) such that, if X knows that p then p is true in w.

If we follow Marcus in her attempt to strengthen the structural similarity between knowing and believing, we might want to require in addition to (BP) that the following hold:

(SBP) There is a possible world w such that, if X believes that p, then p is true in w.

The latter is of course our stronger belief-possibility thesis.

But suppose again that we come across a speaker who assents separately to “Cicero was bald” and “Tully was not bald” in that order. Using the proposed revised disquotational principle, we would first ascribe to the speaker belief in “Cicero was bald”, assuming that sentence to be consistent with his other beliefs, but not ascribe belief in “Tully was not bald”. But if we switch the order of the sentences so that “Tully was not bald” is first assented to, we get the opposite result that the speaker believes that Tully was not bald but not that Cicero was bald. This dependence on the order of evaluation is surely hard to swallow.

My own view on the matter is that there is no simple relationship between assent and belief. A person may assent to a given proposition, and yet non-verbally act as if the opposite were true. In many such cases we are inclined to assign greater weight to her non-verbal actions. In her later works, Marcus has expressed doubts concerning any disquotational principle that asserts “a privileged and overriding status of certain speech acts as belief indicators in language-using agents” (1983, p. 335). In the 1983 paper she is still apparently undecided concerning the correct notion of rationality. In the paper from 1990, however, she expresses unequivocal support for what she calls a “broader notion of rationality” (p. 138), according to which non-verbal behavior also counts as an indicator or counter-indicator of belief. On this new view, “[i]t is not supposed that the act of sincere assent even where evoked must be an overriding indicator of belief” (1990, p. 141).<sup>7</sup> I believe therefore that she would not consider the mere failure of the strong belief-possibility thesis to provide the basis for a simple principle linking assent to belief to be a convincing argument against that stronger thesis.

## 8. Consequences for the theory of belief revision

---

<sup>7</sup> In her 1990 paper, Marcus advances a dispositional account of belief according to which “x believes that S just in case under certain agent-centered circumstances including x’s desires and needs as well as external circumstances, x is disposed to act as if S, that actual or non-actual state of affairs, obtains” (p. 140). For a discussion, see Engel (1999). This account is used as part of an account of belief that does justice to the plausible pre-systematic intuition that animals and other non-linguals can also have beliefs (and desires). I wonder, though, whether Marcus is not guilty of overkill. She advances an account of belief having essentially two components. Belief is construed as, on the one hand, being a disposition and, on the other hand, as taking something non-linguistic as its object. It seems to me that either component in isolation would actually do the job of securing the possibility of belief and desire in non-linguals. Non-linguals (like lower animals) can obviously be disposed to act, and so they can have beliefs in that sense. If Marcus is correct, non-linguals can also stand in relations to non-linguistic facts, so this condition, too, is satisfied. If so, why invoke *both* a dispositional *and* a non-propositional condition on belief?

The strong belief-possibility thesis, if accepted, has implications that go beyond the dispute concerning direct reference. The discussion concerning whether or not a fully rational believer can end up believing an impossible proposition is not uniquely tied to identity cases. As Isaac Levi has convincingly argued, observation is plausibly construed as routine expansion whereby a proposition is accepted in a routine fashion (e.g. Levi, 1980). This may, in his view, lead to inconsistency in the set of full beliefs, for what comes to be believed in this fashion, *p* say, may contradict beliefs that the agent has acquired previously. So, the agent may come to believe both that *p* and that non-*p*. Given logical closure of beliefs, she will end up believing a blatant contradiction.

Now it may be claimed that what is special about the identity cases is that the agent will be unable to detect the impossibility, even if he or she is logically omnipotent. But that this is a genuine difference could be questioned; for it could be argued that the agent, while in the inconsistent state (there is only one inconsistent state assuming beliefs to be closed under logical consequence), is unable to pursue any rational inquiry at all, including inquiry aimed at detecting inconsistency. In a recent debate with Levi in the journal *Synthese*, I have argued that this is indeed a consequence of Levi's theory of routine expansion in connection with his claim that the inconsistent belief state is useless for inquiry and deliberation, that it is "epistemic hell".<sup>8</sup>

This is a troublesome consequence. If inconsistency means hell, how can it ever be legitimate for a rational person to enter that state, and on what basis could consistency be regained? It seems that an inconsistent belief state, if there were such a thing, would be an intellectual point of no return. I suspect that philosophers defending the possibility of having inconsistent beliefs based on our supposed intuitions about the matter underestimate the grave implications of their view (e.g. Altrichter, 1985).

Central in Levi's earlier theory was his claim that it is possible to give principled advice for how a person should extricate herself from inconsistency. His theory of "coerced contraction", as detailed in his 1991 book, was intended to provide such guidance. However, as I pointed out in my *Synthese* paper, his proposal simply won't

---

<sup>8</sup> See Olsson (2003). The expression "epistemic hell" stems from Peter Gärdenfors (1988).

work: again, once in epistemic hell, there is no rational basis for inquiry, and this goes in particular for inquiry and deliberation into how to escape inconsistency.

In his response to my criticism, Levi devised a new theory of contraction from inconsistency which, unlike his earlier attempt, makes extrication from inconsistency a matter not of deliberation but of routine, the idea being that the agent can precommit herself to a routine for how to handle inconsistency. The routine will automatically, as it were, catapult the agent from the inconsistent state without there being any need for deliberate efforts on his or her part.<sup>9</sup> Yet, there are reasons to be seriously unsatisfied with Levi's new proposal as well, if only because it seems psychologically unrealistic to think that people would be equipped with such elaborate contingency plans for inconsistency-handling. Also, I fail to see how such a routine could be preprogrammed for taking care of all eventualities, that is to say, all the various ways in which an inconsistency may arise. Considering the grave consequences of ending up in epistemic hell, extensive contingency planning would be imperative. A detailed assessment of Levi's new theory would require a longer discussion, which is best left for another occasion.

This is where Marcus's belief-possibility thesis in its stronger form comes in. Perhaps we should, in Marcus's spirit, think of the inconsistent state not as *epistemic* hell but as something that is not epistemic or doxastic at all. Perhaps, then, an inconsistent state should not be regarded as a state of *belief*. That would open up for a new solution to Levi's long-standing problem of inconsistency, a solution that *prima facie* requires less delicate footwork than Levi's own recent proposal. It would be a radical solution, to be sure, and yet, to paraphrase a famous fictional detective, when all the non-radical possibilities have been examined, the remaining possibility, however radical, has a good chance of being true. At the very least, this approach does seem worthy of serious consideration.<sup>10</sup>

---

<sup>9</sup> See Levi (2003).

<sup>10</sup> A drawback of this suggestion is that the nice Boolean algebraic structure of potential states of full belief is sacrificed. We cannot anymore think of the space of potential state of full beliefs as featuring two designated states, 1 and 0, and being closed under the two operations of "meet" and "join" (Levi, 1991). While there is still a 1 element corresponding to the belief state in which only tautologies are held true, there is no 0 element corresponding to the contradictory belief state. Since the join of two consistent potential belief states may be inconsistent, closure under it must also be sacrificed, whereas closure under the meet operation can be retained. If potential belief states are represented as sets of sentences of some

How would the proposal for handling belief-contravening observations work in practice? Suppose the inquirer, while being in belief state K in which non-p is believed, routinely expands by p by relying on a trusted source. What is the new state of belief after the routine has been carried out? On the current proposal, that belief state cannot be the inconsistent state for there is no such state of belief. We may instead ask what *consistent* belief state most plausibly reflects the current commitments of the inquirer. Presumably, that belief state is still K but now augmented with the further belief that the routine in question was invoked giving a certain result, p, that is inconsistent with K. Augmenting K with this information of a report character does not by itself make the new belief state inconsistent. The inquirer may well be in a state in which non-p is believed and also believe that one of his or her previously trusted routines gives a conflicting result.

## 9. Conclusion

I have argued that Marcus's intriguing proposal for how to solve Kripke's puzzle about belief in the context of a theory of direct reference, while not fully satisfactory as it stands, can be amended in a way that takes care of a remaining problem. The amendment amounts to a strengthening of what I have called her belief-possibility thesis. According to the stronger thesis, a fully rational person's beliefs must be jointly possible in order to count as beliefs at all. Once this has been sorted out, Marcus's general observation still stands: we may continue to endorse the attractively simple theory of direct reference on the condition that we accept a tighter connection between belief and reality. There are, not surprisingly, powerful arguments from the standpoint of ordinary language against such a theory of belief. Alrichter (1985) produces a long list of counter intuitive consequences of Marcus's view, some more serious than others, whereas that view receives a more sympathetic treatment in Brown (1991). I believe that what has been said above adds some weight to the Marcus-friendly side of the dispute. As I have argued, the view that beliefs are by their very nature consistent is in line with a philosophically robust view of inconsistency as a state in which all coherent inquiry and deliberation

---

regimented language, the meet of two such states corresponds to their intersection, and the join of the states to their union.

would break down, so that there would be no rational deliberative escape route. That this is so has, I believe, gone largely unnoticed by Marcus's critics, who have been eager to ascribe inconsistent beliefs in a number of cases, often in an effort to explain action. What they have forgotten is to explain how consistency was, or at least could have been, regained. So, while Marcus's conception of belief is to some extent "revisionary", to use her own characterization in the 1990 paper,<sup>11</sup> it may still be a fruitful explication in the sense of Carnap (1950). Its systematic advantages may in the end outweigh any doubts that arise from what we would say in various cases. A fuller investigation into which view is to be preferred, all things considered, will have to await another occasion.<sup>12</sup>

## References

- Altrichter, F. (1985) "Belief and possibility", *Journal of Philosophy* 82(7):364-381.
- Brown, C. (1991) "Believing the impossible", *Synthese* 89(3):353-364.
- Burgess, J. P. (1996), "Marcus, Kripke, and names", *Philosophical Studies* 84(1):1-47.
- Carnap, R. (1950), *Logical foundations of probability*, Chicago University Press.
- Davidson, D. (1980), *Essays on actions and events*, Oxford University Press.
- Engel, P. (1999), "Dispositional belief, assent, and acceptance", *Dialectica* 53 (3-4): 211-226.
- Gärdenfors, P. (1988), *Knowledge in flux: modeling the dynamics of epistemic states*, MIT Press.
- Haack, S. (1978), *Philosophy of Logics*, Cambridge University Press.
- Kripke, S. (1979), "A puzzle about belief", in Margalit, A. (ed.), *Meaning and Use*, Dordrecht.
- Levi, I. (1991), *The fixation of belief and its undoing*, Cambridge University Press.
- Levi, I. (2003), "Contracting from epistemic hell is routine", *Synthese* 135: 141-164.

---

<sup>11</sup> Altrichter (1985), as mentioned, emphasizes the divergence between Marcus's account of belief and presystematic intuition on the matter.

<sup>12</sup> An earlier version of this paper was read at the 3rd International Lauener Symposium on Analytic Philosophy in Honour of Prof. Ruth Barcan Marcus in Bern 2008. I would like to thank the members of the audience for their comments, including Dagfinn Føllesdal, Timothy Williamson and, above all, Ruth Barcan Marcus herself. I consider myself fortunate to have had the opportunity to meet Ruth, a remarkable philosopher and person, before she passed way.

- Levi, I. (2004), *Mild contraction: evaluating loss of information due to loss of belief*, Oxford University Press.
- Marcus, R. B. (1961), "Modalities and intensional language", *Synthese* 13: 303-322.
- Marcus, R. B. (1981), "A proposed solution to a puzzle about belief", *Midwest Studies in Philosophy*.
- Marcus, R. B. (1983), "Rationality and believing the impossible", *The Journal of Philosophy* 80 (6): 321-338.
- Marcus, R. B. (1990), "Some revisionary proposals about belief and believing", *Philosophy and Phenomenological Research* 50, supplement: 133-153.
- Marcus, R. B. (1995), "The anti-naturalism of some language centered accounts of belief", *Dialectica* 49 (2-4): 113-129.
- Marcus, R. B. (2005), "Ruth Barcan Marcus", in Hendricks, V. F., and Symons, J. (eds.), *Formal Philosophy*, Automatic Press.
- Olsson, E. J. (2003), "Avoiding epistemic hell: Levi on pragmatism and inconsistency", *Synthese* 135: 119-140.