Ratio (new series) XII 4 December 1999 0034–0006

EXPLAINING NORMATIVITY: ON RATIONALITY AND THE JUSTIFICATION OF REASON[1]

Joseph Raz

Abstract

Aspects of the world are normative in as much as they or their existence constitute reasons for persons, i.e. grounds which make certain beliefs, moods, emotions, intentions or actions appropriate or inappropriate. Our capacities to perceive and understand how things are, and what response is appropriate to them, and our ability to respond appropriately, make us into persons, i.e. creatures with the ability to direct their own life in accordance with their appreciation of themselves and their environment, and of the reasons with which, given how they are, the world presents them.

An explanation of normativity would explain the various puzzling aspects of this complex phenomenon. In particular it would explain how it is that aspects of the world can constitute reasons for cognitive, emotive, and volitional responses; how it is that we can come to realise that certain cognitive, emotional or volitional responses are appropriate in various circumstances, and inappropriate in others; and how it is that we can respond appropriately. This paper explores an aspect of the last of these questions.

1. Normativity and rationality

The normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons. The normativity of rules, or of authority, or of morality, for example, consists in the fact that rules are reasons of a special kind, the fact that directives issued by legitimate authorities are reasons, and in the fact that moral considerations are valid reasons. So ultimately

[1] The first version of section one was included in papers presented at the Philosophy Colloquium at Berkeley, at the conference in memory of Jean Hampton at Tucson, and at a conference on practical reason in the Humbolt University, Berlin. I learnt from questions and comments of many who participated in those occasions. I owe a special debt to David Silver, who was the commentator on my paper at Tucson, and to Jonathan Dancy for most helpful comments on a later version of the paper.

the explanation of normativity is the explanation of what it is to be a reason, and of related puzzles about reasons.

Reason[2] is inherently normative. That is its central characteristic. Therefore, the accounts of normativity and of reason and rationality, though not identical, are inter-related. An account of rationality is an account of the capacity to perceive reasons and to conform to them, and of different forms of conforming to reasons, and their appropriateness in different contexts. To explain the capacity to conform to reason the account must explain the possibility of error, failure to perceive reasons correctly, and of failure to respond to them once perceived. An account of irrationality is an account of some of the ways of failing to conform to reason, those which render one, or one's behaviour, or emotions, etc. irrational. The core idea is that rationality is the ability to realise the normative significance of the normative features of the world, and the ability to respond accordingly.

In one sense of "rational", we, or anything else, are rational beings to the extent that we possess that ability, which I will call "capacity-rationality". The absence of capacity-rationality does not mean that a creature is irrational. It means that no judgements of rationality apply to that creature. In another sense, we, or anything else, are rational to the extent that in general we use that ability well. There is a further use of "rationality" in which it applies to specific human responses, or their absence. Our actions, intentions, beliefs, emotions, etc. can be severally rational, non-rational, or irrational.

Rationality makes us into persons. To yield an explanation of rationality the core idea has, of course, to be hedged and refined. I will confine myself to four observations which help locate the relations between capacity-rationality and normativity.

First, to be people who are rational in the first sense, that is to be people with the ability to perceive reasons and respond to them, we need a range of capacities which do not directly contribute to our rationality. They include some perceptual ability, and the capacity to control our movements at will. An impairment of our perceptual ability does not diminish our rationality. Nor does lack of muscular control, or other neurological or physical impairments of our ability to move at will. Possession of

[2] I will be using "reason" when it sounds natural to use it. The expression refers variously to all or some of the following: people's reasoning capacity, people's use of that capacity when referred to in a general way, the reasons which apply to people on the occasion(s) discussed. I will sometimes leave it to context to identify the meaning.

at least some perceptual ability, and of some ability to control one's movements at will, are presupposed by capacity-rationality.[3] But they are not themselves constituents of rationality.

There are other preconditions of capacity-rationality. Among them are psychological capacities, such as the possession of memory, the capacity for conceptual thought, and the capacities to form beliefs, and to reach decisions. Some of the preconditions relate to capacities possession of which is a precondition of being either rational or irrational, but where their successful exercise does not render one (or one's judgement, action, etc.) rational, nor does their failure render one irrational. Perceptual failures, muscular failures, failures of memory are examples of preconditions of this kind. The abilities to form beliefs and to reach decisions belong to the other kind. Success and failure in their exercise contribute to the evaluation of the rationality of one's beliefs, actions, etc., as well as to the evaluation of the rationality (in the second sense) of the person concerned.

Not surprisingly the boundary between those who do not meet the preconditions of capacity-rationality, and are neither rational nor irrational, and those who meet the preconditions, and are irrational, is not a sharp one. Someone who lacks a minimal capacity to make up his mind about anything fails the preconditions and is neither rational nor irrational. He lacks capacity-rationality. But those who have a minimal ability to make up their minds, and constantly vacillate, finding it always difficult to make up their minds, and almost impossible not to change almost every decision they take soon after taking it, are irrational.

What is capacity-rationality in itself? It is a capacity to see the normative significance of the way things are, to comprehend what reasons they constitute, and the significance of that fact for oneself.[4] As indicated at the outset, part of the explanation of

[3] My claim is that their possession to some degree, during some part of the creature's life, is necessary for that creature to have capacity rationality at all, not merely for his ability to display that capacity or exercise it. It is true, though, that beyond that minimum lacking the ancillary capacities may impede the exercise of rationality, or make it temporarily impossible.

[4] Needless to say possession of this capacity itself is a matter of degree. Some humans and other creatures can recognise (the normative aspect of) some types of reasons but not of others. Some humans and other creatures can reflect about the fact that the existence of certain facts constitutes reasons, and form general views about the nature of reasons and rationality. Others can do so to a limited degree, or cannot do so at all. There is little point in trying to fix a test of personhood which will endow the concept with relatively sharp boundaries, which it does not possess.

normativity consists in the explanation of this capacity. Here again we encounter the overlap between the explanations of normativity and of rationality.

These remarks lead to my *second* point. Some accounts of rationality identify it with a reasoning ability. That ability consists, at least in part, in the ability to recognise inferential relations. That is it consists of, or includes, an ability to figure out what conclusions follow from given premises (the active side of the capacity), and an ability to recognise that inferences are valid or invalid[5] when this is pointed out to one (the passive side of the capacity). While at least a minimal reasoning ability is among the constituents of capacity-rationality, it does not exhaust it, at least not if the capacity to reason is the ability to construct and understand inferences.

That a certain proposition follows from certain premises is, other things being equal,[6] a reason for not believing the premises without believing the conclusion.[7] Given that rationality is the capacity to realise the normative significance of facts, that is to realise whether they constitute reasons, and which reasons they constitute, and to respond appropriately, the capacity to see the normative significance of inferences is high among the constituents of capacity-rationality. But what of reasoning power in general? One possibility is that the capacity to realise that C follows from P is like the capacity to see that the house is on fire. They are ways of realising the existence of facts which constitute reasons, and their possession is a precondition of being able to become aware of the normative significance of those facts, and to respond appropriately. On this account neither perceptual capacity nor reasoning ability are themselves constituents of

[5] I am using these terms to designate the success and failure of any inference, not merely of deductive ones.

[6] The existence of defeasible valid inferences requires the qualification "other things being equal". Typically when the inference is defeated it yields no reason at all.

[7] It is not, however, not even when the inference is a deductive one, a conclusive reason. Mere knowledge that a set of propositions is self-contradictory is a reason, if one does believe in at least some of them, to refrain from believing in one or more of them so that one's beliefs will not be self-contradictory. When one has no further information about the location of the contradiction and the reasons for it, the only way to know that one conformed with this reason is to refrain from believing any of the propositions in the set. When the contradictory set is large the price of doing so can be very great. Sometimes, it may be impossible to conform with this reason (e.g. if the set includes all one's beliefs). But even when it is possible it may be unjustified, given the price. After all, knowledge that the set is contradictory is no more than knowledge that one of its propositions is false. We know on inductive grounds, that even if our beliefs are consistent at least one of them is false. That is no (adequate) reason to suspend all our beliefs. Why should the fact that our beliefs are contradictory be such a reason?

rationality. They are merely preconditions of its exercise. Alternatively, it is arguable that reasoning is unlike perception in that it is involved in almost any recognition of the normative significance of anything. Even recognising that since C follows from P one has, other thing being equal, reason not to believe P and reject C involves reasoning. The close involvement of reasoning in capacity-rationality would justify, on this view, regarding the ability to reason as a constituent of rationality.

There is no doubt that reasoning ability is closely involved with rationality, even though the pervasiveness of its involvement may be debated. Yet it is doubtful that this involvement would in itself justify the identification of reasoning ability with capacity-rationality. There is, however, a better argument for that conclusion: realising that C follows from P is not merely realising the existence of a fact which is a reason (as in seeing that the house is on fire). It is realising that there is a reason. Understanding that C follows from P is the same as, or at least involves, understanding that, other things being equal, one has reason not to believe P and reject C. That is part of what it means that the one follows from the other.

This argument notwithstanding, this seems to be a case where our concepts are not as neat and tidy as philosophers may wish them to be. While possession of a minimal reasoning capacity is a constituent of capacity-rationality, that ability may be modest. In general, failures of our reasoning powers do not warrant a judgement of irrationality. Failure to see that a conclusion follows from certain premises most commonly merely shows that one is not very bright, or just not very good at reasoning. In most cases it does not show that one is irrational, as failure to realise that if a conclusion follows from premises then one has reason not to believe the premises while rejecting the conclusion does. On the other hand, failure of *elementary* reasoning does establish irrationality: Failure to realise, in normal circumstances, that it follows from the fact that one's destination is not far away that it will not take long to get there, is a failure of rationality.

It follows that there is a certain distance between capacity-rationality and reasoning ability. Good reasoners can be habitually irrational, and, more commonly, perfectly rational people can be bad reasoners. They often make mistakes, but that does not impugn their rationality.[8]

[8] The point discussed in the text helps in dissolving an apparent asymmetry between theoretical and practical reasons. How is it, one may ask, that practical (and aesthetic) rea-

The third point was anticipated in the previous remarks, and helps illustrate them. Our rationality, I claimed, consists in the ability to recognise the normativity of features of the world. That ability expresses itself in the proper functioning, in relevant respects, of our faculties. The point I am striving to make is that our rationality expresses itself not only in our deliberation and reasoning, nor in any other specific act or activity, but more widely in the way we function, in so far as that functioning is, or should be, responsive to reasons. Take a simple example. I mentioned above that habitual failure to take decisions, or to form judgements where they are called for, is a form of irrationality. Imagine that whenever Sylvia leaves her home she locks her front door behind her, walks to the gate, turns back, goes to the door and, to check that it is really locked, unlocks and relocks it, goes out of the gate and immediately turns back, goes back to the door and repeats the action, and so on several times. This form of indecisiveness amounts to irrationality. At some level Sylvia probably recognises that her actions are irrational. But at the same time she is seized by anxiety and doubt. Maybe the key did not turn full circle and did not lock the door, maybe she imagined feeling the bolt move, and in fact it did not, maybe when she unlocked the door to test what she did before she forgot to relock it.

We are all familiar with mild forms of such anxiety. Barristers are supposed to capitalise on it by undermining witnesses' confidence that they really saw what they saw, that they really remember what they remember. At the same time we recognise that the capacity for such self-doubt is itself an aspect of our rationality. It manifests our ability to monitor (mostly below the level of awareness) our reactions and their appropriateness to the circumstances, that is to monitor our responsiveness to reason. But to act rationally we need to preserve a proper balance between resoluteness and openness to doubt. It is not a balance we can decide upon. While Sylvia can grit her teeth and decide not to turn back again for the fifth time, such conscious decisions only minimise her irrationality. To be rational she must act

sons are normative on their face, being facts like: 'Doing A will give you pleasure', or 'you promised to do A', whereas anything can be a reason for belief? That there are clouds and high winds is a reason to believe that it will rain. But there is nothing inherently normative in the fact that there are clouds and high winds. The explanation is that the fact that there are clouds is no reason to believe anything, though that it follows from the fact that there are clouds that rain is likely is a reason, and is normative on its face.

'automatically' in a way which allows for a proper openness to doubt without relapsing into indecisiveness and anxiety. To be rational she, and we, must function properly, and that functioning must be automatic, rather than a product of deliberation and decision.

The fourth and last point to be made here about rationality is by now obvious. I am treating rationality as a unified concept, designating a unified capacity, which straddles the divides between practical and theoretical rationality, as well as between procedural and substantive rationality, and others. Some writers believe that there are two concepts of rationality in use, and presumably they designate two different abilities. Parfit sets the distinction thus:

"To be substantively rational, we must care about certain things, such as our own well-being." "To be procedurally rational, we must deliberate in certain ways, but we are not required to have any particular desires or aims, such as concern for our own well-being."[9] He does not elaborate. What could he mean? It is possible for a person[10] irrationally to fail to have desires or intentions or goals which he has adequate reason to have. It is also possible for someone who generally has the goals which rationally he should have to display irrationality often when deliberating. There is no reason to think that failures of rationality are randomly distributed among the different occasions on which rationality is called for. For example, motivated irrationality, such as self-deception, would manifest itself selectively on occasions where its existence would serve its underlying motive. Parfit's distinction may, therefore, be understood as a distinction between success in being rational in different aspects of one's life. But the context makes clear that he does not mean it in this way. He seems to think that there are two different notions of rationality, each designating a different capacity. His discussion raises the possibility that possession of the rationality identified by one of these notions is independent of possession of the rationality marked by the other. Whether or not this is Parfit's meaning it is a common view, and a mistaken one.

The division between substantive and procedural rationality (and between substantive and instrumental rationality) took hold

[9] D. Parfit, "Reason and Motivation" *Aristotelian Society*, supp. Vol. 71 (1997) 99, at 101.

[10] By their nature persons are rational in the sense of having capacity-rationality.

among philosophers who doubted that reason is directly involved in the choice of ends, but believed that reason has a role to play in practical thought, which the notion of procedural rationality captures. Some allowed that we talk of rationality in a more extensive sense, and claimed that that is merely an ambiguity in the meaning of 'reason' and 'rationality', perhaps resulting from those primitive times when people believed that rationality affects the choice of ends as well. In fact reason affects our choice of ends and the desires we have just as much as it affects our deliberations and our beliefs. We cannot have a desire except for a reason.[11] Once that is allowed the motivation for the division of rationality into two distinct capacities disappears. There is no reason for thinking that the capacities which enable us to discern and respond to reasons for desires are different from those which enable us to discern and respond to reasons for belief. One may well need some special abilities (discriminating eyesight, or palate, a good ear, or a capacity for empathy etc.) to be able to discern various non-instrumental reasons. But these, while presupposed by capacity rationality, or by the ability to use it, are not identical with it, nor are they constituent elements of it. They do not show that there is more than one concept of rationality.

It is not clear what Parfit means to include in "to deliberate in a certain way". The power of reasoning in and of itself does not establish any degree of rationality. Imagine a person who as a hobby picks on arbitrary collections of propositions, perhaps chosen randomly from newspapers and magazines, and works out various other propositions which follow from them. Putting on one side the suspicion that the very activity displays irrationality, does he show himself to be good at procedural rationality? I do not think so. Our person is good at recognising the existence of inferential relations. He is not changing his beliefs in any way, not even conditionally (i.e. the story does not assume that he acquires beliefs of the sort: were I to believe these premises I would have a reason to accept these conclusions). For all we know he does not understand the normative significance of inferences. Perhaps procedural rationality is meant to include more than such reasoning. There may be no obstacle to enriching the notion to include steadfastness of resolution and other aspects of

[11] Though occasionally people have urges which are unreasoned. I have argued to this conclusion in "Incommensurability and Agency" in *Incommensurability, Incomparability, and Practical Reason* ed. R. Chang (Cambridge, Mass.: Harvard UP, 1998).

proper functioning of the kind alluded to above. The question is whether once enriched the capacities included under the heading of "procedural rationality" will be different from the capacities which make one so-called substantively rational.

To repeat: to those who believe that no rational capacities are involved in persons having goals the answer is obvious. But since we desire only what we think of as worth desiring, our desires are among our responses to perceived reasons. This is true of a desire to drink when thirsty as much as of the desire to become a good teacher. In conceiving desires and in adopting and maintaining goals we deploy all the capacities which are involved in so-called procedural rationality, and there seems to be no other capacity involved, at least none which can relate to rationality. The reasoning ability and other capacities which make people rational in forming beliefs about scientific matters, or about the weather, or anything else which can be said not to be in itself normative, are the same abilities which make people rational in the way they adopt and maintain goals. Therefore, there is only one kind of rationality.

As was allowed above there may be factors which may interfere with the display of rationality in one area but not in another. But these will not be enough to establish complete independence of one's success in being rational in the different domains, and they will not track the supposed distinction between procedural and substantive rationality. Typical examples of such selective failures of rationality are those occasioned by motivated irrationality. They affect one's thoughts, feelings and goals only when the motive leading to the irrationality comes into play. But they can affect one's choice of goals as well as one's 'theoretical' beliefs.

These remarks on the nature of rationality and its relation to normativity form the backdrop to the discussion that follows. They concern capacity-rationality, and though they inevitably have implications regarding the meaning of judgements that this or that action, desire, emotion, attitude, belief, etc. is rational or irrational, these implications are not straightforward. As they do not affect the rest of the argument they need not be explored here, beyond one point. Obviously to judge a belief, desire, emotion etc. as rational is to note that having them is at the very least consistent with a successful deployment of our capacity for rationality.[12] The standard by which success is to be measured is far

[12] "This was a rational belief for X to have" may mean no more than that. "He rationally came to the belief that …" indicates much more. It presupposes actual use of one's

from clear. It is doubtful that there is only one standard employed on all occasions. It seems likely that we recognise a range of standards, and on each occasion we implicitly employ the one we find fitting in the circumstances. The most demanding standard regards as rational only those beliefs, etc. which are consistent with complete identification of all the reasons which apply to the situation[13] and a perfectly appropriate response to them. Everything else is judged irrational.[14] This standard seems to be used only rarely. A very lenient standard regards any belief to be rational unless holding it results from a failure (successfully) to employ one's rational capacity which involves gross mistakes, etc. Some standards relativise to one's age (what would be rational for a child to believe may be irrational for an adult), educational background (given that you have a degree in physics it is irrational for you to believe that), position in life, or to some other social variable. Common are standards which tie irrationality to blame. One's beliefs are irrational when one is blameworthy for having them. The best of those employ something like the legal test of negligence: a belief is irrational if and only if holding it displays lack of care and diligence in one's epistemic conduct.

2. The centrality of the ability to reason

When studying reasons we study normative aspects of the world. When discussing rationality we discuss our perceptions of, and responses to, reasons. Our ability to reason is central to our rationality in all its manifestations, that is regarding reasons for belief, action, emotion, or anything else.

One way to bring out the point is by reminding ourselves that emotions, attitudes, desires, and intentions have a cognitive content and cognitive presuppositions. Their rationality depends, in part, on the rationality of the beliefs which contribute to them, and which are presupposed by them. This dependence is asymmetric. Beliefs do not depend on feelings, desires or intentions in

rational capacity, and asserts that that has been successful. There are many variations and nuances in attributions of rationality and irrationality to beliefs, etc.

[13] Even this demanding standard allows that epistemic reasons vary with context. Therefore, even according to it one's beliefs may be rational and false.

[14] There is no reason for the standard for irrationality to be the contradictory of the standard for being rational. The two can be logically independent, allowing for beliefs which are neither rational nor irrational.

the same manner. The rationality of beliefs enjoys a certain primacy for being involved with other forms of rationality, which it does not presuppose.

More specifically, at least a rudimentary reasoning ability is involved in all rational responses to reason, simply because they are responses under the control of the agent, though admittedly I am using "responses" loosely here, to indicate that rationality depends on appropriateness or intelligibility in the circumstances. Emotions, desires, intentions, or beliefs are rational depending on (a) whether they belong to a rational agent; (b) whether their occurrence is under the control of the agent; and (c) whether they are appropriate or intelligible given the reasons for and against them, as these reasons are, or as they are reasonably perceived by the agent.[15]

It is worth stopping to comment on the second of these conditions. It embodies several of the points belaboured in the previous section. I will address my remarks to the case of rationality in the endorsement of beliefs, but they apply with minor modifications to rationality in one's intentions, emotions, decisions, actions, etc. Control consists in the proper functioning of a person's rational faculty, the proper functioning of the person's ability to recognise and respond to reasons, rather than in any particular performance, such as an action, or a deliberate decision. Control is manifested when a belief is adopted, or endorsed in a process in which the ability to recognise reasons and respond to them (proper appreciation of and response to perception, or to testimony, for example) is active. It can be active even when beliefs are formed without deliberation or awareness, but when the agent's critical faculties would have stopped their formation had they been rationally suspect.

Is it not a distortion to claim, as I just did, that the proper functioning of a faculty, even of rationality, depends on being in control of our emotions, actions, beliefs, and the will? After all the central use of "control" relates to exercise of the agent's will: we control what we do, and how we do it, and we control our emotions, to the extent that we intentionally hold our emotions in check by an effort of will, etc. But there is a wider, more basic, use of control. People's being out of control means that their will itself is not under their control, and our will and

[15] This point is subject to the baseline question discussed above.

beliefs are out of control when they are systematically irrational.

Capacity rationality is a more fundamental capacity than the will, which is the capacity for intentional action, for forming intentions and taking decisions. Rationality is like dispositional abilities, that is abilities which are manifested when the circumstances are right. (E.g., a rope which can take a 100 kg weight is a rope which will take weights up to 100 kg. without snapping, unless . . .). Our fundamental psychological abilities are of this kind, except that they are subject to complex possible interferences, many of them due to psychological factors: memory can fail when people are depressed, etc. Rationality is like that: it is the ability to respond appropriately to (perceived) normative aspects of the world, and this means that rational beings respond appropriately to perceived normative aspects of the world, when no failure of attention, emotional upset, mood, memory, will, etc. interferes.

Like other dispositional capacities rationality is a capacity which displays itself when the occasion presents itself, so long as no distorting factors interfere. That is what the second condition for the rationality of beliefs, actions, etc. signifies. People's beliefs (and again these remarks apply with some changes to other objects of our rationality) are rational only if they are formed and maintained while the people involved are in control of their formation, and continued endorsement. That is, while their reason controls how they come to adopt or endorse their beliefs. Reason controls the formation and endorsement of beliefs when, whether or not their formation or endorsement involves deliberations, beliefs are formed in processes which stop people from having them when their formation or endorsement is not warranted by reasons, as the agents see them, given their understanding of the situation they are in. A different aspect of control (and all these are matters of degree) is exercised when reason makes people endorse propositions which they are aware of compelling or at least adequate reasons to endorse.

Some people will say that when reason is in control we engage in subconscious reasoning. Whether or not such claims can be vindicated, the exercise of our rationality must be represented in the same way that reasoning is represented. The exercise of reason which manifests its control over our beliefs, emotions, intentions, desires, etc. is subject to the same rules that govern explicit reasoning. To that extent capacity for reasoning is central to

rationality and is involved in all its manifestations. This raises the question of the standing of principles of reasoning.

3. Can reason be justified?

In setting out (at the beginning of this essay) the contours of the problems of explaining the nature of normativity I did not once refer to the justification of normativity. Is not that the primary task of a theory of normativity or of reason?

It is not easy to make sense of the very quest for the justification of normativity. We can ask whether this fact or that is a cogent reason for action or belief, etc. We can raise more general questions about types of facts: For example: does the law (i.e. the fact that one is legally required to perform an action) constitute a binding reason for action? Do people have good reason to conform to the practices of their country? But what is it to justify reason as such? Presumably the question is whether we are ever justified in holding anything as constituting an "objective" reason? Or, whether it is possible for anything to be a reason? Or, whether there are any facts which are reasons?[16]

These explanations of the quest for the justification of normativity do not, however, dispose altogether of the difficulty in understanding it. Even once anxieties about the objectivity of reasons are put on one side questions remain. They are not dissipated by looking for the justification of statements that this or that is a reason for action or belief. Such justifications take many shapes and forms: burning the cat would be cruel (and that is a reason not to burn it) because it would inflict gratuitous suffering, and so on. These are the mundane arguments for this reason or that, which we are all familiar with. The quest for the justification of normativity cannot be whittled down to the normal arguments for the truth of a statement about this reason or that. It is a search for the vindication of the methods of reasoning employed in such mundane arguments, or for the discovery of a super principle which justifies confidence in the whole enterprise of reason, the whole enterprise of discerning reasons and responding to them. Is that a meaningful and a sensible quest?

[16] Alternatively, perhaps the question meant is whether it makes sense to talk of anything being a reason? But this question comes very close to the explanatory questions I mentioned in the previous section. They explore what is the sense of normative discourse. I for one find little reason to doubt that normative discourse is meaningful.

One common reply is that every argument to debunk reason would be self-defeating, for it will have to use reason and thus its own validity depends on the assumption that it seeks to challenge, i.e. that reason is justified. Recently Tom Nagel has advanced a whole array of arguments in support of the objectivity, universality and reality of reason (these are his terms) including a version of the argument that challenges to reason are self-defeating.[17] Nagel's argument rests on two pillars. First that "one cannot criticise something with nothing".[18] The second is the fact that we cannot escape relying on reason.[19] Both pillars are suspect.

Reductio ad absurdum is a familiar form of argument which, at least prima facie, need not presuppose anything. It refutes a supposition by deriving a contradiction from it. As presented *reductio* arguments quite commonly presuppose premises and rules or methods of argumentation, which are accepted as uncontroversial. But need this be the case? Is there some reason why one must rely on some premises, rules or methods of argument other than those which are refuted by the *reductio*? I do not think so. In a way no *reductio* argument does. They all take the following form: Using rules (or methods) of argument R_1 to R_n a contradiction can be derived from premises P_1 to P_m. Therefore, at least one of the premises or at least one of the rules of inference is false or invalid. As I said, commonly in such arguments all but one or a few premises are accepted as being true, hence the conclusion is that at least one of the remaining ones is false. But strictly speaking such arguments do no more than impugn one of the premises and rules of inference. Of course, *reductio* arguments use rules of inference, not least in their final step: All these premises and rules yield a contradiction. So at least one of the premises is false or at least one of the rules invalid. That does not matter to my point, since the rules of inference relied upon are themselves put in doubt by the argument. Nothing is relied upon without being cast into doubt at the end.

[17] In chapters 2 and 4 of *The Last Word* (New York: O.U.P. 1996)

[18] *Ibid.* p. 20.

[19] Nagel is, of course, aware of the fact that the inescapability of a belief in the objectivity of reason is no proof of its correctness: see. p. 33. He rests his case on the claim that no sceptical conclusion follows from the fact that explanation must come to an end, and that "the language, and the truth of some other form of subjectivism is not shown by the fact that justification comes to an end at certain points at which there is natural agreement in judgements. Nothing about the framework of thought is shown by these facts". (p. 34). I agree with these propositions. They are consistent with the less sanguine view expressed in the main text.

To use *reductio* to challenge the validity of reason one would reduce the number of premises to the minimum, including only logical truths or indubitable other truths, and use only the most basic rules of inference. If *they* give rise to a contradiction then it follows that at least one of them is false or invalid. Would not such a conclusion be tantamount to a sceptical argument against reason itself? For example suppose that using substitution and modus ponens only, one derives a contradiction from (x) x = x. Could not such a proof lay claim to being a refutation of the cogency of reason?

Moreover, if we reject Nagel's claim that "one cannot criticise something with nothing", then the fact that we cannot escape relying on reason may not be a very powerful point in its defence. It may turn out to be the case that while we cannot avoid relying on reason, we know that we are doomed to rely on an incoherent system of thought.

From the fact that if there is such a radical *reductio* argument it will refute the validity of reason and the value of rationality it does not follow that they stand refuted. One can perhaps conclude (a) that Nagel, and others, have failed to produce general reasons for the impossibility of a successful sceptical argument against reason, and yet hold (b) that no such argument exists, and therefore that we have no reason to doubt the cogency of reason. But that view does not seem to be quite adequate to the situation.

First, it is not entirely true to say that no challenges of the kind described exist. Paradoxes, some old some new, such as the liar paradox, or Zeno's paradoxes, or the sorites paradox, have puzzled thinkers. Such paradoxes seem to be of the required kind. If not singly then cumulatively they challenge the coherence of reason. They take assumptions that lie at the very foundation of our conceptual thought and reduce them to a contradiction. They challenge the coherence of the concept of truth, the concepts of change and of time, and of the concepts of identity, of objects and of possessing properties. It will take a brave man to say that they were all solved successfully. But we need not argue about that. What is of interest is that for long stretches of time people did not know how to solve them, and knew that they did not know how to solve them. Yet they carried on regardless. Rational thought did not stop, was not abandoned, in spite of awareness that, for all one knew, there were contradictions at its foundations. I do not know of any

serious, let alone successful, argument that that was irrational, that it was irrational of people to carry on using reason, in spite of being aware of unresolved paradoxes concerning its basic features.

How are we to understand this reaction? I am less interested in its historical explanation than in the question whether it is a rational reaction on general grounds, independent of the specific historical context. One possible explanation is that even if the use of reason is incoherent and self-contradictory, we are condemned to carry on using it. To be sure we can abandon reason, but we cannot reason our way into doing so. We abandon reason, or it abandons us, when we suffer brain damage, stupefy ourselves with hallucinatory and other drugs, etc. People can take action deliberately in order to be rational no more, but they cannot get there simply by reasoning their way into scepticism about reason. But that answer is unsatisfactory. If we can abandon reason, or cause it to abandon us, why is it not the case that exposing the contradictions at its foundations constitutes an argument for doing so? One cannot answer that such an argument relies on reason. It relies only on the fact that reason is self-defeating and self-contradictory.

One may point out that the step from the proposition that reason is self-defeating, and self-contradictory, or simply that it is unjustified, to "let's blow our brains out" relies on the rationality that has just been defeated. Perhaps the bankruptcy of reason cannot be a reason for abandoning reason. But at least it means that there is no reason not to do so. We cannot conclude that we may do this if we want to, if that means that we have reason to believe that abandoning reason is permissible. But we can hold that we may do so, meaning simply that there is nothing against doing so – no reason against doing so, since there are no reasons for anything. That conclusion is devastating enough. It does not leave things in equilibrium. It is not neutral as between abandoning reason and not doing so. For, if the debunking arguments are successful, then to choose to continue to rely on reason is either to choose to be self-deluding, or to choose arbitrarily with open eyes to follow a debunked mode of life. It is to act against the spirit of the action. It is to rely on a method when by the light of that method itself it should not be relied upon. That does not establish a reason for not relying on it, but it places reliance psychologically at a disadvantage.

4. Arguing against substantive principles of reason[20]

We cannot reason ourselves into abandoning faith in reason not because it is impossible for us to abandon faith in reason as a result of a successful argument that we should do so, but because no such argument can be sound. *Reductio* arguments of the kind I mentioned do not constitute a *reductio* of rationality. It is useful here to distinguish between the formal notion of reason, and substantive doctrines about the nature of reason. The formal notion of reason is fixed by the very abstract and essential characteristics that mark the kind of thinking which is governed by reason. We have no word for it other than "thinking". It is, however, thinking in a narrow sense, the sort of thinking that we refer to when saying: "wait a minute. I am trying to think", and not the sort of thinking that day dreaming, free association, fantasising, etc. are.[21] The formal notion of reason singles out a type of thinking which is marked by the discipline it is subjected to, a discipline which enables one to distinguish instances of successful thinking, from flawed thinking. Possibly thinking in the narrow sense can be adequately characterised by two central essential properties:

(a) It is thinking which is subject to evaluation as correct or incorrect.

(b) The standards by which success of episodes of thinking is judged depend on the reliability of the process of thinking which meets them in yielding justified intentions, decisions and beliefs, that is ones which are adequate, given the normative aspects of the world.[22]

In short reason is a discipline which governs thinking, or a type of thinking. Thinking in conformity with reason is successful thinking, and thinking which does not conform to reason is unsuccessful as

[20] The considerations advanced below are presented informally. I assume an interdependence of words, meanings, and concepts, but do not clarify it. Nor do I try to make more precise the notion of a principle of reasoning. Given a generous understanding of the notion not all principles of reasoning are constitutive of meanings or concepts. That does not affect the case put here which is sound as applied to those principles of reasoning which are constitutive of meanings and concepts. If anything, the contingency of other principles of reasoning is even easier to establish, but I will not consider them here.

[21] Though other forms of thinking may borrow parts of the discipline of reason which marks the narrow notion of thinking.

[22] Notice that not all thinking aims at justifying beliefs or intentions. It can be part of the telling of invented stories, fanciful imagining, etc. It is, however, thinking of the relevant kind if it is governed by standards the use of which can justify intentions and beliefs.

an instance of thinking of that kind. The substantive doctrine of reason spells out the content of that discipline. Rules of inference (deductive and non-deductive alike) and the central concepts they depend on are a central part of the substantive doctrine of reason.

I will argue that the paradoxes of reason cast doubt on the substantive doctrine of reason, on the cogency of the concepts and rules of inference it employs, not on the possibility of reason in itself. That is they do not impugn formal reason. Do they cast doubt on the way we reason? Or on our understanding of the way we reason, on our theory of reason? Does the substantive doctrine of reason I referred to consist of the principles which govern our reason or our attempts at an explicit articulation of those principles?

They may do either. In any case the two are not entirely separate. Once we develop explicit accounts of the nature of reason, of its substantive doctrines, those accounts, while they never exhaust the forms of reasoning we engage in, do affect how those who are more or less aware of them and accept them, reason. Solutions to the liar paradox cannot be said simply to improve our understanding of how we reason all along. They change our reasoning practices, modify the rules used in them, or restrict their application. And the same goes for solutions to most other logical paradoxes. This is particularly likely to be true of the ancient paradoxes. Regarding them it is unlikely that the persistent failure of attempts to solve them is due merely to misunderstanding of how we actually reason, and therefore unlikely that solutions point merely to misunderstandings rather than to the need for a change in the concepts which generate the paradoxes.

If I am right then reason, i.e. the doctrines of reason, can be successfully challenged, and we respond to such challenges by modifying it, modifying reasoning practices and the principles which govern them. I am not suggesting that they are modified only in response to paradoxes and sceptical challenges. However, so long as they do not bring new paradoxes in their wake, modifications which free our practices from paradoxes constitute advances or improvement in our reasoning practices, and in the principles of reason we use.

On this view reasoning principles are social principles, evolving roughly in the ways in which social practices generally evolve. But they can also be challenged on grounds of incoherence, or

unreliability. They can be improved in response to such challenges. In this picture, changes in logical and conceptual principles of reasoning parallel changes in inductive clues and scientific methods of experimentation. Even though changes in principles of reasoning involving conceptual shifts occur mostly in informal ways and are harder to document, they respond to pressures similar to, though more diffuse than, those affecting scientific methods.

One reason for viewing with suspicion the claim that reasoning principles are based on practices which evolve in ways similar to that of other practices which govern our life is the familiar philosophical doctrine that regards changes in reasoning practices as nothing but corrections of mistakes. Such corrections occur when the changed practices comply more closely with the universal principles of reason. This objection admits that not only the explicit articulations of principles of reasoning can be faulted and can change, but that actual reasoning practices can be affected by contradictions and paradoxes, and can change to avoid them. But the objection denies that that process should be regarded as a process of change in *the substantive principles of reason*. They are universal and timeless. The changes we observe are in our imperfect attempts to conform to them.

The objection presupposes, however, that there is one and only one set of correct or valid principles of reason. This seems to me implausible, or at least in need of qualification. Principles of reasoning and of meaning and reference are, of course, intimately related (i.e. if we abandon a principle of reason – e.g., the excluded middle – we change the meaning or content of the propositions which are governed by these principles of reasoning). Therefore, if there can be systems of concepts such that (a) none of them is better than any of the others, and (b) there is no possible system of concepts and reasoning which is better than they are, then there is more than one ideal or correct system of concepts and reasoning, ideal in that they cannot be improved upon. Systems meeting these conditions also meet the further condition, namely that each includes concepts which are not part of the other. Therefore, such systems are incommensurate.[23]

[23] My invocation of meaning incommensurability carries no sceptical implications. I do not claim that those who understand a system of concepts and reasoning (or, for that matter, a theory) cannot understand others which are incommensurate with it. Nor do I believe that incommensurability implies the possibility of incompatible truths, each vindicated within its own system or language, and each refuting the other. The possibility of

The mere possibility of incommensurability among systems of reasoning (expressed in different languages or segments of languages) does not cast doubt on the universality and timelessness of the principles of reasoning. However, once that possibility is admitted it becomes difficult to resist the thought that there is indeed an indefinite number of incommensurate systems of reasoning, and incommensurate languages or segments of languages expressing them, each with its own ideal or correct principles of reason. The number and identity of the historically instantiated systems of concepts, and of the rules of inference and reasoning associated with them, are matters of historical contingency. Moreover, it is possible for more than one correct system to be instantiated. That makes it hard to deny not only that the historical instantiation of principle of reasoning is a matter of contingent fact, but that the principles themselves are historical products emerging at particular points in time. The alternative is to assume the existence of an infinite number of sets of principles of reason, most of which it is impossible for us to find out about until such time, if ever, when the related new language evolves.

Let it be granted that ideal or correct principles of reason are historical products, that is, practices which arise in time: must we also concede that practices riddled with paradoxes embody principles of reason, rather than that they embody mistaken principles which are accepted as if they were principles of reason, but which are not? Admittedly it is puzzling to think that there can be paradox-ridden principles of reason. Adding that they are imperfect or defective principles of reason, rather than removing the oddity of this view, adds to it. Is it not a contradiction in terms to think of imperfect or defective principles of reason?

Nevertheless, we must accept that this is so, for the alternative is unacceptable. The first and most radical way of understanding the situation of people whose practices of reasoning are infected by paradoxes has to be rejected. It would be wrong to say that the people whose practices of reasoning were riddled with paradoxes did not have any principles of reasoning enabling them to distinguish cogent from erroneous arguments, and rightly guiding

incommensurate systems of concepts presupposes that truths that can be stated in one are compatible with truths which can only be stated in the other, though the tests of compatibility may involve extending the range of concepts in one or both to make sense of the very notion of compatibility and incompatibility between incommensurate propositions.

them in their thinking. It is evident that they engaged in thinking guided by reason just as much as anyone else does.

A more moderate understanding of their situation will have it that their practices of reasoning were mistaken, but had limited validity in being imperfect approximations of the sound principles of reason. On this view we can say that they were, by and large, guided by the sound principles of reason, except that their practices failed to incorporate and follow them adequately, thus leading them on occasion to incorrect applications of the right principles. Apart from other weaknesses, this way of understanding their situation presupposes that there is a unique set of sound principles of reason, whereas – as was suggested above – there seem to be indefinitely many incommensurate sound sets of principles of reason.

The most promising suggestion is a modification of the second proposal, to allow for a plurality of incommensurate sets of principles of reason. It says that practices of reasoning infected by paradox can rightly be regarded as mistaken approximations of some sound principle of reason or other. Being such approximations we can say, as in the second proposal, that the people who followed them could be in part vindicated. They can be said to have been imperfectly guided by the sound principles of reason which their practices approximate. This suggestion purchases the advantage of not allowing that principles of reason can themselves be imperfect at too easy a price. In the absence of further criteria, and it is not clear what they might be, any practice of thinking approximates some set of principles of reason or other. This may not be an objection to the suggestion we are examining. Possibly no practices which allow one to distinguish correct from mistaken thoughts, episodes of thinking, or transitions of thought, can be discounted. All of them are practices of reasoning in the minimal sense we are exploring, i.e. they approximate some sound principles of reason, and their practitioners can be said to be guided by the sound principles these practices approximate.

Nevertheless, I find this suggestion unacceptable for it seems that any imperfect (i.e. paradox infected) practices of reasoning approximate not one but many sound principles of reason, many of which are unknown to us or to the practitioners, and, given our time and place, many are unknowable by us or by them. In these circumstances, denying that people are governed and guided by the principles they have, and insisting that they are really

guided by principles beyond these, principles which they do not know, and perhaps cannot know, is an unhelpful verbal trick to avoid saying that principles of reason can be imperfect. The suggestion serves no other purpose. The "more honest" course is to keep the connection between principles people follow and the practices they engage in. This connection allows us to explain how people can be guided by principles which in some sense they do not know, i.e. of which they are not (fully) aware, and the content of which they cannot articulate. The price of allowing that principles of reason can be imperfect seems no price at all.

The preceding argument for the historical character of principles of reason presupposes meaning incommensurability. Is the presupposition justified? The issue is complex and it is easy to argue on the one hand for, and on the other hand against meaning incommensurability. Not everything which can be expressed in one language can be expressed in all others without extending their resources, either by adding new words, or phrases, or enriching their grammar. Once, however, enrichment is allowed what is there to stop a language from being enriched to whatever degree may be necessary to enable it to express whatever the other does, and to include all the concepts of the other? Can it be enriched by simply absorbing the other language as a part? Is not that the way the "languages of science" came to enrich the standard natural languages which, because of the existence of such specialised segments, can now express what they could not express without them?

If this is where the argument resides then I suspect that there is no 'principled' solution to this problem. That is, it cannot be solved by arguments which disregard the contingent and historical nature of languages (both natural languages, and specialised segments of them like the languages of science or of law). The issue of incommensurability of meaning turns out to be the problem of the identity of languages, and that is a historical matter. Whether something is a segment of another language or a separate language can depend not on relations of meanings and of rules of grammar, but on how the two developed and who uses them and when. But if the issue of meaning incommensurability is one of historical development then meaning incommensurability is possible and therefore principles of reasoning are themselves historical products.

I will, therefore, proceed on the assumption that the substantive principles of reason are historical products which can be

challenged on grounds of self-contradiction, incoherence, and unreliability, and which can change to avoid such challenges. This allows scepticism more scope than Nagel allows it. It does not, however, allow for an attack on reason or rationality itself. Such an attack would have to be addressed to the formal concept of reason, and show that it is incoherent.

Before we consider this problem, one last word regarding critiques of substantive doctrines of reason. The preceding discussion aimed to show that such doctrines are capable of being criticised. When paradoxes are discovered principles are revised to avoid them. This in itself does not constitute a justification of those principles which are not infected by paradox. Does one not need to provide them with some justification other than the timid response "so far no paradoxes affecting them have come to light"? Yes, and no. It is certainly possible to explain why such principles are valid. But the explanation is not of a kind normally thought of as justification. Rather the explanation will relate to the constitutive role of such principles, reconciling the fact that they are constitutive of a mode of thinking with the possibility of alternative, incommensurate, sets of principles, and noting that the validity of the principles does not guarantee that whatever thoughts they are manifested in are about "an independent, objective reality."[24]

5. The standing of formal reason

Is it possible for a sceptical argument challenging the coherence of the formal notion of reason, or the case for its use, to be successful? Many of the attacks on reason and rationality which we witnessed in the course of history are meaningful and arguably justified. Such, for example, are attacks on rationalism claiming either that people have come to exaggerate what can be achieved simply by conformity with principles of reason, or that the imagination and fantasy, or feelings and emotions, are of value, as well as thought in the narrow sense. None of these nor other similar claims need be disputed. Nor need we engage in an argument about the instrumental importance of rational thought, e.g., whether people or the human race generally can survive for long without it. The question is whether it is possible to prove that

[24] See my "Notes on Value and Objectivity ".

rational thought is inherently bankrupt. If it has instrumental value, the debunking argument goes, it has it in spite of its failure to live up to its aspiration to be a reliable guide to truth. Can one show that rational thought as such, not merely this or that substantive doctrine of rational thought, is incoherent?

Two possible routes towards this conclusion suggest themselves. Sceptical argument following the first route would prove that no account of the content of reason can escape a *reductio* argument showing it to be incoherent. While the refutation of any specific account of reason has to proceed via a *reductio* argument, an argument that any account of reason is subject to a successful refutation need not itself employ a *reductio* method of argumentation. Any form of proof would do.[25] The second line of attack would consist of arguments showing that the very formal notion of reason is incoherent. I do not believe that we have been confronted by any such arguments. But it may be useful to speculate on the effect they may, if successful, have.

Think first of the impact of the paradoxes affecting the substantive doctrines of reasoning. Their solutions did not lead to wholesale abandonment of rational concepts or principles of reasoning. Rather, they led to modifications of existing principles and practices of reasoning. This is what one would expect if one believes that all adjustments forced by arguments would tend to be the minimal necessary to satisfy the reasons forcing them. In terms of familiar metaphors: we start where we are and we proceed to adjust our position from there. We move not towards some unique ideal but towards a system of concepts and of reasoning which is readily reachable from our starting points and which seems to avoid the difficulties which force us to move from our initial position. It seems reasonable to assume that more radical paradoxes affecting the cogency of the formal concept of reason will, if they ever materialise, lead to similar partial adjustments.

It is true that the strategies of challenging formal reason seem to allow a more far-reaching conclusion. They seem to allow the conclusion that no principles of reason whatsoever can be valid. But appearances are misleading. To entertain that thought is to entertain the thought that propositions and concepts are discrete with no logical or conceptual relations between them. That

[25] Pyrrhonist sceptical arguments are of this kind, as they aim to show that if a proposition is supported by reason so is its negation.

would make mastering any of them impossible. Hence the most that a challenge to formal reason can succeed in doing is forcing adjustments to the concept of formal reason.

There is, however, a difficulty in conceiving of such an adjustment as anything other than a correction of a mistake regarding the one and only notion of reason. When considering the principles of reason my suggestion was that the pressure of paradoxes leads one to move from one's initial system of concepts and of principles of reasoning, through the smallest adjustment which happens to suggest itself, to what is strictly speaking an incommensurable alternative system. Both systems are instantiations of the formal notion of reason, though the second is (so far as we know) superior to the first in not being infected by paradoxes. When we abandon, through some adjustment, our formal notion of reason such a picture is impossible to sustain, for there is no genus which covers both the earlier and the later formal notions of reason. It seems as if strictly speaking we must regard the previous notion as a mistaken (because affected by paradoxes) version of the second. That is, we must assume that there is but one, though possibly not yet correctly understood, notion of reason.

The imagined radical paradoxes do not lead to a change in the concept of reason, but to the correction of mistaken beliefs about its nature. We cannot distance ourselves from the formal concept of reason, and adopt another in its stead. This enterprise does not make sense. This conclusion should not be misinterpreted. It is not that we are captives of "our" concept of reason and cannot abandon it even if it is mistaken. We can so to speak "abandon it", though what we abandon is not the concept of reason but the views we happen to hold about its nature.[26] These views are what sceptical arguments can hope to refute. The very possibility of formal reason cannot be refuted for the notion of such a refutation does not make sense. It assumes either that concepts are discrete with no conceptual relations between them, or that there can be an alternative to formal reason, and that is nonsense.[27]

[26] We can of course "abandon" it in the trivial sense of ceasing to reflect about it.

[27] We can of course become creatures which cannot think (in the narrow sense of that word). That, however, is a fact about us not about the concept of reason.

6. Conclusion

Let me take stock. We saw (in section 3) that the thought that normativity as such should be defended and justified often amounts to a demand for the justification of reason. It encompasses questions I did not touch on; primarily doubts about the objectivity of reason, and about the possibility that people may be motivated by reason.[28] Putting such doubts to one side we saw that it is possible to advance sceptical arguments against any of the principles of reason. Such principles are historical products in the same way that languages and other systems of concepts are. They can be replaced by others which are, hopefully, free of paradox. Such changes are normally achieved not through wholesale rejection of principles of reasoning and of the systems of concepts which gave rise to them, but through their adjustment and modification. There cannot, however, be sceptical arguments against reason itself. For while it is possible for human beings to stop engaging in thinking, in the narrow sense of the word, and even to lose the ability to do so[29], it makes no sense to think that the concept of thought or of reason can be rejected or be found defective. Our understanding of it can be found wanting. It can change, but what can change is the understanding of a concept which remains the same.[30] It follows that the validity of specific normative principles can be called into question. And that we – human beings – can become creatures incapable of being guided by normative considerations. But so long as we are capable of rational thought we are capable of being normatively guided, and, while we can explain the nature of reason and normativity, there is no such enterprise as justifying normativity.

[28] See Joseph Raz, *Engaging Reason* (Oxford: OUP, 1999) chs. 5 and 6.

[29] It is plausible to suppose that we can stop thinking altogether only if we lose the ability to think.

[30] To avoid doubt let me add that throughout history the methods of reasoning employed by people have changed, hopefully improved, in many ways which have nothing to do with paradoxes, and sceptical arguments of any kind. Such improvements can happen as people's knowledge and understanding of the world they live in increases. The discussion above does not bear on such developments.