

COPYRIGHT NOTICE:

Jon Elster: Reason and Rationality

is published by Princeton University Press and copyrighted, © 2009, by Princeton University Press. All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher, except for reading and browsing via the World Wide Web. Users are not permitted to mount this file on any network servers.

Follow links for Class Use and other Permissions. For more information send email to: permissions@press.princeton.edu



In analytical approaches to human behaviors, the same Latin word, *ratio*, is at the root of two intellectual traditions that are at once very different and interconnected.

On the one hand, there is the tradition that opposes reason to the passions and, more recently, to interests. Seneca's treatise *On Anger*, for instance, is organized around the opposition between reason and passion, whereas the French moralists of the seventeenth century added the notion of self-interest. La Bruyère, in a famous

passage, summed up their mutual relationships this way: “Nothing is easier for passion than to overcome reason, but its greatest triumph is to conquer a man’s own interest.”¹ The idea of reason is intimately connected to that of the common good.

On the other hand, there is the still more recent idea of rational choice, which is opposed to the diverse forms of irrationality. The rational actor is one who acts for sufficient reasons. These reasons are the beliefs and desires in light of which the action appears to be appropriate in a sense that I shall discuss at length. The idea of rationality is often but wrongly related to that of the actor’s private good or self-interest in

¹La Bruyère (2007), p. 98.

the moralists' sense. Anyone who is pursuing the common good can—and even ought to—do so in a rational manner.

Acting in conformity with reason, in the singular, and acting for good reasons, in the plural, are two different things insofar as reason is objective, whereas reasons are subjective. From an external point of view, we can evaluate a policy as being in conformity with reason or not. From an internal point of view, one can evaluate an action as being rational or not.² From this difference it follows that only rationality can be used for explanatory ends. It is only insofar as the agent has made the demands of reason his own that the latter may give rise to, and possibly explain, specific behaviors. The

² For this distinction, see Williams (1981).

assessment of the actor and that of the observer need not coincide.

Although they are different, the two norms encounter a common obstacle, namely, the passions.³ They also have a common com-

³ Although it is quite classical, various objections to this opposition have recently been raised. Concerning the relationship between the emotions and rationality, see in particular Damasio (1994) according to whom individuals suffering from a brain lesion have both deficient rationality and a reduced ability to feel emotions. But it does not seem proven that this is a causal connection rather than a simple correlation. On the relationship between the emotions and reason, see especially Rudenfeld (2001), according to whom reason needs the help of passion in order to overcome self-interest. This was also Clermont-Tonnere's view in the Constituent Assembly of 1789: "Anarchy is a dreadful but necessary stage, and it is the only time when one can arrive at a new order of things. Uniform measures would not be taken in calm times" (*AP* 9:461). The night of August 4th has led some to draw the same conclusion. The debates of the Constituent Assembly nonetheless include ambiguities that seriously complicate matters; one has to agree with Toqueville, for

ponent, which is the idea of acting in accord with well-founded beliefs. Finally, they have in common the fact that they are the object of a certain deference on the part of the actor. The origin and nature of this deference are not the same, but in both cases it is a matter of deference with regard to a source of normativity.⁴ The operation of mechanisms of deference is complex. For the moment, let it suffice to say that their effect is sometimes to subvert the object of deference.

It might be objected that comparing a principle concerning normative political

example, when he writes that “the night of August 4th [. . .] was the combined product, but in doses that cannot be determined with precision, of fear and enthusiasm” (2004b, p. 593).

⁴I borrow this phrase from Korsgaard (1996), who uses it, however, in a narrower sense.

philosophy with another that concerns the explanation of individual behavior is wrongheaded. One modest but sufficient reply to this objection would be to say that considering the usual confusions on this subject, conceptual clarification is worth pursuing for its own sake.

More ambitiously, I shall reply that clarification also has its place in political debate. Is it true, is it coherent, to say that the common good can be realized only through the pursuit of private goods? Is it true that the more rational actors are, the better reason's demands are met? Or must we see, inversely, the rationality of individuals as an obstacle to reason? Take, for example, the "voter's paradox," which results from the fact that the rational actor has no rea-

son to vote.⁵ In fact, the chance of having an influence on the outcome of the election is clearly less than the risk of dying in a traffic accident on the way to the polls. Moreover, those who are in the best position to understand the logic of this line of reasoning—in particular, professional economists—choose the cooperative strategy less often in the “prisoner’s dilemma,” of which voting is a classical example.⁶

Whereas the theory of rational choice has been elaborated and developed with great precision, the same cannot be said of the idea of reason. The conception that I am going to propose is not based on a ca-

⁵ See Blais (2000).

⁶ Frank, Gilovich, and Regan (1993).

nonical definition, because there is none. It represents a personal—but, I hope, not too idiosyncratic—synthesis of classical texts.

Let us begin with a remark of La Bruyère's: "To think only of oneself and of the present time is a source of error in politics."⁷ To correct this error, we have to consider both other people and the future. More precisely, we must substitute an impartial attitude for the partial perspectives constituted by egoism and myopia.

The idea that reason requires an impartial treatment of individuals corresponds to well-known principles. To resolve the questions of distributive justice, Leibniz proposes the following maxim: "Put your-

⁷ La Bruyère (2007), p. 358.

self in the place of everyone.”⁸ In recent theories, this amounts to saying that the choice of a just organization of society must take place behind a “veil of ignorance,” an idea that can be interpreted in several ways.⁹ For utilitarianism, each individual must count as one, and none as more than one. For John Rawls, we have to choose the form of society that favors the least advantaged, whoever they might be. Another impartial idea is that of universal rights, embodied in the two declarations of 1776 and 1789.

Less emphasis has been put on the idea, which is just as important, that reason requires impartial treatment of temporal instants. In itself, no date can be accorded

⁸ See Elster (1975), pp. 127–30.

⁹ See, for example, Fleurbaey (1996).

special privilege. Let us take first an absurd example:¹⁰ always preferring goods that come on Thursdays to those that come on Wednesdays, solely because of a preference for that particular day of the week. As we shall see, this is not contrary to the principles of rational choice, but it is certainly contrary to reason. The simple preference for Thursdays is a reason, but reason also demands the reason for that reason. And obviously there is none.

Let us now take a less absurd example: preferring to receive a hundred dollars today rather than two hundred dollars a year from now. This preference is not necessarily contrary to reason. If my life expectancy is less than a year, it is perfectly well

¹⁰ Borrowed from Parfit (1984), pp. 124–25.

founded. If on the other hand it results simply from the fact that our “telescopic faculty” is deficient, as economists say, it is contrary to reason. From an objective point of view, a person who takes into account the long-term consequences of present actions has a better chance of leading a long and happy life than one who considers only immediate effects. We shall see that this fact has no relevance from a subjective point of view.

We can consider in this perspective the idea of “interest properly understood” such as it is used by Tocqueville, for example. Once again, in view of the absence of explicit definitions in classical authors, we must attempt a synthesis of their ideas. It seems to me that interest properly understood includes at least two components.