

London: Subjective Rationality

Contents

A. What subjective rationality really means 12

B. Why the notion of subjective rationality is crucial for the social sciences 13

C. An Inquiry into the Explanation of Social Behavior 17

- Guest Lecture - 23

Raymond Boudon

89/6

D. When sound methodological principles lead to false beliefs 29

E. When respectable ideas and theories lead to false beliefs 36

This discussion paper was presented as a guest lecture at the MPI für Gesellschaftsforschung, Köln, by Prof. Boudon, Université Paris-Sorbonne on April 6, 1989

References

FA80/50  
Max-Planck-Institut  
für Gesellschaftsforschung  
Bibliothek

Max-Planck-Institut für Gesellschaftsforschung  
Lothringer Str. 78  
D-5000 Köln 1  
Federal Republic of Germany  
Telephone 0221 / 336050  
Fax 0221 / 3360555

MPIFG Discussion Paper 89/6  
ISSN 0933-5668  
June 1989

EXX 0531 1330200  
1000000 0531 1330200  
1000000 0531 1330200  
1000000 0531 1330200  
1000000 0531 1330200

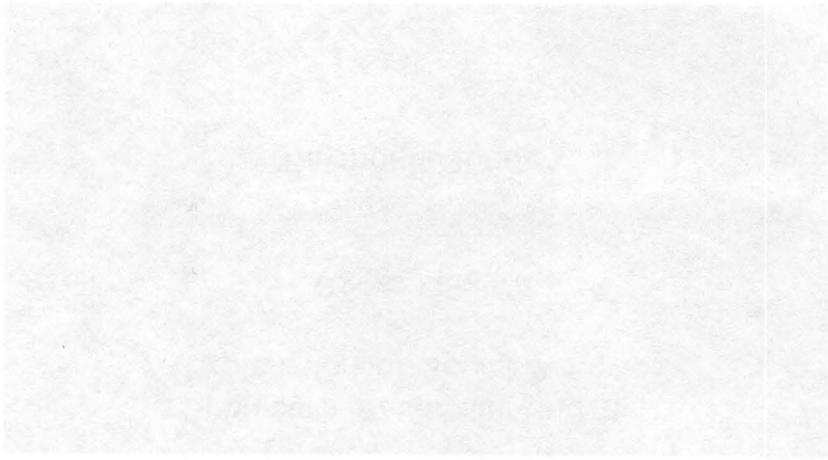
1000000 0531 1330200  
1000000 0531 1330200  
1000000 0531 1330200

18399000  
PLH-1/8906

Bibliothek  
für Gesellschaftsforschung  
Max-Planck-Institut

02/0897

1000000 0531 1330200  
1000000 0531 1330200  
1000000 0531 1330200



Contents: I have chosen for this lecture covers a limited objective, i.e. trying:

A.	What <i>subjective rationality</i> really means	5
	- to underline the importance of the notion, coined by Herbert A. Blaus, of <i>subjective rationality</i> ;	
	- to clarify its definition and meaning.	
B.	Why the notion of <i>subjective rationality</i> is crucial for the social sciences	10
	- to show that this notion is actually an implicit cornerstone of many classical	
a.	An imaginary dialogue between Hume and Downs	12
b.	Second example: the explanation of magic	17
	and finally to suggest that this notion is crucial in the analysis of many types of social phenomena.	
c.	Further classical examples	25
	As a matter of fact, we will consider notably the important case of the rational explanation of false beliefs.	
C.	<i>Subjective rationality</i> and the explanation of false beliefs	28
a.	When sound methodological principles lead to false beliefs	29
b.	When respectable ideas and theories lead to false beliefs	36
	his notion of <i>subjective rationality</i> .	
c.	When respectable mental procedures lead to false beliefs	39
	One of these is the following:	
References	"In a broad sense," he writes, "rationality denotes a style of behavior that is appropriate to the achievement of given goals, within the limits imposed by certain conditions and constraints".	45

The title I have chosen for this lecture covers a limited objective, i.e. trying:

- to underline the importance of the notion, coined by Herbert Simon, of *subjective rationality*,
  - to clarify its definition and meaning,
  - to show that the intuition lying behind this notion is actually an implicit cornerstone of many classical sociological theories,
  - and finally to suggest that this notion is crucial in the analysis of many types of social phenomena.
- As I must be selective, I will consider notably the important case of the rational explanation of false beliefs.

#### A. What *subjective rationality* really means

H. Simon (1982) has proposed several definitions of his notion of subjective rationality.

One of these is the following:

"In a broad sense," he writes, "*rationality* denotes a style of behavior that is appropriate to the achievement of given goals, within the limits imposed by certain conditions and constraints".

And he adds:

"The conditions and constraints referred to in the general definition may be *objective characteristics*, or they may be *characteristics of the organism itself* that it takes as fixed and not subject to its own control. The line between the first case and the other two is sometimes drawn by distinguishing *objective rationality*, on the one hand, from *subjective or bounded rationality*, on the other" [Simon (1982), II, 8.1)].

I am not sure this definition is entirely satisfactory. But, rather than discussing it, I will introduce the notion of *subjective rationality* by a deictic definition, i.e. by an example. I will use an example used by Simon (1982, II, p. 134) himself and borrowed from Feldman's (1963) work in cognitive experimental psychology.

Subjects are asked by an experimenter to predict the outcomes of a head-and-tail game. They are informed, however, that the coin used in the game is biased and that head and tail will appear respectively with probabilities .8 and .2.

In most cases, the subjects choose a wrong solution: they generate a sequence of outcomes governed by the same probabilities as the series they are supposed to predict. In other words, they choose to predict randomly the outcome "head" with probability .8 and the outcome "tail" with probability .2. By so doing, they

predict correctly the outcomes with probability .68.<sup>1</sup> This is a poor result compared to the result they would get if they would choose to predict head at all moves, since their probability of predicting correctly each outcome would in this case be equal to .8.

This example seems to me extremely important for several reasons.

First, it illustrates a case where the behavior of a subject is governed by reasons, which, although they are *objectively* wrong, are perceived as *good*. Very often the importance of this case is ill-recognized because it runs against a current view grounded in an old philosophical tradition which assimilates groundedness and objective validity.

This raises of course the question as to how and with the help of which criteria these reasons, which have the curious property of being both invalid and good, can be identified. The answer to this question is not immediate, but we check readily that we recognize the existence of such *good reasons* in the linguistic expressions we use normally: thus, we would likely explain the behavior of the subjects in the experiment I have reported by a sentence such as "they had *good reasons* to choose this wrong solution since ...", rather than by a formula of type "they had really no reasons of choosing this wrong solution, but ...". In other words, we consider normally their behavior as rational rather than irrational.

---

<sup>1</sup> Since  $(.8 \times .8) + (.2 \times .2) = .68$ .

Why are the reasons of the subjects perceived as good while they are wrong? Because they tried to answer the question they were confronted with by making a *guess*, a *conjecture*, by applying a *theory* or a general *principle* valid in many cases.

Suppose for instance the experimenter would ask them to predict a mathematical series governed by rules as:

- rule 1:  $y$  is the value of the first member of the series;
- rule 2: add  $x$  to the  $n$ -th member of the series to get the following member.

If the subjects would decide to use these rules, they would guess the outcomes correctly. Or suppose they would be asked to guess the outcomes of a *normal* head-and-tail game played with an *unbiased* coin. In that case, the strategy which was wrong in the case of a biased coin becomes right. By tossing an unbiased coin in order to predict the outcomes obtained by the experimenter, the subject has a probability of .5 of predicting correctly the outcome.<sup>2</sup> This is neither better nor worse than the outcome he would obtain by predicting at each move either head or tail.

On the whole, the subjects made the conjecture that, in order to replicate a model, a good strategy is to generate the copy by applying the very rules which govern the production of the model. This conjecture is obviously valid in many circumstances. It is valid

---

<sup>2</sup> Since  $(.5 \times .5) + (.5 \times .5) = .5$ .

when the model to be replicated is a mathematical series. It is even valid when the model is a head-and-tail game of the normal unbiased type. It is also obviously valid in all kinds of other circumstances. It can also be invalid, however, as here, when the head-and-tail game is biased. But, on the whole, it is much easier to create situations where the principle in question works than situations where it does not.

This explains why few people would probably accept the idea that the subjects are irrational, though their reasons were objectively ungrounded. Their reasons are perceived as *good* because they are universal in the sense that any subject in the same position, i.e. any subject who would not have been trained in probability calculus, would naturally come to the idea of using the generative rules of the model to produce the copy.

So, a first virtue of this example is to clarify the Simonian notion of subjective rationality better than the abstract definition I gave earlier. The example suggests namely that we tend normally to consider invalid reasons as good when these reasons are valid in many circumstances though not in all.

Another important point conveyed by the example is that *subjective rationality* is the product of the discordance between the complexity of the world and the cognitive capacities of the subject.

But the general conclusion to be drawn from the example is that, except in simple and marginal cases, action includes *theories, conjectures or principles,*

in other words, that sociology as well as economics and the other social sciences should develop a more cognitively oriented theory of action.

Finally, and this will be the *Leitmotiv* of this lecture, while these conjectures, theories or principles may be true, valid or well-adapted, they may also without contradiction be untrue, invalid and still grounded.

B. Why the notion of *subjective rationality* is crucial for the social sciences

Why is the notion of subjective rationality crucial for the social sciences?

Because, as Max Weber (1951) and Popper (1967) among others have suggested, trying to substitute rational explanations for the irrational explanations of behavior naturally produced by ordinary knowledge is one to the main tasks of the social sciences and even one of their main sources of legitimation.

But, as long as we do not perceive the importance of the notion of subjective rationality, and endorse a narrow definition of rationality - reasons are good when they are objectively good - we tend to see too easily behavior as irrational. And this is true not only of ordinary social knowledge, but of scientific sociology as well.

To take a classical example: partly because he had a narrow definition of rationality (see his definition of what he calls *logical* actions), a Pareto (1935) saw most social actions as irrational, as governed by feelings rather than by reasons. For the same reason many modern sociologists see easily behavior as irrational, i.e. as inspired by causes located beyond the control of the social actor rather than by reasons.

By contrast, because he had a broad view of rationality including implicitly the notion of *subjective rationality*, a Weber could defend the idea that explaining behavior amounts in most cases to disentangling the reasons of the actors.

By combining these statements, we reach in other words the conclusion that, *trying to substitute subjectively rational explanations of social behavior for the current irrational explanations is a major task of the social sciences.*

In order to avoid a possible misunderstanding, it should be made clear, however, that the interest of this substitution is *cognitive* rather than *moral*. In other words, the main legitimation for trying to interpret behavior rationally is not that this generates, say, a more optimistic picture of man. It lies rather in the fact that, when a behavior appears to an observer as strange or unfamiliar, his first move is to interpret it in an irrational fashion, as the effect of causes rather than reasons. If he could subsume reasons under the observed behavior, he would namely not perceive it as odd. This is why many impor-

tant social scientific contributions take the form of substituting a subjectively rational for an irrational explanation of social behavior.

I will take some classical examples to illustrate this point, which is particularly important since, under the influence notably of Freud, many people think on the contrary that depth and progress in the explanation of behavior consists in discovering latent unconscious causes under patent but supposedly fallacious reasons.

a. An imaginary dialogue between Hume and Downs

My first example will introduce an imaginary discussion between David Hume and Anthony Downs.

In a fascinating text in political theory, Hume (1963) develops the assumption that political parties bring together people according to three basic mechanisms: interests, solidarity, and *principles*.

In the first case, people are motivated to affiliate to a given party because they have similar social positions and consequently common interests. To use Dahrendorf's vocabulary, political parties have in this case the function of giving solid political expression to *latent groups*.

The second mechanism, the mechanism of solidarity, which Hume calls rather *affection*, is clear enough: people have normally a sense of belongingness toward

social groups and wish to protect their collective identity. In this second case, political parties help maintaining social groups in existence. Hume adds at this point that he has no difficulty whatsoever understanding these first two mechanisms. In other words, he would have accepted as easily the Marxian idea of class consciousness as the Durkheimian idea of collective consciousness.

But as a good political observer, Hume remarks that parties can also be built around *ideas* or rather *principles*. And he adds that not only he has little to say on the question as to how this strange mechanism operates, but that explaining the existence of political parties grounded on principles is probably the most difficult task the behavioral sciences are confronted with:

"Parties from *principle*, especially abstract speculative principle (...) are, perhaps, the most extraordinary and unaccountable *phenomenon* that has yet appeared in human affairs" (p. 58).

At this point Hume becomes hyperbolic. He gives the reader the feeling that he has discovered in this field of political theory a fact as extraordinary as the so-called problem of induction in the field of the philosophy of knowledge.

Why does Hume find the existence of political parties grounded on principles so puzzling? Because he has in mind a narrow view of rationality. According to this

narrow view there is only one way of knowing whether it should be endorsed or not: determining the outcomes of this program and checking whether they are acceptable or not, whether a better alternative program could be devised or not. From this point of view the parties which gather people around principles seem to walk on their head: instead of checking whether the consequences of a political program are good or not, the voters wonder whether it is grounded on the right principles.

In fact, Hume proposes an *irrational* explanation of this behavior which appears so strange to him. We would affiliate to ideological parties because we would be pushed by obscure psychic forces to agree or disagree with other people.

"(...) such is the nature of the human mind, that it always lays hold on every mind that approaches it; and as it is wonderfully fortified by an unanimity of sentiments, so it is shocked and disturbed by any contrariety. Hence the eagerness with which most people discover in a dispute; and hence their impatience of opposition, even in the most speculative and indifferent opinions" (p. 59).

But, interestingly enough, Hume is far from being satisfied by this explanation.

The paradox raised by Hume was solved by A. Downs (1957) when he showed that, the choice between, say,

two political programs, cannot be rational in the narrow sense. Anticipating with certainty what the consequences of the two programs will actually be is namely very difficult. Moreover, it is difficult to know whether they will be applied in the form where they have been presented. Finally, even if the voter would know the consequences of the programs which are proposed to him, even if he could be sure the programs will be literally applied, he can still not know what is better for him. In other words, in such circumstances, evaluating the *principles* on which the two programs are grounded is a good alternative to the impossible task of checking the consequences of the program. Exactly as for Simon's subjects imitating the experimenter was a good alternative to the mathematical solution, which is more satisfactory but unaccessible to the subject.

Downs uses at this point the word *ideology*. The rational voter *should* be ideological: this formula summarizes the central theorem of his *Economic Theory of Democracy*, the theorem which made the book influential. It shows that what appeared to Hume as mysterious and unintelligible is in fact easy to understand.

Though he contributed to destroying the influence of the classical narrow view about rationality, Downs himself appears as dominated by this view, however. This explains his choice of the word *ideology* in the formulation of his main theorem. What Downs has taught us can be more plainly described by saying that, as Simon's subjects, the voter tries to meet the situation he is faced to with the help of reasonable con-

jectures or theories, for instance the conjecture that a program grounded on certain principles will likely lead to consequences congruent with these principles. Of course, this is not always true: an equality-oriented policy can bring more inequality; by aiming at a goal one can reach the opposite outcome. But the assumption of congruence is in many cases a reasonable one, and, as a general guideline, better than the opposite assumption.

One could even go farther: for people with an interest in politics, judging a program on its principles can be a good conjecture. For less interested, it can be rational to use simpler *signals*: checking whether the candidate looks sincere for instance. In this case, the voter makes the conjecture that there is a correlation between this *signal* and the outcomes he is interested in, exactly as a businessman uses the vita of a candidate to a job as a *signal* of his future inobservable achievements.

On the whole, this Hume-Downs imaginary discussion can be considered as paradigmatic for several reasons.

- It illustrates the strength of the classical view of rationality, which even Downs does not entirely get rid of.
- It shows that, as soon as a more liberal view of rationality is adopted, the explanation of behavior becomes in many cases much easier: a type of behavior which Hume considered a mystery becomes easily understandable thanks to Downs' rational theory.

- It suggests that even though the first impression given by many behaviors can be an impression of irrationality, the actors can have *good reasons* to do what they do.
- It shows finally that these good reasons have the status of conjectures, principles or theories which most people with the same level of information and/or interest in the question they are confronted with would endorse.

b. Second example: the explanation of magic

A second, more paradoxical though equally classical example can be used to illustrate the crucial importance of the notion of subjective rationality for the analysis of social behavior: the case of the interpretation of magic. In fact, magic is a crucial phenomenon for any discussion about rationality because believing in magic is often perceived as a canonical example of irrationality: "how is it possible to believe in such causal relationships?".

Here as in the previous example, however, the substitution of a subjective rational explanation for an irrational one generates a definite feeling of progress.

Magical beliefs were actually interpreted by many writers as irrational. Thus, to Lévy-Bruhl (1960), they would show that the so-called primitive men have a mental constitution different from ours. This cause would explain for instance why they interpret verbal

associations as causal relations. Lévy-Bruhl's theory of what he called the "primitive mentality" is typically a causal irrational theory, in the sense that the behavior under examination, i.e. belief in magic, is explained as the effect of psychic causes located beyond any control of the subject.

Many people felt uncomfortable with this type of explanation, however. For causal explanations of behavior often appear to be *ad hoc* and arbitrary, at least when the causes are not reasons. Of course, some behaviors can be explained in a genuine causal non-rational way, as when it is observed for instance that the consumption of some chemical substance has a given effect on mood. But in this case and in similar cases, the causality can be empirically demonstrated, while in a case such as Lévy-Bruhl's, it is not. Here, the cause, i.e. the so-called primitive mentality, is inferred in a circular fashion from the very effects it is supposed to explain.

Interestingly enough, those who were dissatisfied with causal explanations of magical beliefs à-la Lévy-Bruhl tried often, not to find a rational explanation of these beliefs, but rather to deny the very existence of magical beliefs. This stance was taken for instance by the philosopher Wittgenstein (1975); but also by many professional anthropologists, among whom Beattie (1964) is perhaps the best known. Their common argument is that the supposed magical beliefs are actually not beliefs: they are not statements about the world, but symbolic statements about the subject's wishes. The primitive does not believe *really* that rain ritu-

als, say, would have the effect of bringing rain on his crops. By these rituals he rather express his wish that rain fall on his crops: "die Magie aber bringt einen Wunsch zur Darstellung; sie äussert einen Wunsch", writes Wittgenstein (1975).

This type of theory is neither less nor more *ad hoc* than the previous one. As well shown by Horton (1982), it contradicts the beliefs of the "primitive" themselves, who appear to be absolutely convinced that their magical rituals are not only effective, but indispensable, even though they see clearly that these rituals are only complementary to the technical operations without which no crop would grow. But Horton has not only demonstrated that primitive *really* believe in the effectiveness of rain rituals, he has also produced a fascinating indirect argument against the expressive theory of magic when he has shown that Christianity, in spite of its success in many areas in black English-speaking Africa, has often failed to replace local beliefs because it had the defect to the eyes of the Africans of not proposing a toolbox of magical devices which seemed to them indispensable in the conduct of everyday life.

In fact, the good theory is the third one, i.e. the *subjectively rational* theory developed by several modern writers but suggested curiously enough in similar terms both by Durkheim (1979) and Max Weber (1922). I said "curiously enough", because the two sociologists seldom agreed.

According to this theory, one should first take seriously the fact that those who believe in magic have no knowledge of a number of theories that Westerners have developed after centuries. Why would a "primitive" African have any knowledge of the conservation of energy, or of the Fisherian designs of experiment? Secondly, everyday action needs theories, and magic is a theory which provides guidelines to the actor. According to Durkheim, magical theories would be *applied* theories exactly as many of our technical devices are inspired by scientific theories. Of course, these guidelines are not as efficient as those of science. But showing that magic is inefficient requires the mobilization of methods of causal analysis which were developed by science.

Moreover, writes Durkheim, don't we observe that scientists often keep their faith in a theory intact even when it appears as contradicted by facts? Anticipating very precisely ideas developed by contemporary philosophers of science, as Lakatos (1970), Kuhn (1970) or Feyerabend (1976), Durkheim (1962, p. 508-528) suggests that scientists have good reasons for doing so. They can namely always hope that minor changes in the theory will make it compatible with facts, or doubt whether the facts contradictory with the theory are genuine facts or whether they are rather artefacts, etc. Moreover, a long time will often elapse before it can be ascertained which of these typical situations is created by the discovery of facts contradictory with the theory. So that it is as rational to try to save the theory as to try to replace it.

For the same reasons, magicians can keep confidence in their theories even though they do not always work. Exactly as scientists, they will without much difficulty devise auxiliary hypotheses to explain why the theory has failed.

Evans-Pritchard's (1968) work on the Azande contains effectively many concrete examples which confirm Durkheim's theory and show that the Zande magician does not behave in a way very different from the modern scientist, as described by modern philosophy and history of science.

On the other hand, even if one can legitimately assume - again on the basis of Evans-Pritchard's work for instance - that the primitive has some intuitive knowledge of the procedures by which a causal relationship is confirmed or rejected, it must be recognized that this knowledge is often practically of no help. Even in our world where methods of causal analysis are well mastered, many people and among them many scientists appear to believe in all kinds of unconfirmed causal relations.

Thus, many people see a causal link between all kinds of practices and state of health or length of life. The reasons for such beliefs lie simply in the fact that in such cases collinearity is often so powerful that it is practically impossible to check seriously whether or not  $x$  has really an effect on  $y$ . Thus, it has apparently been discovered recently after many years that the idea according to which stress would be a main cause of stomach ulcer is a magical belief, so

to say. This magical belief was endorsed by many people, however, including many scientists.

To this, a subtle argument explicitly developed by Durkheim (1962, p. 527) can be added. Ethnologists have observed that the rain rituals are celebrated during the rainy season. Wittgenstein (1975) interpreted this fact as supporting his theory that rain rituals are perceived by the primitive as having an expressive rather than instrumental function. The alternative interpretation of this fact proposed by Durkheim is much more ingenious and interesting: the confidence of the primitive in the effectiveness of their rituals, says Durkheim, is reinforced by the fact that rain is actually more frequent in the period of the year when crops start growing, need rain and when consequently they celebrate these rituals. Durkheim's idea is in other words that collinearity would reinforce the magical causal beliefs of the primitives, i.e. that their causal assumptions can be confirmed by genuinely observable correlations.

On the whole, Durkheim suggests that the magical beliefs of the primitive are not different from ours. The difference is that, because of the development of science, a number of old beliefs have become obsolete. As a consequence, when we see other people still believing in them, we have a strong feeling of irrationality. But as soon as we evoke the causal relationships in which we believe ourselves on such existential problems as health for instance, we become conscious of the fact that magical and scientific knowledge coexist in primitive societies.

Max Weber (1922, p. 227-228) wrote on his side that to the primitive, "the actions of the rainmaker are exactly as magical as the actions of the firemaker". He meant by this statement that the primitive has no reason to introduce the asymmetry between rain- and fire-making which we introduce ourselves, precisely because the process by which kinetic is transformed into thermic energy is familiar to us. Reciprocally, the primitive to whom this process is unfamiliar has no reason to see a basic difference between fire- and rain-making.

So, Weber's as well as Durkheim's implicit diagnosis on magic is that the primitive should not be considered to be less rational than ourselves. When they are confronted with existential problems, they develop theories and conjectures, which they derive from the socially available body of knowledge. They are eventually ready to forget them as soon as they are proposed better theories. But as long as these theories are not available, they have also good reasons for believing - as scientists - even in theories which repeatedly fail. Moreover, reality can reinforce rather than contradict their beliefs, even when these beliefs are false, as Durkheim has noted in one of the most brilliant parts of his analysis of magic.

We have here an example where substituting a subjective rational theory for a set of irrational interpretations gives a strong feeling of scientific progress. By contrast with the other two types of theories, the rational theory of magic does not give this impression of being *ad hoc* and arbitrary. It rests upon a set of

psychologically easily acceptable statements and appears to be congruent with available data.

c. Further classical examples

Mentioning Durkheim in the frame of the present discussion is also interesting for another reason: that Durkheim is seldom seen as a writer who would have insisted on the idea of explaining actions and beliefs rationally. But while this is true of his theoretical writings, he often offers in his empirical analysis rational - in the subjective sense - explanations of social behavior.<sup>3</sup>

To take another example: when Durkheim (1962) tries to explain in his *Suicide* why economic booms seem to be associated with higher rates of suicide, he introduces an explanation of the *subjective rationality* type, not very far away from ideas developed by A. Hirschman (1980) in his theory of the tunnel effect. Durkheim's assumption is namely that the anticipations and expectations of social actors are grounded on good reasons: during a period of stable economic development, they tend to start from the principle that they can expect, say, for the year to come, the same gains as the year before, while during an economic boom, when the situation of many people appears to be getting better, they will change their conjectures as to which objectives can be reached and aimed at.

---

<sup>3</sup> A point which I have tried to develop in Boudon (1981).

Durkheim introduces implicitly at this point a very brilliant hypotheses, namely that people would extrapolate from the tangent to the curve at each point of time. So that in the first part of the ascending phase, before the inflection point, their expectations would tend to be under-optimistic, while in the second part they would be over-optimistic. This is at least my interpretation as to why Durkheim predicts an increase in disillusion and consequently in suicide rates in the second part of the ascending phase of the business cycle, but not in the first.

I have tried to show for my part with the help of a formal model [Boudon (1986)], why, as hypothesized by Durkheim, but also by Tocqueville and others, an economic boom will in many circumstances produce less satisfaction than a stable business cycle, because it will generate over-optimistic conjectures.

I could offer about Marx the same remarks as I have just made on Durkheim. In the first pages of the *German Ideology*, he develops a famous irrational theory of beliefs: People believe in all kinds of false ideas because obscure social forces make them see the world in a distorted fashion. But in many of his analyses, in contradiction with his general theory of beliefs, Marx produces a rational explanation. Probably because he felt, as Durkheim, although he did not want to recognize it, that explaining behavior or beliefs means finding the *good reasons* behind behavior and beliefs.

Thus, at one point in *Capital*, Marx (1867) wonders why workers accept so easily exploitation. A behavior running against the very interests of an actor is typically explained in most cases in an irrational fashion. Against his own principles, Marx sketches in *Capital* a rational explanation, however, which I reconstruct in the following way: to know that they are exploited, the workers should see the difference between their salary and the value of their work. But to determine the value of their work, they would have to master a complex economic theory, as well as a great deal of empirical information. As this is impossible and as they still want to know whether their salary is fair or not, they turn toward a natural substitute: comparing their salary to the salary of the individual producer. But, by so doing, they introduce of course a bias in their estimation, the individual producer using more time to produce, say, a pair of shoes than the worker in a shoe manufacture. On the whole, they will agree to leave to the capitalist the gain resulting from the division of labor. But they will do so with good reasons.

One need not of course endorse the theory of the surplus value and still admire the ingeniousness with which Marx explains how the social actor can with good reasons act against his own interest: even masochism can be grounded on *good reasons*.

---

A point which I have tried to develop in  
[unclear] (1981).

C. *Subjective rationality* and the explanation of false beliefs

The examples I have drawn from Weber and Durkheim as well as Marx lead me to the final part of this lecture, where I will try to underline the importance of the notion of subjective rationality in the analysis of beliefs.

Beliefs, notably false beliefs, are currently interpreted as the product of causes located beyond the control of the subject rather than of reasons. Marx, in his theoretical writings at least, Freud and Pareto among the classics illustrate clearly this paradigm. In the same way, modern sociologists of knowledge analyze in most cases beliefs as the product of obscure and invisible social or psychological causes. Alternatively, they content themselves with exploring the correlations between independent variables and beliefs and do so, this is at least my guess, because they endorse currently a naturalistic irrational view of men.

The notion of subjective rationality sketches, on this subject of beliefs also, an alternative paradigm, which I think is potentially much more fruitful. It follows immediately from the fact that good reasons can be objectively invalid, so that the social actor can with good reasons believe in false ideas.

So, an interesting theoretical task of the sociology of knowledge would be to identify and classify a population of examples where mental procedures typical of

subjective rationality produce false beliefs. While I cannot undertake this task here, I will present some examples to suggest the potential interest of such developments.

More precisely, I will consider three cases, namely the cases where respectively *sound methodological principles*, *sound ideas or theories* and finally *sound intellectual procedures* lead to false beliefs.

a. When sound methodological principles lead to false beliefs

We saw an example of the first case in the experiment reported by Simon I presented earlier. The subjects started from the idea that, in order to reproduce a model, the best thing to do was to produce a copy following the rules according to which the original was built. The error comes here from a natural application of an apparently innocent principle.

Another seemingly innocent principle, a principle of vital importance which we are in many occasions entitled to follow, is the following: When we have to make up our mind on an empirical question, for instance as to whether a given  $x$  is  $y$ , we would normally check empirically, if we are in the position of doing so, whether or not  $x$  is actually  $y$ . If I am asked whether a given book is on a given table, the best thing to do is to look at the table and see whether the book is actually there. The application of this natural prin-

ciple can lead to false views about the world, however.

To illustrate this point, I will draw a first example from a polemical discussion which, once more, was started in France in the last few weeks on the question as to whether the cognitive level of youngsters tends to decline or to climb as an effect of the overall increase in education. I will leave aside the trivial point that checking the cognitive level of a population is more complicated than checking whether a book is on a table or not. For, even if we assume that cognitive levels can be easily observed, strong illusions about their evolution can still appear with good reasons in many heads.

To introduce my point, I will build a highly simplified model: I will assume that there is a variable such as the capacity to learn and that in each of the successive cohorts of youngsters this capacity is distributed in the same way. I will suppose moreover that each year the number of students taking a given grade increases, and that this increase corresponds to a population located lower on the hypothetical distribution from one cohort to the next. Although this model is caricatural, it is obviously not without relation with the real world.

Suppose now a teacher is asked whether the cognitive level of the secondary school students is increasing or decreasing. He would naturally answer this question by reference to his own experience. Now, according to the assumptions of the model his conclusion would be

entirely unambiguous: he can see directly that the level of his students is declining regularly from one year to the next. Moreover, he could easily check that his impression is general, since most of his colleagues would reach the same conclusions for the same reasons. On the whole, all teachers, except those who see the world through biased glasses, would consider as an empirical evidence that the cognitive level of high school students is declining.

This pessimistic impression of the quasi-unanimity of our teachers is obviously not incompatible with the fact that the average cognitive level of the youngsters between, say, 15 and 20 is increasing regularly according to the model. For, except in the case where we would be ready to accept the very unlikely assumption that education would not only *not* increase knowledge but even destroy it, the model predicts also a general increase of the cognitive level. More concretely, the model shows that all English teachers for example would be convinced by their immediate experience that the level of their class in English is decreasing regularly over time. This opinion would be grounded on good reasons. But on the whole more people would know more English over time.

So, the false beliefs of our teachers are in no way illusions. On the contrary, they perceive the reality around them *as it is*. The false belief comes from the application of a principle which is in most cases alright, i.e. that observing whether  $x$  is  $y$  is a good way to determine whether  $x$  is  $y$ . This principle is effectively most often valid. But, as here, it can

also lead with good reasons to beliefs in false statements.

Many examples of the same vein could be mentioned where subjective rationality produces false ideas or beliefs, where in other words subjects appear to be endorsing false ideas or beliefs with good reasons. I will add two further short examples in order to suggest that the idea is general.

According to a widely accepted point of the Keynesian theory, an increase in taxes has normally a deflationary effect. For the following reasons: it generates a decrease in the purchasing power, which in turn generates a decline in the demand for goods and services, and finally a decline in the prices of goods and services. Now, in a survey conducted by Katona (1951), a sample of businessmen were asked whether an increase in taxes has deflationary or inflationary effects. A majority of them answered it should increase inflation. Why? Simply because they analyzed the question by reference to a situation familiar to them: when taxes climb, this has the effect of increasing the costs of production the businessman has to bear. Except in the case where the demand for his products is highly elastic with respect to price, he will normally try to transfer at least a part of this increase in his costs of production to the consumer by increasing the price of his products or services. From this simple mental experiment he concludes with excellent reasons that an increase in taxation should generate inflationary effects.

Or consider another example. Many people are convinced that substituting mechanical for human work increases unemployment. Here again, this conviction can be grounded on *good reasons*. A simple mental experiment shows namely that, at the local level, the introduction of new machines will normally in many circumstances destroy some jobs. While, at the general level, it can have a positive net effect on employment, since the machine will have to be produced, maintained, replaced after a while by a better-performing machine which will have to be devised and produced, etc. But this positive influence is visible at a general abstract level, the level at which the economist is located. At the local level where the worker is located, he will perceive on the contrary a negative effect. As in the previous examples this perception is not an *illusion*, however. What the worker sees corresponds to the reality as it is.

In other words a false belief results here, as in the other examples, from the combination of a correct non-illusory perception and of the general apparently innocent principle according to which questions of type "does x produce y?" can be answered by checking whether empirically x produces y.

Questions of this type - "does x produce y?", "is y an effect of x?" - are obviously questions which we ask almost as frequently as we breathe. Causal diagnoses are as vital in everyday life as they are in scientific research. Now, a principle often used to reach a causal diagnosis as to whether, say, x is cause of y, a principle generally considered as safe and valid,

consists in observing whether  $x$  and  $y$  tend to appear together. If I observe a certain number of times that drinking too much wine makes me sleepy, I will normally conclude that wine...tends to make me sleepy.

Like the previous one, this apparently innocent principle can lead to false beliefs, however. This has admirably been shown by a number of experiments conducted by cognitive psychologists. Here is one of them, which was conducted by Shweder (1977). A sample of nurses are presented 100 cards representing 100 supposed patients. On these cards two fictitious informations have been recorded: whether or not the patient shows a given hypothetical symptom; whether or not a hypothetical disease has been diagnosed in his case. The joint distribution is the following:

		DISEASE		
		Present	Absent	Total
SYMPTOM	Present	37	33	77
	Absent	17	13	30
	Total	54	46	100

While the correlation is very low and goes in the "wrong" direction, 85% of the nurses thought that the symptom was an effect of the disease. Why? Because they started from the principle that a number of cases as high as 37 where the two features occurred together was sufficient to show that the disease is the cause of the symptom. After all, if 37 times I feel sleepy

after drinking wine, would I not be convinced that wine makes sleepy?

While finding his paper extremely interesting, I disagree with the author of the article on one crucial point. He claims that the nurses of his experiment illustrate "magical thinking", as he writes in the title of his paper. If magical thinking is defined as it is usually by the fact of seeing causal relationships where there are actually none, it is true that the nurses think magically. But I disagree with him when he seems to interpret the beliefs of the nurses in the irrational way in which Lévy-Bruhl interpreted the magical beliefs of this "primitives": they would see a causal relationship where there is none because they would follow the rules of a primitive logic unconsciously.

In fact Shweder's interpretation was inspired to him because he perceived the answers of this nurses through the glasses of his scientists, because we manipulate currently contingency tables with their three (once  $N$  is known) degrees of freedom, we tend to consider as normal the situations where three degrees of freedom have to be taken into account before a causal statement can be proposed. And we do not see that in many cases it is sufficient to check that  $y$  follows frequently  $x$  to reach a reliable answer to the question as to whether  $x$  is the cause of  $y$ .

So, Shweder's nurses are not more irrational than Simon's subjects in the experiment I reported at the beginning of this lecture. And it is more illuminating

to interpret their answers - on the *subjective rationality mode* - as dictated by the application of principles which are currently valid while in some circumstances they are not.

Finally, using sound methodological principles, such as looking for instance whether x is y in order to know whether x is y, looking whether x and y appear frequently together to determine whether they are related can and often does lead to right ideas. But they can also lead to wrong ones. In other words sound principles can be the cause of false beliefs.

b. When respectable ideas and theories lead to false beliefs

In the same fashion genuinely respectable ideas can lead to false conclusions.

Several surveys on beliefs about supernatural phenomena show for instance a strange result. Believing in the existence of God becomes *less* frequent when the level of education is higher. But other supernatural beliefs appear *more* frequently. Thus, believing in the existence of extraterrestrial beings or in psychic action at a distance seems to increase with the level of education. In a survey [mentioned by Renard (1988)], 48% of the interviewees with a primary education, 62% of those with a secondary education and 73% of those

---

A survey conducted by Boy and Michéas (1986) leads to the same kind of observations.

with a higher education appeared to believe in extra-terrestrial beings.<sup>4</sup>

Several interpretations can be given of these results. The first reaction of many sociologists to this type of findings is to interpret the belief as irrational: more education, less common sense. A more careful analysis shows, however, that this belief can often be interpreted as the consequence of a conjecture grounded on *good reasons*. People with a scientific education have learned, namely, that concepts, entities or mechanisms which were considered for a while as contradictory with the very notion of science have frequently been incorporated into the body of scientific knowledge later. The notion of physical action at a distance is a good example in this respect: the Cartesians developed all kinds of theories to explain by direct mechanical effects the physical phenomena which appeared to reveal the existence of forces acting at a distance, a concept which they considered unacceptable, until Newton made the notion respectable and, by so doing, discredited the Cartesian theories.

Why would the progress of knowledge not incorporate in the same way the notion of a psychic action at a distance? Why would not future scientific research demonstrate the existence of superterrestrial beings? As empirical sociologists have currently a naturalistic vision of the *homo sociologicus*, they seldom care about retrieving his reasons, so that I have no proof that such reasons explain the correlations I mentioned

---

<sup>4</sup> A survey conducted by Boy and Michelat (1986) leads to the same kind of observations.

between beliefs and level of education. But it seems plausible that the respondents with a higher education have finally good reasons for being less skeptical or critical on some subjects than those with less familiarity with the history of science. To summarize this case in an ironic fashion, I would say that the confidence in the scientific virtue of methodical doubt can occasionally lead to more credulity.

Many examples of the same vein could be mentioned. Another respectable genuinely scientific idea is for instance the Keynesian theory according to which stimulating the demand for goods and services, for instance by distributing more purchasing power or reducing taxes, will stimulate the supply of goods and services and consequently reduce unemployment. This respectable and true idea can lead to dangerous beliefs, however, if it is forgotten that Keynes proposed it in an economic situation where aggregate demand was much below the potential supply. For in a business cycle where aggregate demand exceeds potential supply, stimulating the demand will produce inflation rather than a decrease in unemployment.

Thus, false beliefs are often generated by the most respectable and sometimes by the most genuinely scientific theories. When for instance the hidden assumptions of a theory remain unperceived, it can lead many people to believe with good reasons that  $x$  will produce  $y$  while it will actually produce  $z$ .

c. When respectable mental procedures lead to false beliefs

Very often, false ideas can also be generated by the procedures of thought, such as induction for instance. As we know since Hume, though induction is logically ill-grounded, it is vital for everyday life and currently used in science. But it can generate false beliefs as in the case of those children who were asked by Balacheff (1987) to look at two triangles similar in all respects except size, the triangle ABC being greater than A'B'C', and to answer some questions about them. When they were asked whether the side AB is greater than A'B', they answered "yes". "Is BC greater than B'C'?" - answer: "yes". "Is AC greater than A'C'?" - answer: "yes". "Is the perimeter of ABC greater?" - answer: "yes". "The surface?" - "yes". "The sum of the angles of ABC" - answer: "yes".

Impeccable formal deductive procedures such as the syllogism, for instance, although leading in most cases to right conclusions will also eventually be responsible for false beliefs when they are used in an overconfident fashion. I am not thinking here of the sophisticated use of correct deductive procedures which a Pareto (1935) had in mind when he criticized for instance these deductive arguments where the meaning of the same word changes from one statement to the next in the argument. Thus, he contended, the words "nature", "natural", etc. are used in many moral arguments to "prove" that something is good or bad. The arguments are often formally correct but they are sophisticated because these words seem to be used with

variable meanings. What I have in mind here is an entirely different case.

Take for instance the following deductive theory: in the average, the social origin of an individual has an influence on his level of education, his level of education on his social status. From these statements, we conclude easily that, if the relationship between origin and education becomes weaker over time, the relation between social origin and status will also appear to be progressively declining. The argument is absolutely correct. The words have the same meaning in each statement, the statements are logically connected to one another in a satisfactory fashion. So that we have all possible good reasons to accept this conclusion.

But the argument contains also, like any argument, all kinds of implicit statements. In general, these mental harmonics do not disturb the argument more than musical overtones disturb the main line of a melody. But sometimes, arguments can contain, as here, implicit statements which make the conclusion we derive from them much more fragile than we believe. Here, the argument takes for granted that each of the two relationships can change without the other changing. While this type of assumption can ordinarily be made in most cases and remains for this reason implicit, it produces here a false belief, as shown by Boudon (1974). Since the relationship between social origin and educational level cannot plausibly become smaller without the overall number of people being educated increasing, the second relationship - between educational

level and status - cannot likely remain constant. A false belief is generated here by applying a generally valid procedure to a case where it is not valid, just as imitating a copy suffices in general to reproduce it with fidelity, while it can also occasionally deform it.

This example is also interesting because it shows once more that, contrarily to what the Simon's example I started from could suggest, subjective rationality appears not only in everyday intuitive inference but also in scientific thinking. For it is clear that although scientific thinking is obviously more controlled than everyday thinking, a scientific procedure contains always, as in the previous example, invisible implicit steps beside its official controlled steps. These implicit assumptions are always present. In most cases they are both unconscious and harmless. In other cases they lead to false beliefs.

This point is essential, since it underlines one of the basic mechanisms by which genuinely scientific arguments can provide the *good reasons* which inspire and consolidate false beliefs.

So, all kinds of usually valid and vital principles and mental procedures - induction, formal procedures such as the syllogism, sound principles such as the principle recommending to check empirically whether x is y in order to judge whether x is y, etc. - in other words the complex set of theories, procedures, principles which are used by *subjective rationality* are also

basic mechanisms by which false ideas and beliefs can be grounded and consolidated.

Thus, the notion of subjective rationality sketches a paradigm in the frame of which all kinds of behaviors, including behaviors oriented against the interest of the subject, as well as all kinds of attitudes and beliefs, including magical beliefs and more generally beliefs in false ideas, can be more convincingly explained than with other paradigms.

More generally, the notion of subjective rationality seems to me a most important theoretical piece for all social sciences. Substituting subjective rational explanations for irrational explanations of behavior typically is a major form of progress in the social sciences, while reciprocally irrational explanations, because they tend to be easy, *ad hoc*, circular and empty are a main plague of sociology.

The fact that irrational explanations are often considered by sociologists to be natural can be detected for instance at the fact that many studies are based on correlational analyses interpreted in a causal naturalistic fashion. Thus, many sociologists will be satisfied to know that income, or class, say, have an influence on the probability of believing in God or in extraterrestrial beings, or on the probability of reaching a given educational level, and will often not even wonder whether or not reasons can explain the correlation.

The appeal of irrational explanations of behavior to sociologists derives from several facts: the influence of the exoteric Marxian and Freudian tradition, but also the influence of the narrow definition of rationality according to which good reasons should be objectively grounded. The strength of this narrow definition comes itself from the fact that it is deeply rooted in a secular philosophical tradition.

Once the importance of this concept of subjective rationality is fully recognized, the strange division of labor between economists who tend to see the *homo oeconomicus* as narrowly rational and sociologists who often see the *homo sociologicus* as basically moved by forces located beyond his control will be attenuated, this is at least my guess, and replaced by a more interesting ideal-type, I mean the ideal-type of the *motivated actor*, i.e. of an actor who in many cases has good reasons to do what he does or to think what he thinks.<sup>5</sup>

---

<sup>5</sup> A reaction against irrational explanations of behavior has appeared with the so-called new economics, as presented for instance by McKenzie and Tullock (1975). This paradigm proposes to use a narrow definition and theory of rationality usual in classical economics in the analysis not only of economic behavior but of behaviors not belonging to the traditional field of economics. Very often this approach is presented as a reaction against the traditional sociological approach which is held as dominated by an irrational view of man. This verdict is true to some extent at least: the Weberian tradition seems to be less influential in sociology than the traditions which see the social actor as irrational. But the reaction to this verdict is not the right one - it seems to me. It is true that the narrow rational model can be applied to the analysis not only of economic behavior but to other types of behavior as well. But this model should be held as a particular case of a more general model.

On the whole, the notion of subjective rationality sketches a general model cutting across the various disciplines: it expresses the idea that, except in simple situations where the social actor can follow the narrow rational model, he is normally confronted with ambiguous and complex situations and masters them by using theories, principles, conjectures, in other words by being intellectually more active than both sociologists and economists recognize.

For all these reasons, the notion of subjective rationality appears to me to be one of the most strategic concepts for the development of the social sciences.

## References

- Balacheff, N. (1987), "Processus de preuve et situations de validation", *Educational Studies in Mathematics*, 18, 147-176.
- Beattie, J. (1964), *Other Cultures*, London, Cohen & West.
- Boudon, R. (1974), *Education, Opportunity and Social Inequality*, New York, Wiley.
- Boudon, R. (1981), *The Logic of Social Action*, London, Routledge; German: *Die Logik des gesellschaftlichen Handelns*, Luchterhand.
- Boudon, R. (1982), "The Logic of Relative Frustration", *in* *The Unintended Consequences of Social Action*, London, McMillan; German: *Die Widersprüche sozialen Handelns*, Luchterhand.
- Boy, D. and G. Michelat (1986), "Croyances aux parasciences: dimensions sociales et culturelles", *Revue Française de Sociologie*, XXVII, 2, 175-204.
- Downs, A. (1957), *An Economic Theory of Democracy*, New York, Harper.
- Durkheim, E. (1962), *Suicide, a Study in Sociology*, Glencoe, The Free Press, (1897).
- Durkheim, E. (1979), *Les formes élémentaires de la vie religieuse*, Paris, Presses Universitaires de France, (1912).
- Evans-Pritchard, E. (1968), *Witchcraft, Oracles and Magic among the Azande*, Oxford, The Clarendon Press, (1937).
- Feldman, J. (1963), "Simulation of Behavior in the Binary Choice Experiment", *in* E.A. Feigenbaum and J. Feldman (eds.), *Computers and Thought*, 329-346.
- Feyerabend, P. (1975), *Against Method*, London, NLB.
- Hirschman, A. (1980), "The Changing Tolerance for Income Inequality in the Course of Economic Development", *in* A. Hirschman, *Essays in Trespassing*, Cambridge, Cambridge University Press, 39-58.
- Horton, R. (1982), "Tradition and Modernity Revisited" *in* M. Hollins and S. Lukes, *Rationality and Relativism*, Oxford, Blackwell, 201-260.

MAX-PLANCK-INSTITUT FÜR GESELLSCHAFTSFORSCHUNG, KÖLN

## - Publikationen -

- Hume, D. (1963), *Essays moral, political and literary*, Oxford, Oxford University Press.
- Katona, G. (1951), *Psychological Analysis of Economic Behavior*, New York, McGraw Hill.
- Kuhn, T. (1970), *The Structure of Scientific Revolution*, Chicago, Chicago University Press.
- Lakatos, I. (1970), "Falsification and the Methodology of Scientific Research Programmes", in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, London, Cambridge University Press, 91-196.
- Lévy-Bruhl, L. (1960), *La mentalité primitive*, Paris, Presses Universitaires de France, (1922).
- Marx, K. (1867), *Das Kapital*, Hamburg, O. Meissner/New York, L.W. Schmidt.
- McKenzie, R.B. and G. Tullock (1975), *The New World of Economics*, Homewood, Irwin.
- Pareto, V. (1935), *The Mind and Society. A Treatise on General Sociology*, New York, Dover Publications, 2 vols.
- Popper, K. (1967), "La rationalité et le statut de principe de rationalité", in E.M. Claassen (red.), *Les fondements philosophiques des systèmes économiques*, Paris, Payot, 142-150.
- Renard, J.B. (1988), *Les extraterrestres*, Paris, Cerf.
- Shweder, R.A. (1977), "Likeness and Likelihood in Everyday Thought: Magical Thinking in Judgments about Personality", *Current Anthropology*, vol. 18, No. 4, 637-658.
- Simon, H. (1982), *Models of Bounded Rationality*, Cambridge, MIT Press.
- Weber, M. (1922), *Wirtschaft and Gesellschaft*, Tübingen, Mohr.
- Weber, M. (1951), *Gesammelte Aufsätze zur Wissenschaftslehre*, 2. Aufl., Tübingen, Mohr.
- Wittgenstein, L. (1975), "Bemerkungen über Frazer's *The Golden Bough*", in R. Wiggershaus (ed.), *Sprachanalyse und Soziologie*, Frankfurt, Suhrkamp, 37-58.