

Homo Sociologicus: Neither a Rational nor an Irrational Idiot

Raymond Boudon

Université de Paris-Sorbonne (Paris IV).

[Metadata, citation and](#)

ments de la UAB

Abstract

As the social sciences deal with macrophenomena which are caused by individual human actions, they have to use a theory of human behaviour. They use basically three types of theory: the rational-utilitarian theory, the causalist theory which sees behaviour as caused by social, cultural or biological forces, and the «rational psychology» in Nisbet's sense. The three theories are important in the sense that they have shown their capacity to explain convincingly puzzling phenomena. Type I and II theories have been claimed to be potentially general. They are not since there are familiar phenomena they are unable to explain. Type III can by contrast be claimed to be general. Amartya Sen has made the point that the rational-utilitarian theory treats human beings as rational idiots. Type II treats them as irrational idiots. These aggressive metaphors draw the attention on the fact that human beings answer the situations they are confronted to by devising systems of arguments which they perceive as strong: by being cognitively rational.

Key words: rational choice, cognitive rationality, axiological rationality, theories of behaviour.

Resumen. *El homo sociologicus: ni idiota racional ni irracional*

Puesto que las ciencias sociales estudian macrofenómenos causados por acciones humanas individuales, tienen que utilizar una teoría del comportamiento humano. Utilizan básicamente tres tipos de teorías: la teoría racional-utilitaria, la teoría causal que ve el comportamiento como causado por fuerzas sociales, culturales o biológicas, y la «psicología racional» en el sentido de Nisbet. Las tres teorías son importantes en el sentido de que han demostrado su capacidad para explicar convincentemente fenómenos enigmáticos. Se ha pretendido que las teorías del tipo I y II eran potencialmente de aplicación general. No lo son, dado que hay fenómenos familiares que son incapaces de explicar. Las del tipo III, por el contrario, pueden pretender generalidad. Amartya Sen ha argumentado que la teoría racional-utilitaria trata a los seres humanos como idiotas racionales. Las teorías del tipo II los tratan como idiotas irracionales. Estas agresivas metáforas llaman la atención sobre el hecho de que los seres humanos responden a las situaciones con las que se enfrentan ideando sistemas de argumentos que perciben como poderosos: esto es, siendo cognitivamente racionales.

Palabras clave: elección racional, racionalidad cognitiva, racionalidad axiológica, teorías del comportamiento.

Contents

A typical analysis: Durkheim analyses a correlation	Type III model
Three theories of behaviour	Type III model and Weber's methodological writings
Type I or II?	The importance of the discussion
Is Type I general?	References
Choosing Type II instead of Type I?	

A typical analysis: Durkheim analyses a correlation

Durkheim (1967[1897]) observes that suicide rates go down when a serious political crisis appears. They go down in France at the time of the main crises of the Third Republic: the Dreyfus and Boulanger Affaires. They decline both in Austria and Prussia during the tension and the war between the two nations which was to be concluded by the Prussian victory at Sadowa. Why this counterintuitive correlation? Because, Durkheim suggests, when a serious national or international crisis appears, people cannot help worrying at what happens in the outside world and are provisionally deterred from their personal problems. Once this psychological mechanism is made explicit, our confidence in the fact that the correlation reflects a genuine causal relationship is increased.

I think this analysis includes an essential methodological teaching: it exemplifies a procedure characterizing, not only sociology, but all sciences.

An investigation begins often with some macroscopic observation, as a correlation between two variables, not only in sociology but also, say, in the life sciences. Thus, a study on nutrition is likely to start from the macroscopic observation that some type of cancer appears as less frequent among people who eat oil from vegetal rather than animal origin. But, one will be really certain that the correlation reflects a genuine causal relationship only from the moment when it will have been possible to describe the elementary mechanisms responsible for the correlation. And, even when the design of experiment is conducted carefully enough, so that there is little doubt that the correlation is due to a genuine causal relationship, the investigation will still give the feeling of being unfinished, since the question will remain as to from which elementary causes the macroscopic causal relationship come from.

So, a general rule of any scientific investigation in any science is that, in order to explain a macroscopic phenomenon, the analyst should try to identify the ultimate elementary causes responsible for the observed macroscopic phenomenon. As far as the social sciences are concerned, Weber and Schumpeter have given this general rule a name: «methodological individualism», since the elementary causes of any social phenomenon are obviously human individual actions. Once this evident principle is recognized, crucial questions appear: how are human actions validly explained? What are their causes? Is it possible

to identify these causes? How? These questions raise the more general question as to «Which theory of behaviour for the social sciences?».

This question has been and is still intensively debated in the social sciences, because it is crucial. It seems to me the point can be made that the variety of answers the social sciences have given and give to this basic question can be reduced to three basic types.

Three theories of behaviour

Type I. The first Type is illustrated by the so-called «Rational Choice Theory» (RCT), also called the «expected utility model». According to theorists such as Gary Becker (1996) or the late James Coleman (1986), RCT should be considered as the theory of behaviour valid for all social sciences. According to this model, people do what they do because they think that their action will give them a maximum satisfaction. The kind of psychology used in the RCT can be qualified as *utilitarian*. It can also be called *consequential*, since it assumes that people do what they do in consideration of the consequences their action will generate. This type of theory was developed systematically by Jeremy Bentham. It has been revived in the last years.

Type II. According to the second type of theory, behaviour should in most cases be explained as the effect of psychological, cultural or biological forces: by socialization effects, by cognitive biases, as the outcome of biological evolution, as the product of some instinct such as the instinct of imitation, etc. While these forces can be psychological or cultural, the ultimate causes underlying them are biological. They cannot in most cases be disentangled by biology in the present state of the art though. For this reason, the forces thanks to which they explain behaviour have the status of mere conjectures, except when they deal with psychological mechanisms known since ever.

The second Type of model is used by many sociologists and many anthropologists, as Clifford Geertz (1984). They do not deny the fact that individuals have intentions, but they maintain that their intentions are determined by psychological, cultural or biological forces, i.e. by causes beyond the control of the individuals themselves. They contend that intentional as well as non-intentional behaviour should *generally* be explained by causes beyond the control of individuals. The psychology used by this model can be called *causalist*, though this appellation raises the objection that not only forces but reasons can be causes.

Type III. The third model is used explicitly by Max Weber. But it is also used implicitly by Tocqueville and even by Durkheim and by many of their followers. They consider that the cause of an individual action is its meaning in the mind of the individual. They regard individual actions as being in principle understandable, as inspired either by understandable motivations or by reasons. Following Nisbet (1966), I propose to call *rational* the kind of *psychology* used

in this model. This psychology is the psychology developed by Aristotle, and by the French moralists of the 17th century. In a word, it is no other than the psychology used in everyday life, or in institutional life, in police stations or in judiciary institutions for instance. This type of rational psychology is used in the example from Durkheim I mentioned earlier: when the fire burns outside, I forget for a while grumbling about my personal problems: an understandable reaction. Tocqueville shows that in 1789 the French people had reasons to be hostile to the clerics which the British people had not. Hence the macroscopic effect that the French enlightenment was much more anticlerical than the British one or than the German *Aufklärung*. Max Weber has explained that the Roman soldiers had reasons to prefer the monotheistic to the polytheistic cults. This explains the quick diffusion of Christianity in the first centuries in the Roman Empire. Neither Tocqueville, nor Durkheim, nor Weber had a use for causalist psychology. They consider that an action, belief, or attitude can be held as irrational, as produced by hidden «forces», exclusively from the moment when all rational possible explanations have been exhausted.

Of course, the three types I have distinguished are «ideal types»: many works are eclectic in the sense that they use simultaneously the three types of psychology. But still, it can be recognized that, say, Clifford Geertz' analyses make little use of rational or of utilitarian psychology, while Tocqueville's, Durkheim's or Max Weber's analyses, say, have no use for the causalist type of psychology and use essentially rational psychology in the above sense.

This discussion is far from being merely speculative. On the contrary, the answer given to the question as to which theory of behaviour the social sciences should preferably use determines to a large extent the effectiveness of their analyses, as I will try to suggest.

Type I or II?

I will illustrate this point by showing that shifting from one of the three models to another can lead to substitute an illuminating explanation for a questionable one.

In a study on Vietnam in the twenties of the last century, Popkin (1979) wondered why all traditional Vietnamese village societies had adopted the unanimity rule as the basic constitutional rule. He looked at the literature on the question as to why this rule is commonly accepted in village societies, in South East Asia, but also in Africa and elsewhere. In fact the unanimity rule is currently used in traditional village societies around the world. Popkin noted that most available explanations of this phenomenon use a model belonging to the second type of behavioural theory: they assume that people were motivated to devise and apply the unanimity rule under the effect of cultural forces.

They mobilize generally the assumption that in traditional societies people have a weak feeling of their individuality; that they have the impression to exist only as one part of the group; that consequently the group itself is in

their mind the only subject able to legitimate a collective decision; and finally that to them unanimity is, so to say, a criterion ensuring them that the decision comes from the will of the group as such.

In other words, the current explanation of the phenomenon illustrates the second type of psychology: people are supposed to favour the unanimity rule under the effect of impersonal cultural forces. They perceive themselves as a mere part of the group under the influence of these forces and consequently find natural to consider that a decision is worth being enforced only when the group has given an unambiguous verdict.

Two objections, though, can be raised against this type of explanation. *First*: that the unanimity rule is another name for the veto right. In societies where the unanimity rule prevails, every single member has the capacity of saying no. In that sense, his opinion has more weight than under any other rule, as the majority rule. *Second objection*: in traditional village societies operating under the unanimity rule, a collective decision is often taken in an atmosphere of conflict and after a long discussion, as the very notion of *palaver* indicates. On the whole, the causalist explanation introduces an assumption which is both *heavy*, since it postulates the existence of cultural forces able to implant representations in the mind of people, and *controversial*, since it accepts wrongly that individuals have less power under the unanimity than under the majority rule. Finally, it ignores and fails to account for the well-known phenomenon of the palaver which always appears as soon as the unanimity rule is adopted in any context.

Popkin made the point that Rational Choice Theory provides a much more easily acceptable explanation. He started from Buchanan's and Tullock's (1962) theory, a theory which explains why in our societies the majority rule is very generally treated as a basic decision making rule. This derives from the fact that the organization of a collective decision procedure is exposed to two types of costs. If it is decided that all members of the group should agree before a decision is considered as lawful, the time used in the decision-making process will likely be long. At the other extreme, if a «dictator» were appointed in the sense that his decisions would be treated as lawful, the collective decision process would be quick. In other words, the higher the proportion of people in the group whose agreement is requested before a decision is considered as lawful, the longer the time taken to reach a decision. On the other hand, the higher the proportion of people in the group whose agreement is requested before a decision is considered as lawful, the smaller the proportion of people exposed to the risk of seeing a decision enforced with which they would disagree. The theory assumes furthermore that the two curves representing the two kinds of costs as a function of the number of people whose agreement is required, are convex. This assumption derives from the consideration that, when all agree on a decision except a last opponent, convincing him will likely be more difficult than convincing less stubborn opponents. So, it seems plausible to assume that the curve relating the number n of people whose agreement is required to the first type of cost, i.e. the length of the decision process, is a monotonous increasing convex function of n . As to the other cost, similar reasons lead to assume

that the curve representing it as a function of n is convex. Under these assumptions, the sum of the two costs will in broad conditions be minimum when n is equal to half the total number of members. This explains why the majority rule is the rule normally accepted and considered legitimate in many situations.

But there are also circumstances where it is not, even in our societies. Thus, by contrast with the adoption of an ordinary new law, a change in the constitutional law will require, say, a two third majority because, there, the time curve is moved to the right: given the importance of the decision, one wants to be sure that it is not imposed to a too great number of people even if the decision takes time. As to the decision rule on the battle field, it is «dictatorial» in all armies, because here the time curve is moved to the left, since taking too much time to reach a decision may be fatal in a war situation.

Popkin proposes to consider that the question as to why village societies use generally the unanimity rule is a special case of this general theory. The Vietnamese traditional village societies are characterized by an economic regime of the subsistence type. There is no economic surplus; the production covers at most the local consumption. If the majority rule, say, was accepted, nothing would prevent for instance the group from forbidding gleaning. Now, the poorest slice in the population lives from gleaning. So, under the majority rule, some members of the society would be heavily threatened. Such a rule would consequently not be considered legitimate. In other words, the curve representing the second type of costs, the «external costs», is moved far to the right in the case of this type of society. These costs are very heavy except if all agree. On the other hand, time has much less value in traditional village societies than in modern societies, since the subsistence economy requires from the producers only a small part of the time available to them, so that the length of time spent to take a decision is much less important than in modern societies.

Of course, it cannot be supposed that the members of the considered societies have effectively this theory in mind. Still, one can legitimately suppose that they have in mind a rough intuitive version of the theory. They know well that a minority lives from gleaning and that a majority could decide to outlaw this traditional institution for instance. Leo Moulin (1953) has shown that the majority rule was invented in the convents of the Middle Ages on the basis of considerations which anticipated literally on Buchanan and Tullock's theory. On the whole, Popkin's theory is much more easily acceptable than its competitor: it includes no heavy assumption, no contradiction with data, no arbitrary assumption; it is congruent with the relevant data.

So we have here a case where shifting from a Type II model to a Type I, from a model interpreting people's behaviour as the effect of «cultural forces» to a model interpreting it as rational in the sense of the RCT appears as a definite scientific progress. Given the criteria generally used to evaluate a theory, the RCT appears as much better than the alternative one.

Is Type I general?

Should we consider the behavioural theory of Type I as general, though? My answer is no. For the reason that a number of phenomena appear as reluctant to be explained by this model, in spite of the sustained efforts made to explain them by Type I theories.

The case of voting is illuminating in this respect. RCT assumes that people's actions are always motivated by the objective of generating some outcome. RCT assumes, in other words, an instrumental or consequential theory of action: that actors are motivated by the consequences they imagine they are going to produce.

Now, no single vote is likely to have any effect on the outcome of the poll. Why do people still vote? A possible answer is that they tend to have a biased view as to the actual importance of their vote. But in that case it is necessary to explain the origin of this bias. Does it come from some psychological force that would lead most people to believe falsely that their vote has an impact on the outcome while it has actually no impact? I doubt that such a theory has ever been taken seriously. Though some people may have a wrong view of the influence of their ballot, it is very unlikely to suppose that most voters have.

An alternative theory would be that people still bet that their ballot can have an influence (Ferejohn et Fiorina, 1974). The 2000 American presidential election has shown that it can. But even in this very particular case the probability of a particular ballot having an influence was *a priori* close to zero. So that, if we want to elaborate on this assumption, we have to suppose that the utility to most people of the election is so high that it compensates, so to say, the infinitesimally weak probability of any ballot being pivotal. This is the type of argument Pascal used when he tried to convince the atheist that he should bet that God exists. Since not believing in the existence of God if God actually exists would condemn the atheist to damnation, it would be rational for him to believe in God even if the probability that God exists would be as close to zero as wished. As a mathematician, Pascal plays in this famous argument with the mathematical paradox that the infinite multiplied by a quantity as close as wished to zero without being zero is equal to the infinite. But are we ready to assume that the eventuality of candidate X being elected as president rather than Y appears to *all* voters as important as spending the eternity in the flames of Hell?

Other theories belonging to Type I have been tried, such as the theory which assumes that abstention has a cost. The theory is highly speculative, though. In some areas and circumstances, people may want effectively to avoid the eventual reprobation induced by abstention because they are under the control of their neighbours. In others, they are not. Thus, in Paris, nobody knows whether I have voted or not.

Many other theories could still be mentioned, which attempt at reconciling RCT with the fact that a non negligible number of people do actually vote.

Since none of the theories belonging to Type I appears as really convincing, they lead to the conclusion that most people see that voting is a non consequential behaviour and that their vote is not going to make the difference. Common sense tells us that people vote because they have a positive attitude toward democracy and hold elections as an essential ingredient of democracy, so that they vote if they think, say, that X is definitely better than Y. This explanation takes us away from RCT. But it is a good theory in the sense that all the statements it includes are acceptable and that they lead moreover to consequences congruent with data. In particular, it can explain satisfactorily the variation in the rate of participation: it predicts for instance that the rate of participation will generally be lower in the case where many citizens have the impression that none of the candidates offered to them is palatable.

An argument often heard on RCT is that it is better because it allows expressing behaviour by simple equations. As biology, the most active and influential science to day, however shows, equations are neither a necessary nor a sufficient sign of a scientific approach. The time when mechanics were considered as the model of all sciences is over. The aim of science is to describe reality as it is, not necessarily to put it into equations. Now, no voter will recognize himself in the theories proposed by RCT on voting.

Voting behaviour is not a singular case. RCT fails in many other examples, as the ultimatum game. In this game, A is asked to propose as he wishes to share 100 € between himself and B. B can only accept or reject A's proposal. If B accepts A's proposal, each receives the part of the 100 € A proposed to grant respectively to A and B. If B refuses A's proposal, the 100 € remain in the experimenter's pocket. If RCT were valid here, A would propose something like «70 € for me, 30 € for B», since B would accept and A would get a satisfactory outcome from his proposal. In fact, A proposes in most circumstances equal sharing.

In the same spirit, Frey (1997) relates an interesting experiment where two groups are proposed to decide whether they would accept nuclear waste on the territory of their community. The local authorities would be proposed a monetary compensation in the case of the first group. No compensation would be proposed in the case of the second group. Against the prediction of RCT, the rate of people accepting the nuclear waste is greater in the case where no compensation is proposed.

These counter examples are taken among many possible other examples. They are probably sufficient to show that RCT cannot be held as general for the obvious reason that many actions are not inspired by their utility to the actor, while other actions appear even as not consequential. In other words, RCT is valid exclusively when restrictive conditions are met.

Amartya Sen has once said that the social actor as described by the RCT is a «rational idiot». He meant that a real human being whose constitution would correspond to the postulates describing the *homo economicus*, the social actor of RCT, should for instance not vote. As a matter of fact, I know a colleague who, as a true believer in RCT, actually does not vote. I always won-

dered whether he was serious or joking. The social actors of RCT are «rational idiots» in the sense that in a prisoner's dilemma game they would easily generate a catastrophe. Axelrod (1984) has tried to make them a little less idiot in the case of repeated prisoner's dilemma games in the sense that in that case, they can play the tit-for-tat strategy, i.e., offer the other player to cooperate. If the other accepts and if none of them is tempted by the short term benefits that a treason strategy would generate, the catastrophe will be avoided. In real repeated prisoner's dilemma games it can occur though that actors cannot use the axelrodian strategy, as illustrated by the case of the arm race between the US and USSR in the second part of the 20th century. Here, the two actors could not help behaving as rational idiots and went on a long time with the arm race until a brilliant innovation brought the arm race to an end: The Strategic Defence Initiative, the Star War. It had the effect that the other player could not follow. By so doing, the first player was not «rational», but rather «intelligent», in the sense that he destroyed the repeated prisoner's dilemma game. And he was able to do so because he had built a good theory of the situation: the other player is exhausted, if a program of arm development is devised which would appear to him as too costly, he would have no other choice than to propose to end the game.

The difference between a rational and an intelligent actor is that the latter has a good theory of the situation and of the ways able to modify the situation. He is «rational» in the sense, not of RCT, but in the sense where a scientist is rational. He is *cognitively* rational.

Choosing Type II instead of Type I?

The failures of RCT are probably to a large extent responsible for its unpopularity among sociologists. But this unpopularity is also due to the fact that RCT contradicts their favourite model. For many sociologists and anthropologists accept the idea that explaining behaviour means identifying the psychological, cultural or social forces causing behaviour.

Thus, Clifford Geertz states in an influential article that sociologists and anthropologists should follow literally Montaigne's teaching, when he claims that, as norms and values are different through cultures and as people tend in all cultures to find them as valid, people believe in the norms and values dominant in their society under the action of these forces which sociologists and anthropologists call *socialization effects*.

Socialization effects do exist beyond doubt. But they are too easily evoked by sociologists. First of all, their existence is generally demonstrated in a circular fashion. Then, socialisation often does not work. Thus, Weber (1988 [1920-21]) explains that in the Roman Empire, most roman civil servants and soldiers had been educated in the traditional polytheistic roman religion. When monotheistic cults appeared in the Roman Empire coming from the Orient, as the Mithra cult and then Christianity, they were strongly attracted by them. So, socialization apparently did not work properly in this case. Simmel (1892)

observed on his side that socialization can easily have opposite effects: a person who has had a hard youth can as well become egoistic as generous; a person educated liberally as a youngster can educate his own children in a liberal or authoritarian fashion.

Behaviour is also often explained in the social sciences by *psychological* forces. Thus, cognitive psychologists have observed that people consider as evident solutions to simple mathematical or statistical problems which are false in the eyes of the mathematician or of the statistician. In many experiments, people appear as underestimating or overestimating probabilities, as seeing a correlation where there is none, etc. Psychologists explain this phenomenon by evoking psychological *biases*. In most cases, the existence of these biases is grounded in a circular fashion by the effects they are supposed to explain, though. Sometimes, an effort is made to guess where these biases come from. Thus Tooby and Cosmides (1992) claim that the biases in statistical inference they registered in some experiments come from the hunting experiences of men in the early stages of mankind. Such assumptions are highly conjectural. Moreover, even if they could be empirically tested, they would explain exclusively *some* particular biases. They explain the under- or overestimation of probabilities, but not for instance that people in some experiments see a correlation where there is none. Still, these writers have well perceived that the biases themselves needed to be explained.

Dawkins (1976) theory of *memes* is fashionable today. Genes explain why some biological features are transmitted from one individual to another. Dawkins propose to introduce the notion of cultural genes, which he calls *memes*, and to explain cultural evolution by the fact that memes are transmitted from one individual to another, in the same way as biological evolution is explained by the fact that genes are transmitted from one individual to another. Dawkins theory develops actually in a new vocabulary a variant of the classical theories which propose to explain behaviour by socialization effects, or by an imitation instinct, as the theory developed by Tarde in his *Laws of imitation*. The attractiveness of Dawkins' theory seems mainly due to the fact that it took benefit of the success, influence and prestige of the well established theory of biological evolution and that it rewords theories popular among sociologists and anthropologists.

The type II model advertised by Geertz, by Dawkins and by many sociologists and anthropologists according to whom behaviour should be explained by some psychological, biological or cultural forces is valid in many circumstances. But, as the Type I model, it is not general.

In some cases it works beyond doubt. To take an obvious example: when Damasio (1994, 2003) observes that some lesion in the brain provokes a brutal change in the behaviour of an individual, we have little doubt that the change in behaviour is due to the lesion in the brain. Thus, Damasio describes the case of a sociable, open, lively and pragmatic mining engineer who made a mistake while he manipulated an explosive, had his brain damaged and became whimsical and egocentric. In that case, we have no doubt

that the change in the state of the brain is the cause of the change in the behaviour.

We are obviously very far from this unambiguous situation when a sociologist, an anthropologist or a psychologist explains the behaviour of adults by some experience in their childhood or by the effect of socialization, say. Here, the cause is in most cases conjectural. As a matter of fact, the concept of socialization is in most cases descriptive rather than genuinely explanatory. When a sociologist states that Roman peasants believed in many gods because they were socialized in the traditional polytheistic Roman religion, this amounts to saying that, as their parents, they were themselves educated in this religion. This statement registers a fact rather than a causal relationship. Moreover, as I reported earlier, this case raises the question as to why socialization worked properly as far as Roman peasants were concerned, but not with the soldiers and the civil servants.

Or, to come back to the examples of application of the Type II model in cognitive psychology, saying that a subject underestimates some probability or sees a correlation where there is none under the effect of a bias, *because* he is exposed to some bias, does not say more than saying that he is mistaken. The notion of *bias* has the status of a stenographic *descriptive* notion. It could cease to be a mere word if it could be shown that biological evolution, say, has produced a wiring of the brain explaining the bias. Otherwise we are left with a black box.

Generally, words like *meme*, *bias*, *frame*, *socialization effects*, *habitus*, *social representation* and the other concepts of the same type are descriptive rather than explanatory. This point was already convincingly made by Popper (1976) in the case of the concept of *frame*. Such words postulate conjectural biological or evolutionary processes which take the social scientist beyond his realm of competence.

At this point of our discussion, we can conclude that, as shown by Popkin's example and many others, models of Type I can lead to illuminating theories. On the other hand, the Type II models are also useful. Though we know little about their nature, there are such things as socialization effects. But neither Type I nor Type II models can claim to be valid *generally*. Becker (1996) makes a good point when he writes that RCT can be scientifically more acceptable than theories explaining behaviour by cultural, psychological or biological forces. But he is wrong when he makes this statement general and also when he assumes that we would be faced to the dilemma of choosing between Type I and Type II models. In the provocative vocabulary of Amartya Sen, we are not condemned to see the *homo sociologicus* either as a «rational idiot» being rational exclusively in the instrumental sense or as an «irrational idiot» led by forces beyond his control.

I will leave aside the eclectic writers who think they can solve the dilemma at a low cost by assuming that people are rational in the sense of the type I model when they choose the appropriate means to reach their objectives, but irrational in the sense of the type II model when they select an objective or follow a norm.

Type III model

By difference with the behavioural theory of Type I, the Type III theory considers in Schütz's vocabulary *Weil-motive* beside *Wozu-motive*. By difference with Type II, it uses only marginally the notion of forces. In a nutshell, its main difference with the other two models lies in the fact that it uses *ordinary psychology*, rational theory in Nisbet's sense, as its favourite theory of behaviour.

Before discussing the model from a theoretical viewpoint in more detail, I will illustrate it. I will examine carefully an example from Durkheim. This will have the incidental advantage of stressing, against a current view, that Durkheim's actual analyses do not use Type II, but rather Type III models. I already mentioned this point when I recalled Durkheim's analysis of the effect of crisis situations on suicide rates. Hirschman (1982) has elaborated on the psychological mechanism used by Durkheim in this analysis, using also a Type III approach.

In a particularly brilliant analysis, Durkheim wonders why people execute these particular actions which we call magical rituals. They do so because they believe in the effectiveness of magical recipes. A difficult enigmatic question on which many pages have been written is then: why do people believe that rain rituals have the capacity of producing rain, while objectively they have not. While many explanations use controversial psychological assumptions, Durkheim succeeds solving the question by using rational ordinary psychology.

Magical rituals as rain rituals are instrumental, in the sense that they are supposed to produce an effect, rain fall. Of course rain fall is wished because of the effect of rain on the growth of plants. This instrumental side of magical practices raises no particular question. The puzzling aspect of this behaviour lies rather in the belief that magical rituals would be objectively useful.

An easy solution to this problem is to use a Type II model to explain the belief: to assume for instance that the magician and his customers have been socialized to believe that rain rituals are effective. This type of solution was proposed notably by Lévy-Bruhl (1960 [1922]): he postulated the existence of forces he christened «primitive mentality». They would make that the brain of the «primitives», in the vocabulary of Lévy-Bruhl's time, is submitted to inference rules different from ours. Lévy-Bruhl supposes in other words that rationality is culture dependent. The same type of solution was endorsed in modern days by Needham (1972), Beattie (1964) or Sahlins (1995). According to the theories they developed, the rules of inference which we use and consider universal would in fact be particular and characteristic of our own culture. Huntington (1996) goes even as far in this vein as to say that the very notion of universality is a particular feature of Western culture.

With his deep sense of methodology, Wittgenstein (1967) saw that this explanation of magical beliefs is unsatisfactory and proposes an alternative one: that Westerners would wrongly assume that the primitives would believe

in the effectiveness of magic. According to Wittgenstein, magical behaviour should be interpreted as *expressive* rather than *instrumental*: the «primitives» would not really believe in the efficiency of their rituals. They would rather express by their rituals their desire that rain falls.

Wittgenstein's theory is definitely more palatable than the theory developed by Lévy-Bruhl and his followers. It uses ordinary psychology and belongs to our Type III. Wittgenstein was conscious that the widespread Type II theories of magical rituals were scientifically unsatisfactory. Unfortunately, his theory is false in the sense that people reject with force the idea that magical rituals would have no effect: they refuse the expressive interpretation of their magical beliefs. According to Evans Pritchard (1968[1937]), when a Zande falls on a root in the bush, he shows when interviewed that he knows that his fall is the effect of mechanical forces, but he sees it at the same time as the effect of «magical» factors. He knows, as we do, that he would plausibly not have fallen if he had been more careful, but also if he would have been more «lucky». As Evans Pritchard says, this combination of objective and magical explanations is common to the Azande and to Westerners: we also explain some states of affairs by luck. Evans Pritchard uses here ordinary psychology.

So, «primitive» people are convinced that their rituals have an instrumental value. To be able to endorse Wittgenstein's answer, we should be ready to assume that the «primitives» do not believe effectively what they say they believe. But in that case we would have to explain this false consciousness. This is more easily said than done. Moreover, Wittgenstein's theory is incompatible with many observations, as Horton's (1993, 1982): in the English speaking countries of black Africa where he had conducted his investigations, he observed that people believe in magical practices even when they have converted to Christianity. When asked why, they answer that Christianity is good from a soteriological viewpoint, but that it has the defect of missing the magical remedies offered by the magical rituals derived from the animistic religion from which they came.

Both Weber and Durkheim and their followers, as Evans Pritchard or Horton, propose a Type III solution to the problem of the explanation of magical rituals which is much more acceptable than its competitors.

Weber (1976 [1922]) encapsulates the theory he defends by a compact sentence: «to the primitive fire making is as magical as rain making». This means that *to us* Westerners, there is a difference between fire- and rain making. The operation thanks to which the primitive makes fire, rubbing two pieces of wood, is grounded on objectively valid laws. By contrast, we see rain rituals as magical because they are not objectively grounded. *To them*, to the «primitives», the distinction is meaningless: they see fire and rain making as equally magical. Weber means here that we make a difference because we know the laws of the transformation of energy. That the primitives have not learnt them is clear. That they would have discovered them intuitively is implausible. In other words, the «primitives» have probably discovered empirically the way of making fire, but the theory thanks to which they explain this technique

is likely as magical as the theory supporting their rain making techniques. There is no reason to suppose with Lévy-Bruhl that the brain of the «primitives» is wired in a way different from ours. It suffices to take seriously the obvious fact that they have not learnt physics on school benches, and that they have no reason to see intuitively laws that Westerners have discovered after many centuries. There is no reason for them to make a difference between the way they explain fire making and rain making, while for us there is a strong reason.

Durkheim did not know Weber's sentence, but he elaborates on the same ideas. When the «primitives» grow some plant, they use much empirical know-how which has been discovered and transmitted from one generation to the next. But they also need to know why plants grow, wither and die. This cannot be determined empirically. So they need to forge some «biological theory». As they have no access to science in the modern sense, they draw this biological theory from the religious interpretation of the world treated as legitimate in their society. For religion represents the legitimate source of knowledge, as science in Western societies. As to the magical recipes, they are technical procedures derived from this religious theory of the world. So, magical recipes are comparable to the technical recipes Westerners draw from science. The obvious difference is that the recipes derived from science are more reliable. Does not this show, as Lévy-Bruhl postulates, that the brain of the primitives is wired otherwise; that they are not, say, as sensitive as Westerners to contradiction?

Not only, objects Durkheim, they dislike contradiction, but they treat it as any modern scientist does: by inventing auxiliary assumptions. We know from the Duhem-Quine thesis that, when a theory fails to explain some new data, the normal reaction of any scientist is to invent auxiliary assumptions rather than to reject the theory, because the theory explains many data and also because he does not know *a priori* which element in the theory is wrong, so that it is reasonable for him to assume that an auxiliary assumption will likely reconcile the theory with the data. This is what the scientists do, as the history of science shows. This is what they should do, as epistemology shows. This is also what the magician does. In the case where the magical rituals fail, he will assume, say, that they have not been executed exactly as they should have been, etc. Durkheim anticipated very clearly the Duhem-Quine thesis, a thesis widely accepted by historians and philosophers of science.

Another objection raised by Durkheim himself is that, being ungrounded, magical recipes fail in 50% of the cases. Durkheim's answer is that, as the rain rituals are practised in the period of the year where rain is more likely to fall, a correlation between the two variables *days with/without rituals* and *days with/without rain* will be normally observed. The correlation will of course be spurious. But Westerners ground also often their beliefs on spurious correlations.

On the whole, Durkheim's theory is much more acceptable than the alternative theories of magical rituals available on the market. Moreover, it explains easily and convincingly a number of puzzling data: why magical practices were much more frequent in Europe in the 16th or 17th centuries than in the 14th cen-

tury, and more frequent in the modern parts of Europe: more frequent in Northern than in Southern Italy or than in Spain, say. It explains on the whole many comparative data, some of them have been discovered a long time after Durkheim proposed his theory, for instance by Thomas (1973). On the whole, Durkheim's Type III theory works beautifully where several Type II theories have failed.

Durkheim uses here ordinary rational psychology. The primitive people need recipes to grow their crops. Durkheim's central thesis is that they are rational in the same way as scientists are. To the best of their knowledge, they have strong reasons to believe in the effectiveness of their magical recipes.

Evans Pritchard's (1968 1937) or Horton's (1993) work are fascinating because they show also, in agreement with Durkheim's intuition, that the primitives use the same inference rules as any scientist. When they take their oracles, the Azande manipulate logical calculus with a greater virtuosity than any western student. Cole and Scribner (1974) have shown that in Liberia, the «primitives» master the rules of logical inference as well as any American student.

I have tried to show elsewhere, taking several examples from classical and modern sociology, that Type III models are used in many illuminating analyses (Boudon, 1998a, 2003, 2005). The greatest sociologists, Tocqueville, as well as Durkheim or Weber and others, the greatest anthropologists, Evans Pritchard and others, always use Type III models. It is exceptional to see them satisfied with the explanation that some phenomenon is due to the fact that individuals are guided by psychological, social or biological forces

Type III model and Weber's methodological writings

I borrowed my main example from Durkheim as a kind of provocation. In his *Elementary forms of religious life* little remains from what he had stated in the *Rules of sociological method* and generally in his methodological writings. In the latter he prohibits any psychological explanation of sociological facts as the famous *motto* states: *explain social facts by social facts* («expliquer les faits sociaux par des faits sociaux»). In his explanation of variations in the rates of suicide, he always uses ordinary rational psychology though, as in his explanation of the reasons as to why Protestant commit suicide more often than Catholics, bachelors than married people, etc. The same is true of his analysis of magical rituals. In all these analyses, he explains the social phenomena he explores by making them the outcome of individual actions which he treats as understandable in Weber's sense: as explainable by the kind of reasons and motivations which we evoke normally in everyday life. For various reasons which I do not intend to explore here, his methodological writings have little to do with the methodology he uses effectively in his analyses.

This discrepancy does not appear with Weber, though his methodological writings are sketchy. When they are made less sketchy, we get the Type III model illustrated by the examples I have just presented, which, by difference with Type I and II models, can be considered as genuinely general.

So, let us make Weber's methodological ideas more explicit and articulated with one another.

The idea of *comprehension* means that we can in principle find the causes of any piece of behaviour or of any belief, even if at first sight it appears as strange to us, as «irrational» in the current sense of the word. This is the case for instance of magical beliefs. They give an impression of being strongly irrational. Still, as Durkheim and Weber have shown, they can be convincingly explained in the sense that their causes can be identified in a convincing way. These causes are the reasons and motivations people have to believe what they believe. The *motivations* in this case are clear: to have a sufficient crop to be able to survive. The *reasons* why they endorse the idea that magical rituals have been identified by Durkheim and Weber. The «primitives» have at their disposal theories of the world from which they draw magical recipes, exactly as we extract technical recipes from our scientific theories. They confront their theories to the data and eliminate eventual contradictions using the same devices. So, to explain an action, even apparently irrational as rain rituals, means reconstructing the motivations and reasons of the actors.

Another point of the comprehension theory is that reconstructing the reasons and motivations of an actor amounts to building a theory following the rules generally used when building any theory. Weber's example of the wood cutter makes this point clear. If I see somebody cutting wood in his yard, I will explain his behaviour by evoking some possible motivations and reasons. Maybe he is cold and wants to burn the wood in his chimney. This theory will be rejected if the weather is warm and I will have to find some other one, until I reach a theory lending to the actor motivations and reasons borrowed from the ordinary rational psychology and compatible with all the available observational data. In the case where I would fail to find understandable motivations and reasons, and only in this case, I would legitimately propose an irrational explanation as: he suffers from a compulsive trend to cut wood.

Another very important point which can be derived from Weber's notes is that the expression «being rational» can be treated as a shorthand notation meaning actually: «behaving, acting or believing on the basis of a set of reasons perceived as strong and well articulated with one another». The example from Durkheim on magical rituals provides a clear illustration of this point: the magical rituals are explained by a set of articulated reasons in the mind of the «primitives».

Another important point is that some of the reasons explaining an action can be instrumental and some not. The very distinction made by Weber between *instrumental* and *axiological rationality* shows that, to him, rationality does not necessarily take an instrumental form.

The meaning of this notion of axiological rationality has been much discussed. Sukale (1995) went even as far as to say that it is meaningless. My own interpretation is that we can and have an interest to distinguish between instrumental rationality and cognitive rationality. *Instrumental rationality* means:

looking for the best means to reach a goal. *Cognitive rationality* means: finding the good explanation of a phenomenon. As to axiological rationality, I propose to see it as a special case of cognitive rationality. If this interpretation—which is not to be found in Weber's work but which can be regarded as elaborating on his intuitions—is accepted, being axiologically rational means: finding a set of strong and strongly articulated reasons leading to a normative conclusion. I have tried elsewhere to illustrate this idea by an example taken from Adam Smith, where he wonders why the British public in his time was convinced that miners should be definitely paid more than soldiers. He explains this puzzling collective belief by showing that this conclusion was derived in the mind of the public by a set of strong reasons, strongly related to one another: a brilliant example of «axiological rationality» (Boudon, 1998b). I will evoke contemporary illustrations in a moment.

A further important point is that the resources of the so-called deep psychology should not be mobilized because the *forces* thanks to which it proposes to explain behaviour are in most cases conjectures which cannot be tested, so that using deep psychology takes sociological theories away from the realm of scientific theories.

The above principles lead many of the most convincing and illuminating analyses we owe to sociologists. Most of them use in other words the behavioural theory of Type III. All recognize that instrumental rationality appears in combination with non instrumental rationality, the rationality I called cognitive, a special case being what Weber christened axiological rationality.

Also, it can be shown that the phenomena which appear as unexplainable in the frame of the behavioural theory of Type I are in most cases explainable by the behavioural theory of Type III. Frey's paradox can be solved if we see that, if no compensation is proposed, people will accept more easily the nuclear waste on the territory of their community, because they tend to develop the theory that they will consider themselves and will be considered as able to accept a sacrifice for the sake of the general interest. By contrast, when compensation is proposed, people tend to develop the theory that a bargain is proposed to them and that they feel entitled to judge that the bargain is actually not very interesting. Because the proposal generates two theories in the mind of people, those to whom compensation is offered accept the proposal less frequently. That the same observation was made in two different contexts tends to show that the proposal tended to induce the same system of reasons in the people's minds in the two contexts.

I have tried to show that the findings of cognitive psychology generally explained by Type II models could also easily be explained by Type III models: one need not to postulate that the wrong answers are produced by biases with a highly conjectural origin (Boudon, 1996). On the whole, it can be shown both by the *a priori* arguments I have tried to develop, but also *a posteriori* by actual analyses of empirical data, that the difficulties met by the behavioural theory of Type I and II can in many cases be solved by using the behavioural theory of Type III.

The importance of the discussion

To conclude, I will point to the fact that the former discussion is important not only from a scientific but also from a general viewpoint.

Suppose we want to explain why most of us have a negative reaction toward female genital mutilation. An explanation using the behavioural theory of Type II would assume that cultural forces and socialization effects make that we have a feeling of cruelty, while other cultural forces make that this practice was and still is widely adopted in many societies without arousing the same feeling. Such explanations are proposed by important writers as Shweder (2000) or Geertz (1984). It is difficult to imagine an explanation of this difference which would use the behavioural theory of Type I. An explanation using the behavioural theory of Type III would assume, say, that female genital mutilation is a means of social control and that as soon as a means of social control less cruel than another is found we tend to require that the less cruel means is adopted. Durkheim (1960 [1893]) developed this kind of explanation when he made the point that social control tends to become over time more and more liberal. Tocqueville (1986 [1840]) developed the same kind of model when he remarked that our moral sensibility toward slavery has irreversibly changed. «The greatest minds in Greece and Rome, he states, never imagined that slavery could be abolished». They accepted it because they thought it was functional: indispensable to the economic system. Then, forms of social organization appeared which did not have recourse to slavery. From that moment a negative feeling toward slavery developed. In the 18th century, Montesquieu wrote that it was impossible to abolish slavery in the Antilles because this would deeply disturb the sugar market, but he considered it as «against nature». In other words, the feelings and reactions toward an institution result from a set of articulated reasons. The Greek saw slavery as indispensable. It was no more possible to endorse this view once it had been shown that it was not. Still at the end of the 19th century universal vote was considered as a dangerous institution, until it was shown it was not. All these familiar examples illustrate the tension between axiological and instrumental rationality. The political men at the end of the 18th century had strong axiological reasons to condemn slavery and strong instrumental reasons not to abolish it.

I will evoke a final example to illustrate the importance of the cognitive approach to rationality which characterizes the behavioural theory of Type III. The example deals with differences in the moral reactions and sensitiveness of the public to current events.

In an article of December 24, 2004, the *International Herald Tribune* raises an interesting sociological question. It wonders why very serious violations of the law can arouse little reaction in a given context while much less serious violations can generate more reprobation in another. The article reports that «in Germany, the deputy police chief of Frankfurt was found guilty of threatening to torture the captured kidnapper-murderer of an 11-year-old child», though he «had every reason to believe that a kidnapped child's life hung in the balance,

and, in this sense, threatening torture of the young man who had seized him was justified». By contrast, while the US has used torture at the Abu Ghraib prison in Baghdad and in Guantánamo, «there is no strong sign that the public is outraged by these actions. The public reaction can be explained in the two cases by reasons. But, as the German and the American contexts are different, the reasons were different. The decision of the German judge was explicitly justified by reference to the German past.» By contrast, «the American public, victimized by a terrible act of mass murder and witness to the daily terror bombings in Iraq, seems inured to the torture being conducted, supposedly, in its interest.»

The same type of cognitive contextual rationality applies to the difference in the treatment of Stalinism and Hitlerism. One can be an ex-Stalinist not an ex-Nazi, because Stalinism is still analyzed by many minds as a perversion of morally and socially acceptable ideas, by difference with Hitlerism which is seen as inspired by devilish ideas.

Men are neither rational idiots nor irrational idiots. They behave in the way they do because they have all kinds of theories on all kinds of questions. In many cases they have to muddle through the complexity of the questions they are faced to. Durkheim's idea that the magician, the ordinary man and the scientist behave in the same way is a deep idea. It sketches an important general theory of the *homo sociologicus*.

My final guess is that the social sciences are less exciting now than in Tocqueville's or Durkheim's and Weber's time because macroscopic sociology has become impressionistic, while it had a definite scientific character in Tocqueville's, Durkheim's or Weber's work. And the main reason why they produced convincing macroscopic theories is that they used a theory of behaviour of Type III rather than of Type I or II, the dominant theories of behaviour today.

References

- AXELROD, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- BEATTIE, J. (1964). *Other Cultures*. London: Cohen & West.
- BECKER, G. (1996). *Accounting for Tastes*. Cambridge: Harvard U. Press.
- BOUDON, R (1996). «The cognitivist model : a generalized Rational Choice model». *Rationality and Society*, 8, 2, 123-150.
- (1998a). «Limitations of Rational Choice Theory». *American Journal of Sociology*, 104, 3, 817-828. Reproduced in: TERRELL & MEIER (eds.). *Readings in Deviant Behavior*. Harcourt College Publishers, Fort Worth, Tx., USA.
- (1998b). «Social mechanisms without black boxes». In: HEDSTRÖM, Peter; SWEDBERG, Richard (eds.). *Social mechanisms: an analytical approach to social theory*. Cambridge: Cambridge University Press, 172-203.
- (2003). «Beyond Rational Choice Theory». *Annual Review of Sociology*, 29, 1-21.
- (2005). *Tocqueville for Today*. Oxford: Bardwell.
- BUCHANAN, J.; TULLOCK, G. (1962). *The Calculus of Consent*. Ann Arbor: University of Michigan Press.

- COLE, L.; SCRIBNER, S. (1974). *Culture and Thought: a Psychological Introduction*. New York: Wiley.
- COLEMAN, J. (1986). *Individual Interests and Collective Action: Selected Essays*. Cambridge: Cambridge U. Press.
- DAMASIO, A.R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset/Putnam.
- (2003). *Looking for Spinoza: Joy, Sorrow and the Feeling Brain*. Harcourt.
- DAWKINS, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.
- DURKHEIM, E. (1960 [1893]). *De la division du travail social*. Paris: Presses Universitaires de France.
- (1967[1897]). *Le suicide, étude sociologique*. Paris: Presses Universitaires de France.
- (1979 [1912]). *Les Formes élémentaires de la vie religieuse*. Paris: Presses Universitaires de France.
- EVANS-PRITCHARD, E.E. (1968[1937]). *Witchcraft, Oracles and Magic among the Azande*. Oxford: Clarendon Press.
- FEREJOHN, J.A.; FIORINA, M.P. (1974). «The paradox of not voting: A decision theoretic analysis». *The American political science review*, 68, 2, juin, 525-36.
- FREY, B.S. (1997). *Not just for the money: an economic theory of personal motivation*. Cheltenham: Edward Elgar.
- GEERTZ, C. (1984). «Distinguished Lecture: Anti anti-relativism». *American anthropologist*, vol. 86, n°2, 263-278.
- HIRSCHMAN, A.O. (1982). *Shifting Involvements. Private interests and Public Action*. Princeton: Princeton University Press.
- HORTON, R. (1993). «Lévy-Bruhl, Durkheim and the Scientific Revolution». In: HORTON, R. *Patterns of Thought in Africa and the West*. Cambridge: Cambridge University Press, 63-104.
- (1982). «Tradition and Modernity Revisited». In HOLLIS, M.; LUKES, S. *Rationality and Relativism*. Oxford: Blackwell, 201-60.
- HUNTINGTON, S. (1996). *The Clash of Civilizations and the Remaking of the World Order*. New York: Shuster and Shuster.
- LÉVY-BRUHL, L. (1960 [1922]). *La Mentalité primitive*. Paris: Presses Universitaires de France.
- MOULIN, L. (1953). «Les origines religieuses des techniques électorales et délibératives modernes». *Revue internationale d'histoire politique et constitutionnelle*, avril-juin, 106-148.
- NEEDHAM, R. (1972). *Belief, Language and Experience*. Oxford: Blackwell.
- NISBET, R. (1966). *The Sociological Tradition*. Glencoe, Ill.: The Free Press.
- OVERBYE, E. (1995). «Making a case for the rational, self-regarding, "ethical" voter... and solving the "Paradox of not voting" in the process». *European Journal of Political Research*, 27, 369-96.
- POPKIN, S. (1979). *The Rational Peasant. The Political Economy of Rural Society in Vietnam*. Berkeley: University of California Press.
- POPPER, K. (1976). «The Myth of the Framework». In FREEMAN, E. (ed.). *The Abdication of Philosophy : Philosophy and the Public Good*. La Salle, Ill.: Open Court, p. 23-48.
- RADNITZKY, G. (1987). «La perspective économique sur le progrès scientifique: application en philosophie de la science de l'analyse coût-bénéfice». *Archives de philosophie*, 50, avril-juin, 177-198.

- SAHLINS, M. (1995). *How «Natives» Think. About Captain Cook for instance*. Chicago: Chicago U. Press.
- SHWEDER, R. A. (1977). «Likeliness and Likelihood in Everyday Thought: Magical Thinking in Judgments about Personality». *Current Anthropology*, vol.18, n°4, décembre, 637-659.
- (2000). «What about “female genital mutilation” and why understanding culture matters in the first place». *Daedalus*, 129, 4, 209-232.
- SIMMEL, G. (1892). *Die Probleme der Geschichtsphilosophie*, München: Duncker & Humblot.
- SUKALE, S. (1995). «Introduction to Max Weber». *Schriften zur Soziologie*. Stuttgart: Reclam.
- THOMAS, K. (1973). *Religion and the Decline of Magic*. Harmondsworth: Penguin Books.
- TOCQUEVILLE, A. de (1986 [1840]). *La Démocratie en Amérique*, vol. II. In *Tocqueville, De la démocratie en Amérique, Souvenirs, L'Ancien Régime et la Révolution*. Paris: Robert Laffont, Bouquins.
- TOOBY, J.; COSMIDES, L. (1992). «The psychological foundations of culture». In: BARKOW, J.; COSMIDES, L.; TOOBY, J. (eds.). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- WEBER, M. (1976 [1922]). *Wirtschaft und Gesellschaft*. Tübingen, Mohr, 3 vol.
- (1988 [1920-21]). *Gesammelte Aufsätze zur Religionssoziologie*. Tübingen, Mohr, 3 vol.
- WITTGENSTEIN, L. (1967). «Bemerkungen über Frazer's *The Golden Bough*». *Synthese*, 17, 233-253.