DENIS BOUYSSOU

DECISION-AID AND EXPECTED UTILITY THEORY:
A CRITICAL SURVEY

INTRODUCTION

Expected utility theory can be viewed in two different ways. First, as a formal mathematical theory providing representation and uniqueness theorems for preference structures. Secondly, as a decision-aid model, seemingly resting on the formal theory and which emerged in the late 50's: what I will call decision analysis. It is not easy from a historical point of view to understand how a purely formal theory designed to handle mixed strategies in Game Theory gave rise to a decision-aid model. Nevertheless the best-known argument in favour of decision analysis is that it is logically sound and axiomatically based. The aim of this paper is to investigate the links between these two levels, in order to clarify the way the decision-aid model works in practice. After a brief description of the main features of the formal model, I will try to see to what extent the decision-aid model is based on it and how it actually works. In the last section, I will review empirical studies dealing with the decision-aid model.

1. THE FORMAL THEORY

The purpose of expected utility theory (leaving aside the Subjective Expected Utility Theory) is to establish a numerical representation of preferences between probability distributions such that:

\[ a \succ b \iff E(u, a) > E(u, b) \quad \forall (a, b) : \mathcal{X} \text{ where} \]

\[ E(u, a) = \sum_{x \in \mathcal{X}} u(x) \times p_a(x) \quad \text{in the discrete case} \]

(P being a binary relation on a set of probability distributions \( \mathcal{X} \) on a set \( X \)).

The theory specifies axioms which, if verified, produce the desired representation. To attain an expected utility re-

presentation, we need (cf. Fishburn (1970) or (1982)):

- an ordering axiom A1, which makes $P$ (the strict preference relation) a weak order;

- an independence axiom A2, which makes the preference invariant to probability mixtures;

- an "Archimedean" axiom A3, which forbids infinitely liked or disliked consequences.

Schoemaker (1982) gives a wide review of the various interpretations of these axioms and their possible "test".

Several points are worth noting about this theory:

- If the underlying set of consequences has a multidimensional structure $X = X_1 \times X_2 \times \ldots \times X_n$, then, provided certain additional independence hypotheses are verified, it is possible to decompose the utility function $u$ using partial utility functions and scaling constants (cf. Keeney and Raiffa (1976) and Farquhar and Fishburn (1983) for alternative decompositions and assessment techniques).

- A large number of empirical studies lead one to consider that the axioms of the formal theory are a poor descriptive representation of the way people actually choose between risky options (see for instance the criticism of the transitivity of indifference by Luce (1956), that of the independence axiom by Allais (1953) and (1979), Ellsberg (1961), MacCrimmon (1968) and also MacCrimmon and Larsson (1979), Jaffray and Cohen (1982), Blatt (1983)). At the same time, it seems that the expected utility representation is fairly robust on a theoretical level. A large number of theorems have been proved with axioms much weaker than those originally postulated by von Neumann and Morgenstern (1947) - see for instance Aumann (1962), Fishburn (1970), chap. 9-10, Machina (1982), Hausner (1954) and Thrall (1954).

It should, however, be noted that the decision-aid model has to be based on a "double way" representation of preferences using a real valued utility function. Only such a representation allows an operational assessment of utility functions. These theoretical extensions therefore have only a limited interest for decision analysis from a practical point of view.
2. FROM THE FORMAL MODEL TO "DECISION ANALYSIS"

The primordial aim of the decision-aid method based on utility theory is to make useful the relationship of the theoretical model: \( a \succ b \iff E(u, a) > E(u, b) \), where \( a \) and \( b \) represent potential actions in the decision-aid study. In order, therefore, to be able to rank two actions, it is sufficient to have a utility function on the set of consequences \( X \) and to be in a position to evaluate the potential actions by means of probability distributions over this set. The process of reaching such an objective has traditionally been regarded as consisting of four phases:

2.1. Generating a set of goals and of actions

This involves understanding the specific nature of the decision problem, including generating the set of all potential actions and goals and describing the set of attributes which measure the level of attainment of various goals. It should be remembered that when one wishes to bring several attributes into this set, they should ideally have the independence properties needed for the utility function to be decomposed (see below).

2.2. Assessing the possible consequences of the actions under consideration

The nature of the theoretical model underlying the analysis means that a probabilistic description has to be used. When there is no available probability distribution based on frequency theory, recourse to a subjective basis justifying the theoretical model is needed. A variety of techniques have been proposed as ways of evaluating such probabilities. They are surveyed in Speltzer and von Holstein (1975) and Wallstein and Budescu (1983).

When the set of consequences has a multidimensional structure, it is normally assumed that the distributions on each of the attributes are independent in order to be able to apply these estimation techniques.

2.3. Determining the decision-maker's preference structure

This third stage is designed to quantify the decision-maker's
attitude towards risk and the possible trade-offs between the
different attributes, using a utility function. In the multi-
attribute case, the global utility function is only in prac-
tice decomposed additively or multiplicatively, though there
exist lots of other theoretically sound decompositions (see
for instance Farquhar and Fishburn (1983) who also provide
original assessment techniques).

In this connection, Keeney (1974) has proved the following
theorem:

Theorem 1: Let \( u(x_1, x_2, ..., x_n) \) be a utility function
defined over the set \( X_1 \times X_2 \times ... \times X_n \). If every subset of
\( \{X_1, X_2, ..., X_n\} \) is utility-independent of its complement,
then, if \( n \geq 3 \):

either \( u(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} k_i u_i(x_i) \) \hspace{1cm} (1) 

or \( 1 + k u(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} (1 + k_i u_i(x_i)) \) \hspace{1cm} (2) 

with \( u(x_1, x_2, ..., x_n) = 1, u_i(x_i) = 1 \) \( \forall i \in \{1, ..., n\} \)
\( u(x_1, x_2, ..., x_n) = 0, u_i(x_i) = 0 \) \( \forall i \in \{1, ..., n\} \)
\( k_i > 0, \forall i \in \{1, ..., n\} \)
with \( k \geq -1 \) and \( k \in \mathbb{R} \) solution of
\( 1 + k = \prod_{i=1}^{n} (1 + k_i u_i) \). \hspace{1cm} (3) 

When the conditions of theorem 1 are satisfied, the utility
function can consequently be determined completely by estima-
ting \( n \) partial utility functions \( u_i(x_i) \) and \( n \) coef-
ficients \( k_i \) ensuring that each function remains strictly wi-
thin the interval \([0, 1]\).

The estimation of the functions \( u_i(x_i) \) is classically ef-
fected in this framework by comparing various lotteries \( x \).
The basic question asked to the decision-maker is to express
a preference between a sure option - i.e. a gift of 50 $, and
a risky option - i.e. a lottery ticket yielding 100 $ with a
.5 probability or nothing. A utility function being deter-
mined up to a positive linear transformation, it can always be assumed that the utility of 0 $\$ is 0 and the utility of 100 $\$ is 1. The clue of this assessment process therefore lies in the search for an indifference expressed between the sure and the risky option. This indifference can be reached either by varying the amount of the sure option (the variable consequence method) or the probability in the risky option (the variable probability method). Note that these two techniques are theoretically equivalent in that they should lead to the same function. For instance, if the decision-maker prefers 50 $\$ to the lottery $\frac{1}{2} - \frac{100}{1}$, the analyst may ask him to compare either 40 $\$ to the same lottery or 50 $\$ with $\frac{0.6}{100} - \frac{0}{1}$. Once the indifference point is reached applying the principle of the expected utility leads in a straightforward manner to the assessment of one point on the utility curve.

The estimation of the coefficients $k_j$ is also based on a comparison of lotteries, which involve this time several attributes (cf. for instance Keeney and Raiffa (1976)).

It will thus be observed that the techniques used for estimating the utility functions are directly derived from the formal model. Within the expected utility theory framework, if the decision-maker is considered able to compare the lotteries - involving one or several attributes - that are submitted to him in systematically reliable fashion, then this stage does not in principle present major difficulties.

2.4. Ranking of the actions and making the recommendations

Once the possible consequences of the different actions have been assessed and the decision-maker's preference structure determined, the ordering of the actions is carried out by simply applying the expected utility principle: $a \succ b \iff E(u, a) > E(u, b)$. The recommendations must nevertheless take into account the robustness of this ranking with respect to the arbitrariness or imprecision involved in the evaluation of both the probability distribution and the parameters of the utility function.
Concrete examples of the application of these techniques may be found in Keeney (1979) and Keeney and Nair (1977).

3. THE HYPOTHESES UNDERLYING THE USE OF EXPECTED UTILITY THEORY FOR DECISION-AID

Much of the attractiveness of the "decision-analysis" techniques is due to the presence of the formal model, which constitute a "logical" basis for the techniques used.

In particular, the principal advantage of the logic underlying the formal model is that it allows one to end up with a relatively simple, numerical representation of the preferences. But, as I have already mentioned, it was not originally designed to provide a decision-aid. The present section will consequently analyse more closely the hypotheses and the techniques involved in applying the concepts of the formal model in a decision-aid context. It will expose the nature of the preference relation that is assumed to exist, then the techniques used for capturing it, and lastly the ways in which this preference relation may be "enriched".

3.1. Decision-aid and numerical representation of preferences

Axiom A1, present in all the formal theories, posits that a preference relation exists. If one's aim is decision-aid, it is essential to explore exactly what is meant by this "existence". The problem was already posed, twenty-seven years ago, by Rapoport (1956): this author argues that axiom A1 automatically excludes any notion of decision-aid. It is true that, if the decision-maker is supposed to have a clear preference relation, then it is very difficult to imagine what aid he could require. Rapoport concludes therefore that von Neumann and Morgenstern's formal theory for representing preference structures cannot serve as a foundation of decision-aid, as axiom A1 presupposes that all decision problems have been already solved.

Consequently, in order to be able to give a meaning to decision-aid techniques based on the formal theory, it is clearly necessary to interpret axiom A1 less literally.

As Fishburn (1967) and Schoemaker (1982) point out, using the formal theory to make recommendations consists essentially of
suggesting a choice between complex alternatives based on the basic preferences and "tastes" of the decision-maker. The philosophy behind this model becomes clearer in such a perspective. Axiom A1 must be interpreted as stipulating the existence of the decision-maker's "fundamental preferences and tastes" - what I will call his "basic attitudes". These basic attitudes, in conjunction with the axioms A2 and A3, are represented numerically by a utility function, which is then used to evaluate and thus rank the potential actions of the decision-aid study.

As an example of what is meant by those "basic attitudes", let us imagine that one is attempting to help the decision-maker to rank a set of actions with consequences presented in terms of probability distributions on a qualitative scale with $n$ levels $\{e_1, e_2, \ldots, e_n\}$; and that the decision-maker clearly prefers level $e_i$ to level $e_j$ whenever $i > j$. If the analyst wishes to arrive at an ordering in the form of a total preorder of all the actions considered, he will have to obtain $n - 2$ "indications" from the decision-maker on his preferences towards risk on this scale. In fact, given the axioms of the formal model, any lottery can be ordered on this scale by simply associating a probability $p_i$ to each level $e_i$, $i = 2, 3, \ldots, n - 1$, such that:

$$
1 - p_1 = e_1
$$

where $I$ represents indifference, provided that these attitudes have at least a minimal internal consistency and are such that $p_i > p_j \iff i > j$. Once these attitudes have been communicated, one can set $u(e_i) = p_i$, $i = 2, 3, \ldots, n - 1$ and $u(e_1) = 0$, $u(e_n) = 1$, which defines the utility function perfectly and allows one to resolve the original problem completely.

This apparently simple mechanism is nevertheless quite fundamental if one wishes to interpret the formal model as a decision-aid model: but this is rarely stated explicitly.

In order to illustrate this point, I will present here the
axioms shown by Keeney (1980) to be both necessary and sufficient for applying decision-analysis (see also Keeney (1982)).

D1 : There exist at least two possible actions.

D2 : The consequences of each action can be effectively identified. This may subdivided into:

D2a : The decision-maker's objectives are determinable.
D2b : Attributes may be brought into play to measure how far the objectives have been reached.

D3 : The probability of occurrence of the various consequences of each action is determinable.

D4 : The utility associated with each consequence is determinable. If the consequences of the actions are limited by \( c \) and \( c_i \), this last axiom becomes: each consequence can be allocated a real number between 0 and 1 such that:

\[
0 \rightarrow u(c_i) \rightarrow 1 - u(c_i) 
\]

D5 : The decision-maker's preferences must be:

- such that, if two actions can produce exactly the same consequences, then the one that leads to the better consequence with greater probability must be preferred;
- transitive;
- such that the preference between two actions is not modified if one replaces any consequences of one of the two by a lottery which is indifferent to this consequence.

This set of axioms evidently no longer depends directly on the existence of a preference relation between the actions, but it does remain very close to the axioms of the formal model. The interpretation of these axioms would seem to be relatively delicate in a decision-aid situation, since it refers to a preference relation between real actions. The preference relation used in D5 must therefore be seen as an "extension" of the basic attitudes referred to (implicitly) in D4 : in other words, it must be seen as an extension of a preference relation between simple, ideal actions basing the expression of the attitudes.
Less formally, "decision-analysis" can be said to depend on three principles:

1) There exist basic attitudes to the problem in the decision-maker's mind which it is possible to discover by comparing simple ideal actions (i.e. lotteries).
2) These attitudes conform to the axioms of the formal model.
3) The extrapolation of these attitudes to real actions, often of considerable complexity, provides an adequate basis for making recommendations.

The assessment techniques mentioned in § 2) give an idea of the basic attitudes necessary for the functioning of the model. We will come back to this point in § 3.2). Let us simply mention here that these attitudes cannot be used operationally if they are not both:

- "rich" enough to serve as a basis for comparing complex actions, and
- stable and well-defined enough to be detectable in operational fashion.

As in the case of the qualitative scale described above, the richness of these attitudes is generally assumed to be such
3.2. The techniques for assessing utility functions

The aim of the classical techniques described in 2) is to ascertain the decision-maker's basic attitudes, by comparing simple actions, evaluated on the smallest possible number of attributes and involving easily comprehensible probabilities (1/4, 1/2, 3/4). Two sorts of attitudes are normally distinguished, and they are captured in two distinct ways:

- the one concerning the decision-maker's attitude towards risk serves to define the shape of the partial utility functions;
- those concerning trade-offs between attributes allow the coefficients $k_i$ to be defined.

It is clear that the quality of the assessment process depends crucially on the skill and experience of the analyst. Keeney (1977) even qualifies this part of the analyst's work as an "art". The classical method consists of taking each of the questions asked and delimiting as precisely as possibly that parameter (the consequence or the probability, depending on the technique chosen) which represents indifference between the two lotteries proposed. Many authors have insisted, however, on the fact that a systematic sensitivity analysis must be carried out on the parameters estimated in this way, in order to ascertain the robustness of the recommendations. It should be noted that these classical assessment techniques are not the only one that can be imagined. Some analysts have emphasized that the systematic search for basic attitudes of "infinite" richness (that is, allowing all the parameters of a utility function to be specified without ambiguity) could lead to certain biases. In some cases, a lack of knowledge about the problem by the decision-maker, a shortage of time for analysis, or the information available being information of lesser richness being deliberately sought.

One might note that, between probabilities $p_i$, but merely $n - 2$ bounds $p_i^*$ and $p_i^*$, such that:

$$
e_i \xrightarrow{p} p_i \xrightarrow{1-p_i} e_1 \quad \text{and} \quad e_i \xrightarrow{p} p_i^* \xrightarrow{1-p_i^*} e_1$$
These limits will only be coherent with the ordinal preferences expressed if they are such that \( p_{i*} > p_{j*}, \forall i > j \).

Obviously, in this last case, the recommendation will not be as rich as in a complete preorder, since each level can only be allocated an "interval" of utility.

No theoretical consideration can lead to one or other method being preferred, since everything depends on the decision-maker's basic attitudes to the problem under consideration. Such attitudes (1977) shows clearly that in certain cases, for a given level \( e_i \), there exists an interval \([p_{i*}, p_{i*}^2]\) within which all actions of the type \( p \in [p_{i*}, p_{i*}^2] \) are perceived as equivalent. The reader is referred to Vedder (1973) for an attempt to formalise these types of situations.

Keeney and Raiffa (1976) point out that the assessment techniques make the decision-maker reply to questions that are sometimes extremely delicate, and they thus consider it as a useful means to "require the decision-maker to reflect on his preferences and to hopefully straighten them out in his own mind" (p. 190), because he has to examine his feelings concerning the consequences of the actions. Unquestionably, to be able to express these fundamental attitudes requires a major reflection by the decision-maker, and this certainly contributes to enriching his perception of the problem. This "maieutics" (cf. the explicit reference to Socrates in Keeney and Raiffa (1976), p. 9) which the analyst submits the decision-maker to - and which reminds one of psychotherapy (cf. Fishhoff (1980)) - must be analysed as a training process (cf. Howard (1980)). It can consequently be seen as a deformation - which has non-negligible consequences on the way of regarding the ascertainment of basic attitudes in a decision-aid study. It can effectively be argued - by what is mainly an act of faith - that the difficult reflection imposed on the decision-maker in the assessment process will not produce simply an explicitation of latent attitudes, but rather an enriching of them. It often results in a rethinking of preconceived ideas. The very sort of questions that the deci-
sion-maker is asked has, in my view, a similar effect. Having recourse to lotteries, which use "simple ideal actions" to explore the decision problem, certainly upsets considerably the decision-makers' habits who seem "naturally" to prefer to argue in terms of real actions (cf. most descriptive studies of decision processes, especially Hirsch et al (1978) and Jacquet-Lagrèze et al (1978)). Furthermore, it must be recognized that the analyst who uses a decision-aid model necessarily affects the decision process well before he makes his formal recommendation. As Roy (1983) remarks (chap. 2), to speak of the neutrality or objectivity of the analyst raises many questions. This is all the more true in our particular case because the axioms of the formal model act as a "consistency guideline" of preferences during the information-gathering process (as we will see later). If ever some of the basic attitudes of the decision-maker seem incompatible with the axioms, the analyst is duty bound to try to rectify the situation (by repeating questions, for example, or encouraging the decision-maker, or seeking redundant information). A large number of practitioners have emphasized the analyst's educative role and the consequent dangers of manipulation (cf. Howard (1980) and Keeney and Raiffa (1976), pp. 189-191). It would clearly be pointless to underestimate the intellectual probity of the analysts, but it is nevertheless essential to question the point - and the practicality - of looking for

stable and structured basic attitudes, their extrapolation to real actions is no longer a simple logical deduction from the axioms of the formal model.

3.3. The extrapolation of the basic attitudes

Once it has been established that:

- the basic attitudes conform to the axioms of the formal model,
- they are rich enough to define the utility function without ambiguity, and can be ascertained operationally,
- the decision-maker is prepared to conform to axioms A1, A2, and A3,

then the validity of extrapolating the attitudes observed to the set of potential actions in the decision-aid study is guaranteed by the representation theorems mentioned in 1).
DECISION-AID AND EXPECTED UTILITY THEORY

For Friedman and Savage (1948), the only "positive" test of the theory is to compare a decision-maker's real choices with the choices deduced from the extrapolation of previously-observed attitudes. This whole question is therefore principally an empirical one, an idea I will come back to in 4. I would like simply to point out here that the validity of this extrapolation depends crucially on the nature of the attitu-

des (that is, as games) implies the danger of introducing major biases into the replies obtained, given the importance of social and even quasi-religious conceptions on ideas on games. This danger will be increased if - rightly or wrongly - the decision-maker sees the actions in terms of imprecision, error or indetermination rather than uncertainty or risk. The very fact of showing the decision-maker the advantages of representing his preferences concerning complex consequences (for example, the cost of a nuclear power-station to be built in ten years' time) in terms of simple money-based games is indicative of the analyst's role of education and/or training.

In summary, these techniques allow data gathered to be extrapolated only if the real actions are perceived in the same way as the ideal ones (as probability distributions) at this stage of the decision-process and if the decision maker wishes his preference structure to conform to the axioms of the formal model. The first of these two conditions is equivalent

badly be largely determined by the social and cultural context (rather than the personal one - cf. Phillips and Wright (1977) and Wright and Phillips (1983)). The second one implies that the precise status of the axioms has to be discussed more thoroughly.

3.4. The role of the axioms

The axioms of the formal model play a double role in the decision-aid model. First of all, as Fishburn (1967) points out, they form a "consistency guideline" during the gathering of the basic attitudes. As we have already mentioned, if the decision-maker produces opinions during the assessment process that are inconsistent with the axioms, he will be asked to bring him to reflect on the questions asked and thus arrive, hopefully, at a set of attitudes that are compatible with A1,
A2 and A3. These axioms thus specify the sort of information that will work for the model.

The axioms of the formal model also play another role which, although less often mentioned, is perhaps just as important. As we saw in 3.3, the validity of the extrapolation depends on whether the decision-maker is prepared to discuss the problem within a preference structure governed by the axioms. The restriction that this provides is not a trivial factor for it is clearly impossible to ask the decision-maker if he agrees with all the consequences of applying the axioms of the formal model to the basic attitudes he has expressed.

In other words, this point may cause difficulties unless the decision-maker accepts (implicitly or explicitly) the normative nature of the axioms. Between the ascertainment of the basic attitudes and the acceptance of the recommendations by the decision-maker, these is an "act of faith" in the axioms. To my knowledge, only Howard (for example (1980)) recognises explicitly the need for normative axioms in the decision-aid model (and not merely in the formal model). Keeney (see for example (1980), p. 387) argues explicitly against such a need.

This problem is not specific to the model we are considering. In all decision-aid models, the decision-maker can only expect to receive non-trivial recommendations if he accepts that he or the analyst must structure his preferences with respect to certain conventions or norms.

It is difficult to judge the normative attractiveness of the axioms of the formal model - especially since this problem lies in with the more general one of rationality, which has already caused a number of disagreements (cf. the discussions between Allais and the "American School", as reported in Allais and Hagen (1979)). But in this connection, the formal model does present a non-negligible advantage: that of making the nature of the conventions proposed absolutely clear. Axioms A1 and A2 also incontestably present a certain normative attractiveness at a first stage of the analysis. A1 means that the preference structure has to be both complete and transitive, while A2 would seem to constitute a particularly simple rule for comparing two complex actions. As for A3, it imposes structural limits which do not seem too restrictive. Even if many authors (cf. for example Samuelson (1952)) have recognized the a priori normative acceptability of A2, others, for instance Allais (1952) and Meld (1952),
have criticised it by means of particularly simple examples (the controversy on this question is further pursued in Econometrica, Oct. 1952, pp. 61 and ss.). More recently, and in a slightly different context, the normative implications of Savage's independence axiom have been severely criticised by Toulet (1982) (see also the survey of Hagen (1983)).

MacRae (1986) has demonstrated that most individuals tend to "correct" whichever of their choices violates the axioms of the formal model. This experiment has been criticised and replicated by Slovic and Tversky (1974) who show that only 6% of individuals (experiments 1, problem 1) agree to change choices that violate the independence axiom when clear arguments in its favour are put to them (versus 25% of people who violate this principle when it is criticised in their presence). MacCrimmon and Larsson's recent study (1979) provide similar evidence in that it shows the limited influence of the independence principle on many individuals (cf. figs. 4 and 6, pp. 358 and 360).

More generally, one can say that the smaller the "distance" between the "cognitive style" of the decision-maker and the principles of the axioms, the more natural normativeness will seem to him. It is clear, however, that this merely amounts to saying that decision-aid causes least problems when it is least useful. The reason is that when the decision-maker's cognitive repetition style is close to that of the axioms, applying them to the problem can only slightly improve his perception of the problem. The dilemma would only be solved if the transitivity, completeness and independence principles could be considered as being contained "in embryo" even in preference structures stemming from undeveloped cognitive and/or perceptive abilities. As for the transitivity - the least controversial - a priori - of the three principles the "money pump" argument pinpoints the irrational nature of non-transitive structures, which in general expose the decision-maker to heavy monetary losses without compensating advantages. Burros's detailed axiomatic analysis (1974) does show however that a decision-maker may knowingly wish to employ a non-transitive structure without suffering from any significant losses of money, under very general conditions. Burros argues that the only disadvantage of keeping intransitivity is that one cannot use methods drawn from utility theory to solve decision-problems. Setting the "money-pump" argument aside, one has to ask whether the transitivity
of the preference relation naturally appears, whatever the decision-maker's level of perception. An intuitive response might be to reply yes. In fact, the first order stochastic dominance relation, which corresponds to a very low level of perception (cf. Fishburn (1978), in this case it is possible to show that every bounded, continuous and monotone utility function is compatible with the ranking) is transitive, which is also the case for the highly structured preferences (the set of utility functions compatible with the preference relation is reduced to one up to a positive linear transformation) used in utility theory. Bisdorff (1981) studies this problem by using a typology of the different levels of perception possible, with each level corresponding to perception operations (concatenation, additions, partitions, ...) of increasing complexity, thus allowing one to compare actions. By deliberately adopting a method involving "teaching" the decision-maker perception operations so that he can carry out progressively richer comparisons of the different actions, Bisdorff shows clearly that if transitivity is present at the two extreme levels of the cognitive capacity of a decision-maker, intermediate situations can produce examples of "semi-cyclical" preferences: that is, ones where non-comparability upsets the transitivity of the strict preference relation (cf. Bisdorff (1981), pp. 166 ss.). This phenomenon can be interpreted intuitively as corresponding to the fact that two actions $a$ and $c$ can only be compared if they have enough "points in common", and that this is not necessarily the case, even when $a \succ b$ and $b \succ c$. It is natural in such situations to conclude from $a \succ b$ and $b \succ c$ either that $a \succ c$ or that $a$ and $c$ cannot be compared.

These considerations help to place the attraction of the normativeness qualities of the axioms of the formal model in perspective, by showing that much of their interest comes from their "divinatory power". Indeed, Bisdorff (1981) shows that, under very general hypotheses, the preference structures implied by a high level of perception, include the structures produced by more basic perceptions. Using the decision-aid model then leads to "rediscover" "intuitive" preference situations, which is very satisfying for a decision-maker. Nevertheless, the preference structure brought to light by the decision-aid model may involve a non-negligible part of arbitrariness if the decision-maker's knowledge of the problem is weak. If the "divinatory" effect of these methods certainly contributes to their being accepted in practice,
recognizing this can only persuade the analyst of the need to consider the normative character of the axioms of the formal model as merely relative.

This rapid analysis clearly does not enable one to come to a definitive assessment of the normative character of the axioms making up the formal model. Each decision-maker must ultimately decide for himself whether or not his preference structure should conform to this convention: its principal merit, in my view, would therefore seem to lie in its explicitness.

4. A FEW EMPIRICAL STUDIES THAT TEST THE VALIDITY OF THE DECISION-AID MODEL

As we have just seen, the value of the recommendations of "decision analysis" depends crucially on the decision-maker's having a set of consistent basic attitudes towards his problem and on being able to ascertain them operationally. In this section, I will review a number of empirical studies dealing with this problem.

Shortly after the publication of von Neumann and Morgenstern's book (1947), economists (Mosteller and Nogee (1951)) and psychologists (Davidson, Suppes and Siegel (1957)) addressed the question of whether it is possible to establish utility functions empirically on the basis of choices made by subjects in laboratory conditions. The economists' attention was however concentrated during this period on the normative advantages and the realism of the independence axiom of the formal model, constituting the principal difference between the rationality assumed by von Neumann and Morgenstern and the classical ordinalist conception of rationality. With the appearance of decision analysis in the late 1950's, research was concentrated almost exclusively on the problem of encoding subjective probabilities, as the first empirical studies seemed to prove that it was possible to construct utility functions operationally (even if this entails the use of stochastic preferences (cf. Mosteller and Nogee (1951))). Despite the numerous applications and successes of decision-analysis, the problems of knowing whether basic attitudes actually exist and of ascertaining them, have again been much discussed over the last decade. Recent studies tend to show that assessing a utility function is a more delicate task than the early studies had implied. Schematically, they can be grouped around
two poles. The first one concerns decision-makers' perceptions and manipulations of the probabilities, and the second one, their basic attitudes towards simple choices involving risks.

4.1. Perceptions of probabilities

A very large number of studies have been devoted to a critique of existing techniques for encoding probabilities (see Speltzer and Von Holstein (1975)) and to the problem of the perception of environments involving risks. These questions are important for two reasons. First, the formal theory assumes that the set of actions possesses properties of a mixture set (cf. Hershstein and Milnor (1953)). In other words, for this theory to be operative, perceptions of actions must conform to the principles used in calculating the probabilities. Secondly, it is a necessary condition that one should be able to allocate a precise probability to each of the various consequences of an action before one can apply the principles of expected utility.

A large number of studies, done mostly by experimental psychologists, have shown convincingly that most individuals produce evaluations of situations that are far removed from the basic principles of probability theory. Most frequently, such evaluations are based on simplistic heuristic methods, and contain major internal contradictions (cf. Tversky and Kahneman (1973, 1980), Kahneman and Tversky (1982), Bar-Hillel (1973, 1980) and a work which reprints the essential of most of the recent studies: Kahneman, Slovic and Tversky (1981)). Given the wide diffusion of these studies, I will not comment on them here.

4.2. The perception of simple actions and of the basic attitudes

The aim of the present section is to review a certain number of recent empirical studies, whose results can be interpreted in terms of existence and stability of basic attitudes.

I will not go into the problem of the independence axiom here, given that many studies have shown that it is violated even when simple actions are being compared. More fundamentally, I will show that the very existence of stable basic attitudes causes difficulty in certain situations.
Within a decision-aid objective, the most disturbing experimental results are the great instability of preference judgments expressed as ratios that are consistently reversed in sign. An example of this instability is the "preference-reversal" phenomenon demonstrated by Lindman (1971) and Lichtenstein and Slovic (1971, 1973) and confirmed by Grether and Plott (1979) (further tests of this phenomenon may be found in Pommerehne et al (1982) and Reilly (1982), Lichtenstein and Slovic (1983) provide a clear review of this topic, but see also Von Winterfeld (1980) who seems not to have encountered this phenomenon). This preference reversal stems from the fact that an individual's preference for one lottery or another is crucially affected by the way he is asked to express it (as a preference or as the minimum price he would be prepared to sell the lottery for). As an example, Lichtenstein and Slovic (1973) asked players in a Las Vegas casino to consider the following two lotteries:

\[
\text{Game A: } \frac{11}{12} \text{ plaques, } \frac{1}{12} \text{ plaques} \\
\text{Game B: } \frac{2}{12} \text{ plaques, } \frac{10}{12} \text{ plaques}
\]

where the value of the plaque is 25 \( \ell \). Game A offers a very high probability of winning, combined with highly improbable, but not impossible, important losses. Game B offers the possibility of winning large sums, both with much lower probability. A majority of players claimed to be indifferent between A and B, but 88% of players would sell B at a higher price than A (even more surprisingly, 87% of the players claiming to prefer A sold it for less than \( B \)). As Grether and Plott (1979) point out, this phenomenon not only calls into doubt whether stable basic attitudes exist at all, but more generally it poses a serious problem for all models using the idea of preference in a descriptive way. Their detailed study shows that the only plausible explanation for this phenomenon remains the very great sensitivity of the replies obtained to the assessment method used (and not that the subjects wrongly perceived the probabilities or were not motivated enough). Similar conclusions are reached by Schoemaker and Kurenther (1982) (table 6, n. 613), Hershey and Schoemaker (1980) and Hershey, Schoemaker and Kurenther (1982) (experiments 4 and 5), grouped together under "context effects". These studies point out that preferences between two lotteries can be greatly affected by whether they
are presented as games or as insurance problems. Hershey and Schoemaker (1980) (table 1, p. 121) observe that 80.5% of their subjects prefer L2 to L1, where

L1 is: "you run the risk of losing $1000, with a probability 1/100", and
L2 is: "you can insure yourself for $10 against this risk"

whereas only 56.1% of subjects preferred L3 to L1, where

L3 is: "you lose $10, with certainty".

It is as if referring to "insurance" induced a strong dislike for risk in certain subjects, by bringing into play certain "social norms" in favour of the idea of insurance within a very different framework. Conrath (1973) has also observed major "context effects" caused by the way to present lotteries (see also Slovic, Fischhoff and Lichtenstein (1982), Tversky and Kahneman (1981) and Zagorski (1981, 1975)).

In the same vein, Tversky (1975), Kahneman and Tversky (1979) and Hershey, Schoemaker and Kurenther (1982) show that a large number of factors affect the decision-makers' basic attitudes (they often decide under laboratory conditions; but it would seem probable that this does not affect the issue (cf. Schoemaker (1980), chap. 1, and (1982))). This is of course in contradiction with the axioms of the formal model, and prevents any effective ascertainment of the basic attitudes. Amongst these factors, and leaving aside the context effects, one can quote:

- the overevaluation of certain consequences (Kahneman and Tversky (1979), problems 1 to 8);
- the isolation of those parts which are common to different lotteries (Kahneman and Tversky (1979), problems 10 to 12).

For two reasons, however, the implications of these studies for the prescriptive models are not obvious (their implica-
tions for economic theory and risk-benefit analysis is clearly discussed in Arrow (1983)). First, defenders of decision analysis have always emphasized the enormous difficulties involved in assessing a utility function (cf. Keeney (1977)). The decision-maker must, in particular, recognize (with the help of some interaction with the analyst) the normative attraction of the underlying axioms, and a large number of redundant questions must normally be asked before a coherent set of basic attitudes can be reached. For obvious reasons, in the studies described above, the subjects fill in a questionnaire without interference from the analyst. Since there is no interaction at this stage, these studies use "intuitive" preferences that can be modified when the interaction does actually happen, emphasising the normative character of the axioms. This question is all the more debatable, given that Kahneman and Tversky (1979) (p. 277) state clearly that such modifications should happen very often. They argue their model exclusively from a decision-theoretic perspective. In part, this implies that the only reason the decision-makers deviate from the formal model of utility theory is that they do not have the opportunity to notice their "inconsistencies".

Secondly, the "context effects" seem natural (one's behaviour in a casino is not the same as in an insurance broker's office). Many authors (cf. Keeney and Raiffa (1976), pp. 50-55, and Keeney (1980)) underline the need to define the consequences of the different actions without ambiguity, and, the context of the decision can generally be taken for granted in a real study.

Nevertheless, these empirical observations cannot be ignored, and there are reasons to think that the effects they describe (in particular favourable situations) are not completely absent in real studies. First, their results tie in with psychologists' observations about human information-processing about the use of simplistic heuristics and the inability to process large numbers of elements of information at the same time (Miller (1965)). Thus the example of the preference-reversal phenomenon is probably due to the fact that when B is evaluated at its selling price, the imposing amount of the 79 plaques is merely the starting-point of the analysis, and is mentally reduced so as to take into account the possibility of losing 5 plaques. But in contrast, deciding a preference on the basis of a direct comparison between A and B uses different information-processing techniques
and will thus tend to favour the lottery where the probability of gain is largest (which is, in fact, close to 1). These anchoring phenomena are apparently visible in most human behaviour (cf. Kahneman and Tversky 1974 for another example). In the same way, the reference effect is directly comparable with psychologists' ideas on the workings of human sensorial devices, which tend to work on relative values rather than absolute ones. At the very least, then, coherent and stable basic attitudes can only be obtained by resisting the "natural" behavioural tendencies of almost every subject (see Raiffa 1968) on this subject: as far as the perception of probabilities is concerned, it would seem probable that even being aware of such tendencies rarely enables a coherent perception of a situation to be realized, at least not by the subjects themselves. However, these indications are not directly applicable to the theoretical questionability, these studies do show how very difficult most people find it to consider actions as lotteries independently from everything else (the context and reference effects). Even though the context is generally clear in decision-aid studies, it follows that one can never be sure that the basic attitudes observed are free from extraneous factors (nor, therefore, that the decision-maker will not reject in their totality recommendations that do not correspond to his fundamental cognitive styles).

If one examines the process of assessing a utility function in laboratory conditions, one soon notices again violations of the axioms of expected utility theory and inaccurate perceptions of actions considered as lotteries.

The first thing that can be noticed is that the range over which the utility function is assessed plays a very important role, directly connected with the reference effect. When studying utility functions assessed using lotteries involving losses (i.e. values below the status quo, reference point or objective) Tversky (1975), Hershey, Kurethier, Schoemaker (1982), ex. 3, Payne and al. (1980) and (1981) and Jaffray and Cohen (1982) observed that they tended to exhibit much more risk-proneness than those assessed with lottery involving only gains (whereas the assessment on absolute values rarely allows to explain choices made around the status quo). In decision-aid studies, the range of assessment of the utility function is generally determined by the range of the values taken by the consequences of the actions being studied; and this tends to reduce the problem. Nevertheless,
some of the choices made seem to involve consequences linked to the phenomenon in question. As an example, Keeney and Nair (1977) assess the utility function for the cost of different sites for nuclear power-stations by considering the differential cost of each site vis-à-vis the cheapest one. One can reasonably ask whether the function obtained would have been the same if the most expensive site had been the one used.

The first systematic study taking account the possibility of great instability in the basic attitudes during the assessing process seems to have been Allais's one. His results were drawn from an experiment carried out in 1952, but were only published (in part) in 1979 (Allais (1979), appendix C). Although interpreting these results in their present form is difficult, they do seem to imply that functions estimated by means of the variable consequence method (index B1/2) differ considerably from those estimated by means of the variable probability method (index B 200). Allais observes that the functions assessed by means of the variable consequence method have no inflection point in the region considered and are generally concave, whereas the variable probability method produces functions with a very definite S-shape (Allais (1979), charts 7 to 16, pp. 646 ss.). It is important to point out that these utility functions were assessed using sums of money and probabilities that had little intuitive appeal: from 1000 FF (1952) to 1,000,000,000 FF for index B 1/2, and from 0.25 to 0.999, with four different assessments being carried out at 0.9, 0.98, 0.99 and 0.999 for the index B 200. But only when all the results of this study have been published can we gain an overall idea of its impact and validity.

assessment method has subsequently been confirmed, and more convincingly demonstrated, by Hershey, Kurenther and Schoemaker (1982), MacCord and de Neufville (1983), MacCord (1983) and Kamkar (1979). Hershey, Kurenther and Schoemaker (1982) (exp. 1) show that the variable consequence method produced utility functions generally more risk-prone than the variable probability method does (the attribute studied being losses of money). MacCord and de Neufville (1983) confirm that the method exerts a great influence. Their study has two advantages compared with Hershey, Kurenther and Schoemaker's (1982). First, each of their 23 subjects was intervie-
ned separately, and for each one a utility function for financial gains was assessed (from $0 to $10,000) using both methods. Secondly, these individual interviews seem to have enabled the subjects to adapt most of their choices which were not compatible with the formal theory, and thus bring in the normativeness of the axioms. MacCord and de Neufville (1983) explicitly state this possibility, but do not say to what extent it influenced the results. Their findings can be summarised in the following way:

- The method used for assessing the utility function has a vital influence on its shape (on one point, the difference was as much as 100%).

- This influence can never be interpreted as random fluctuation in the replies obtained. This is because:

- The variable probability method gives a utility function that is, respectively, below/above the one that the variable consequence method gives, depending on whether the individual is risk averse/prone (cf. MacCord and de Neufville (1983) concerning the way in which this classification was carried out);

- The functions obtained by means of the variable consequence method are very sensitive to the probabilities chosen as reference ones. Whether or not the individuals are risk averse, the higher the probability of winning the maximum amount, the higher the values of the function for a given consequence. This last observation is perfectly compatible with the anchoring phenomena noticed in the preference-reversal problem.

Hershey, Kurenther and Schoemaker's results (1982) are difficult to compare with those of MacCord and de Neufville (1983), since they deal only with financial losses and financial gains respectively. The presence of the "reference effect" leads one to imagine that behaviour is probably very different in the two zones. However, posing the hypothesis of the "reflection effect" (cf. Kahneman and Tversky (1979), problems 3', 4', 7' and 8'; but see Hershey and Schoemaker (1980) for a critical analysis) changes matters: one can conclude from Hershey, Kurenther and Schoemaker's results (1982) that, as far as gains are concerned, the variable consequence method would have produced more risk averse utility
functions that those obtained by means of the variable probability method. This is perfectly compatible with the data obtained by MacCord and de Neufville; but it would be interesting to verify the results by means of further studies.

At first sight, these empirical observations are particularly disturbing (see also the comments in Arrow (1983)). It follows from the formal theory that all the assessment methods reviewed above are to be considered as equivalent and should lead to identical utility functions apart from random noise. It is then legitimate to wonder whether the utility functions assessed during decision-aid studies do not represent predominantly the influence of the assessment techniques used. Given these criticisms, it is not surprising that numerous descriptive extensions of utility theory have been proposed, in general trying to cope with a large number of empirical findings (cf. Kahneman and Tversky (1979), Kamar-kar (1979), Handa (1977), Chew (1983), or Machina (1983), or Bell (1982), plus the theories bringing in the notion of regret, e.g. Loomes and Sugden (1982) or (1983), or Bell (1982)). None of them, however, have offered convincing explanations of the very considerable instability of the basic attitudes revealed by the "context effect". In any case, the normative and prescriptive advantages of these "descriptive" extensions remain minimal. Fischhoff, Goiten and Shapira (1982) (p. 331) conclude from these studies that it is possible that we do not have structured preferences, and that at some fundamental level, our values may be incoherent. In this connection, MacCord and de Neufville (1983) note that the replies given to their questions were relatively stable over time... while continuing to violate utility theory. This fact that...
consciously letting his preferences violate them in any important way at a given time.

5. DISCUSSION

Facing these empirical studies, the status of the decision-aid model based on expected utility theory has to be reconsidered. In fact we saw that the validity of this decision-aid model crucially depends on the existence of a set of stable and consistent "basic attitudes", whereas section 4 leads us to think that these attitudes are rarely stable, and sometimes non-existent.

It should be pointed out straightaway that none of the empirical studies just mentioned involved the "art" necessary for assessing utility functions (see however MacCord and de Neufville's attempt). It is possible a priori to deny any value in these studies outside of their description of individual's intuitive behaviour in laboratory conditions. This criticism indicates the absolute need to establish experiments protocols that make great use of the normative attraction of the axioms through the discussion of contradictions, the use of redundant questions, etc. The results quoted above nevertheless seem so fundamental (and to a certain extent autonomous of the empirical context) that it is unlikely that they could ever be totally excluded from decision-aid studies. In another context, their importance have been stressed by Arrow (1983).

Therefore, to be able to interpret in this perspective how decision analysis works in practice, one is bound to admit that the assessment process in the strict sense (which tries to capture the basic attitudes of the decision-maker) is not the most important part of the decision-aid studies. It seems in fact that the most difficult and important task is the necessary first step of determining the general characteristics of the decision-maker's preferences concerning the attribute being considered, and in particular his attitude towards risk. We saw that many types of common attitudes reduce considerably the number of degrees of freedom for choosing the utility function by imposing drastic restrictions on its functional form (cf. Pratt (1969), for example $u(x) = a x + b$, $u(x) = a e^{cx} + b$, $u(x) = (x + b)^c$, etc.). Consequently, if one has managed to ascertain these general
characteristics, the estimation itself of the parameters of the function selected merely involves a very small number of questions about lotteries, and these will in principle be free from the complications described above (for example, in the case of an exponential type utility function, a single question allows to specify it completely. In the case of a unique assessment there is no risk of inconsistency). Consequently, the problem generally shifts from assessment in the strict sense to analysing the general attitudes towards risk of the decision-maker.

Completely a hypothesis like: "the decision-maker's risk aversion is constant over the attribute considered", as this would imply an infinite number of questions, even supposing that his preferences were structured enough to reply to them (an unreasonable supposition, in general). Thus, the analyst's problem is not that of describing basic attitudes as accurately and subtly as possible, but comes back to "negotiating" with the decision-maker how to structure his preferences on the basis of a mutually-agreed convention couched mainly in qualitative terms.

In order to be able to interpret decision analysis studies, one is bound to drastically reverse their whole perspective from a purely "descriptive attitude" to what Roy and Bouyssou (1983) called a "constructive attitude", consisting mainly of structuring a preference relation on the basis of certain conventions proposed to the decision-maker, taking into account what seems to be the most stable part of his fundamental attitudes and tastes. Roy and Bouyssou (1983) conclude from their recent study that the formal model can only be used to legitimize the decision taken, especially in the form of a "constructive" model which can only work in practice if it is embedded in a constructive approach - and this applies indeed to every decision-aid model.

What I am not arguing is that the decision-analysis model is not useful in a constructive framework. Indeed, the success of a large number of studies (see for instance the references in Keeney (1982) or Schoemaker (1982)) shows how untenable such a position would be. But if one believes that a constructive approach is inevitable, it is essential to admit that the above described method is not the only concei-
vable one to help a decision-maker structure "rationally" his preferences (for another type of convention, see for instance Hagen (1983), pp. 14-16).

Ultimately, the whole decision-analysis process depends to every great extent on the decision-maker's willingness to adopt such an attitude towards risk when formalising preference distributions.

It is then important to stress that the decision-analysis model does not draw its force from the existence of the formal model, which indeed, I would claim, confers no special legitimacy per sé, but from the type of convention it offers to the decision-maker for structuring his preferences.

ACKNOWLEDGEMENTS

We are grateful to Professors Bernard Roy and Fred Wenstrøp for their detailed review of earlier versions of this text.

LAMSADE, Université de Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, F-75775 PARIS CEDEX 16.

REFERENCES


Bisdorff, R. : 1981, Evaluations distributionnelles sur une échelle de préférence qualitative, Thése de 3e cycle, Université de Paris-Dauphine.


Keeney, R.L.: 1979, 'Evaluation of proposed sotrage sites'.


Luce, R.D.: 1956, 'Semi-orders and a theory of utility dis-


Miller, G.A. : 1956, 'The magical number seven, plus or minus two, some limits on our capacity for processing information', Psychological Review 63, no 2, 81-97.


Thrall, R.M. : 1954, 'Application of multidimensional utili-
Decision Processes, John Wiley and Sons.

Toulet, C. : 1982, 'Respect de l'axiome d'indépendance dans
l'incertain total : Pessimisme, optimisme et indécision'.

Tversky, A. : 1979, 'Intransitivity of preferences'. Psycholog-

Tversky, A. : 1977, 'On the elicitation of preferences : Des-
criptive and prescriptive considerations', in D.E. Bell,
R.L. Keeney and H. Raiffa (eds.), Conflicting objectives
in decision, John Wiley and Sons.

for judging frequency and probability', Cognitive Psycho-
logy 5, n° 2, 207-232.

Tversky, A., Kahneman, D. : 1974, 'Judgement under uncertain-
ity, Heuristics and Biases', Science 185, 1114-1131.

and the psychology of choice', Science 211, 542-573.

Vedder, J.N. : 1973, 'Multiattribute decision making under
uncertainty using bounded intervals', in J.L. Cochrane and
M. Zeleny (eds.), Multiple Criteria Decision Making, Uni-
versity of South-Carolina Press.

Wallenstein, T.S., Budescu, D.U. : 1983, 'Encoding subjec-
tive probabilities - A psychological and psychometric review',
Management Science 29, n° 2, 151-173.

von Winterfeld, D. : 1980, 'Additivity and expected utility
in risky multiattribute preferences', Journal of Mathema-
tical Psychology 21, 66-82.

Wold, H. : 1952, 'Ordinal preferences or cardinal utility ?', E-
conometrica October, 661-663.

Wright, G.N., Phillips, L.D. : 1983, 'Personality and probabi-
listic thinking', Submitted to British Journal of Psychol-
ogy.

Zagorski, M.A. : 1975, 'Risky decision : Attention effects
or masking effects ?', Acta Psychologica 39, 487-494.

NOTES


2. that is to say a representation such that:
   \[ a \rightarrow b \Leftrightarrow E(u, a) > E(u, b) \] and not only
   \[ a \rightarrow b \rightarrow E(u, a) > E(u, b). \]

3. These lotteries can be viewed as "ideal" actions as they are defined in totally unambiguous fashion by their probability distributions. This is not often true for a "real" action in a decision-aid study.

4. This ranking is said to be "latent" because the decision-maker is unable to express it directly. It is nevertheless contained in his basic attitudes in the presence of the axioms.

5. It is true that even when there is no much discrepancy between the cognitive abilities of the decision-maker and the axioms, decision analysis can still provide useful decision-aid through a clear structuring of the decision problem which, in particular, simplifies the calculation

6. This section will only deal with the problem of ascertaining the basic attitudes and not with the empirical studies analysing hypotheses generally accepted in utility theory concerning attitude towards risk (on this point see Schoemaker (1980) and his bibliography, and Krzysztofowicz (1983)).