DECISION-AID AND UTILITY THEORY:
A CRITICAL SURVEY (*)

CAHIER N° 52
novembre 1983

D. BOUYSSOU

# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESUME:</td>
<td>1</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>II</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1. THE FORMAL THEORY</td>
<td>1</td>
</tr>
<tr>
<td>2. FROM THE FORMAL MODEL TO THE DECISION-AID MODEL</td>
<td>3</td>
</tr>
<tr>
<td>3. THE HYPOTHESES AND THE FUNCTIONING OF THE DECISION-AID MODEL</td>
<td>7</td>
</tr>
<tr>
<td>4. A FEW EMPIRICAL STUDIES THAT TEST THE VALIDITY OF THE</td>
<td>19</td>
</tr>
<tr>
<td>DECISION-AID MODEL</td>
<td></td>
</tr>
<tr>
<td>5. DISCUSSION</td>
<td>28</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>31</td>
</tr>
</tbody>
</table>
AIDE A LA DECISION ET THEORIE DE L'UTILITE :
UNE SYNTHESE CRITIQUE

RESUME

La théorie de l'utilité peut être vue soit en tant que théorie formelle de représentation numérique de structures préférentielles, soit en tant qu'outil d'aide à la décision, au travers des techniques de la "Decision Analysis". L'argument essentiel en faveur de la "Decision Analysis" tient en l'existence du modèle formel. Ce cahier vise à clarifier les hypothèses qui sous-tendent l'utilisation des concepts de la théorie formelle dans une optique d'aide à la décision. De nombreuses études empiriques récentes permettant d'apprécier la validité de ces hypothèses sont également passées en revue.

Mots-clés : Théorie de l'utilité ; Decision Analysis ; Estimation des fonctions d'utilité ; Aide à la décision.
Utility theory can be viewed either as a formal theory providing representation theorems for preference structures or as a decision-aid model, usually referred to as "Decision Analysis". The main argument in favour of decision analysis is the existence of the formal theory. This paper intends to clarify the hypotheses underlying the use of the concepts of the formal theory within a decision-aid framework. Numerous recent empirical studies dealing with those hypotheses are also reviewed.

Key-words: Utility theory; Decision analysis; Utility function assessment; Decision-aid.
INTRODUCTION

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 statistical theory came to have a practical application.

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

\[ a \mathrel{P} b \iff E(u, a) > E(u, b) \quad \forall (a, b) \in \mathcal{X} \]

where

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

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

The theory specifies axioms which, if verified, produce the desired representation. To attain an expected utility representation, we need
- 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.
2. FROM THE FORMAL MODEL TO THE DECISION-AID MODEL

The primordial aim of the decision-aid model based on utility theory is to put into practice 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:

1) Understanding the specific nature of the decision problem, including generating the set of all potential actions and describing the set of all consequences. It should be remembered that when one wishes to bring different attributes into this set, they should ideally have the independence properties needed for the utility function to be decomposed (see below).

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 frequentist data, Savage's model (1951) provides the theoretical basis justifying recourse to personal (or subjective) probabilities. A very large number 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, as this makes the estimation techniques much easier to apply.

3) Determining the decision-maker's preference structure. This third stage is designed to quantify the decision-maker's attitude towards risks...
and the possible trade-offs between the different criteria, using a utility function. In the multi-attribute case, the global utility function is only in practice decomposed additively or multiplicatively.

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

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

either \[ u(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} k_i u_i(x_i) \tag{1} \]

or \[ 1 = k u(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} (1 + k_i u_i(x_i)) \tag{2} \]

with \[ u(\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_n) = 1, u_i(\bar{X}_i) = 1 \ \forall i \in \{1, \ldots, n\} \]
\[ u(x_1, x_2, \ldots, x_n) = 0, u_i(x_i) = 0 \ \forall i \in \{1, \ldots, n\} \]

\[ k_i \in ]0, 1[ \ \forall i \in \{1, \ldots, n\} \]

\[ k_i \in ]0, 1[ \ \forall i \in \{1, \ldots, n\} \]

with \( k \geq -1 \) and \( k = 0 \) solution of

\[ 1 + k = \prod_{i=1}^{n} (1 + k_i). \tag{3} \]

When the conditions of theorem 1 are satisfied, the utility function can consequently be determined completely by estimating \( n \) partial utility functions \( u_i(x_i) \) and \( n \) coefficients \( k_i \) ensuring that each function remains strictly within the interval \([0, 1]\).

The estimation of the coefficients \( k_i \) and the functions \( u_i(x_i) \) is effected by comparing various "ideal" actions (i.e. defined in totally unambiguous fashion by their probability distributions as opposed to "real" actions).

Let us consider two lotteries, \( L_1 \) and \( L_2 \):

\[ L_1 \longrightarrow \frac{1}{1} \left( x_1^O, x_2^O, \ldots, x_j^O, x_j^O, x_{j+1}^O, \ldots, x_n^O \right) \]
In accordance with the hypotheses of theorem 1, \( L_1 \) and \( L_2 \) can be compared in the same way, whatever the values of \( (x_1^0, x_2^0, \ldots, x_n^0) \) (which will therefore be omitted from now on). Two classical techniques are available for estimating the utility function: the "variable consequence" one and the "variable probability" one. Let us denote \( L_1(x_j) \) the lottery that produces consequence \( x_j \) with absolute certainty, and \( L_2(\tilde{x}_j, p, x_j) \) the lottery producing consequence \( \tilde{x}_j \) with probability \( p \) and consequence \( x_j \) with probability \( 1-p \). The variable consequence method consists simply of specifying a probability \( p_0 \) and two consequences \( \tilde{x}_j \) and \( x_j \), and asking the decision-maker for the consequence \( x_j^* \) representing indifference between \( L_1(x_j^*) \) and \( L_2(\tilde{x}_j, p_0, x_j) \). If he is able to specify this consequence \( x_j^* \), one can then use the principles of expected utility to set: \( u_j(x_j^*) = p_0 u_j(\tilde{x}_j) + (1-p_0) u_j(x_j) \). For reasons to do with perceptions of the probabilities, the value most often chosen for \( p_0 \) is 1/2. The procedure is initiated with \( \tilde{x}_j = \tilde{x}_j \) and \( x_j = x_j \). If \( L_1(x_j^*) \) is indifferent to \( L_2(\tilde{x}_j, 1/2, x_j) \), then \( u(x_j^*) = 1/2 \). The procedure is then iterated by considering the lotteries \( L_2(x_j^*, 1/2, x_j) \) and \( L_2(\tilde{x}_j, 1/2, x_j^*) \) and looking for the lotteries \( L_1(x_j^{**}) \) and \( L_1(x_j^{***}) \) which are respectively indifferent to them. We then have \( u_j(x_j^{**}) = 0.25 \) and \( u_j(x_j^{***}) = 0.75 \). In this way, reiteration may be used to obtain as many points as are desired for estimating the function \( u_j(x_j) \).

The variable probability method is theoretically equivalent to the above method. It consists of specifying a probability \( p^* \) and asking the decision-maker which consequence \( x_j^* \) is such that \( L_1(x_j^*) \) and \( L_2(\tilde{x}_j, p^*, x_j) \) are indifferent. Once this consequence \( x_j^* \) has been obtained, one can set \( u_j(x_j^*) = p^* \). The procedure is then iterated by considering other probabilities \( p^* \). For the two methods considered, the reply to each of the questions asked involves in practice progressively reducing the size of an interval on the values of \( x_j \) that contains the value representing indifference between \( L_1 \) and \( L_2 \) being sought. This technique has the advantage of escaping largely from the bias brought in
when one seeks the indifference value immediately. In addition, the analyst's task may be considerably simplified - cf. Pratt (1964) - since the functional form of the utility function may be linked with simple qualitative considerations concerning the decision-maker's attitudes towards risks.

The estimation of the coefficients $k_i$ is based on simple concepts, even if putting the methods that I am about to survey into practice raises important problems (cf. Roy and Bouyssou (1983)). Under the hypotheses of theorem 1, let us consider the two lotteries

$$L_i \circ \frac{1}{l} (x_1, x_2, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots, x_n)$$

$$\mathcal{L} \circ \begin{cases} 
1 \rightarrow (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n) \\ 
1-p \rightarrow (x_1, x_2, \ldots, x_n) 
\end{cases}$$

The expected utility of $\mathcal{L}$ is by definition equal to $p$, whereas that of $L_i$ is $k_i$ in both the additive and the multiplicative forms. This implies that if the decision-maker is able to define the probability $p$ ensuring indifference between $L_i$ and $\mathcal{L}$ then one can set $k_i = p$. The procedure can then be iterated by considering all the lotteries $L_i$, $i \in \{1, \ldots, n\}$. If $\sum_{i=1}^{n} k_i \neq 1$, the complete estimation of the utility func-
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 AND THE FUNCTIONING OF THE DECISION-AID MODEL

Much of the attractiveness of the "decision-analysis" techniques is due to the presence of the formal model, which implies a "logical" basis for the
then it is very difficult to imagine what aid he could require. Rappoport concludes therefore that von Neumann and Morgenstern's formal theory for representing preference structures cannot serve as a foundation of decision-aid, as axiom $A_1$ 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 $A_1$ 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 $A_1$ 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 $A_2$ and $A_3$, 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, let us imagine that one is attempting to help the decision-maker to rank a set of actions evaluated 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:

```
  e_1 \quad p_i \quad e_n
  \quad 1-p_i \quad 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 $u(e_i) = p_i$, $i = 2, 3, ..., n - 1$ and $u(e_1) = 0$, $u(e_n) = 1$, which defines the utility function perfectly and allows one to resolve the original problems completely.
such that, if two actions may 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 $D_0$ must therefore be seen as an "extension" of the basic attitudes referred to (implicitly) in $D_4$; 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.
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 the recommendations.

The techniques described in 2) give an idea of the basic attitudes necessary for the functioning of the model. We will come back to this point in 3h.)
As in the case of the qualitative scale described above, the richness of these attitudes is generally assumed to be such as to allow any action to be ordered in an unambiguous way. In other words, the decision-maker must be able to order all the actions in at least "latent" fashion. Indeed, this ordering follows on directly from axioms $A_2$ and $A_3$. The total preorder mentioned in $A_1$ can therefore be considered as a "latent total preorder", to be made explicit by the analyst. Roy and Bouyssou (1983) describe such a model as "descriptive", in the sense that it seeks to apply maximum precision to the task of analysing an existing preference relation. In my view, this term "descriptive" applies even if one only assumes the existence of a set of basic attitudes, since the axioms (already verified, often implicitly, at the beginning - cf. 3d)) mean that this hypothesis is equivalent to that of the existence of a complete preference relation.

As far as the existence and the nature of these basic attitudes are concerned, many of the questions raised up till now will only be answered by means of empirical studies. Before pursuing this point further, it is important to further analyse the mechanisms underlying this method.

b) 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 criteria allow the coefficients $k_1$ to be defined.

It is clear that the quality of this assessment process depends crucially on the skill and experience of the analyst. Keeney (1977) even qualifies
ceived 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. Uncontestably, 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. Fischhoff (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, one that can, for example, result in a rethinking of preconceived ideas. The very sort of questions that the decision-maker is asked has, in my view, a similar affect. 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 this general sort of 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 emphasised 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 under-estimate the intellectual probity of the analysts, but it is nevertheless essential to question the point – and the practicality – of looking for stable basic attitudes at all costs in such an educative process. If one refuses to assume the existence of a set of stable and structured basic attitudes, their extrapolation to real actions is no longer a simple logical deduction from the axioms of the formal model.

c) 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). 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 attitudes gathered on the basis of ideal actions. As Hull, Moore and Thomas (1973) remark, representing ideal actions as lotteries (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 in 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 to requiring a mainly empirical analysis, one which will probably 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.

d) 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, the analyst must try to bring him to reflect on the questions asked and thus arrive, hopefully, at a set of attitudes that are compatible with $A_1$, $A_2$ and $A_3$. 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 c), 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, there 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 strength of the axioms of the formal model - especially since this problem ties 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 $A_1$ and $A_2$ also incontestably present a certain normative attractiveness at a first stage of the analysis. $A_1$ means that the preference structure has to be both complete and transitive, while $A_2$ would seem to constitute a particularly simple rule for comparing two complex actions. As for $A_3$, it imposes structural limits which do not seem too restrictive. If many authors (cf. for example Samuelson (1952)) have recognized the a priori normative acceptability of $A_2$, others, for instance Allais (1952) and Wold (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).

MacCrimmon (1968) carried out an experiment which implies 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 (experiment 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 Laarson's recent study (1979) provide similar evidence in that it shows the limited influence of the in-
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 \( a \text{ P } b \) and \( b \text{ P } c \). It is natural in such situations to conclude from \( a \text{ P } b \) and \( b \text{ P } c \) either that \( a \text{ P } c \) 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, one can show 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 "re-discover" "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.

a) **Perceptions of probabilities**

A very large number of studies have been devoted to 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. Hertstein 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 (1972), 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.

b) **The perception of simple actions and the ascertainment 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.

(*) 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).
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 prescriptive objective, the most disturbing experimental results are the great instability of preference judgements expressed on actions that are very close to those used for assessing utility functions. The most striking 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) (but see also Von Winterfeld (1980) who seem 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:

- **A**: 11/12 plaques, 1/12 - 24 plaques
- **B**: 2/12 plaques, 10/12 - 5 plaques

where the value of the plaque is £25. 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, but 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, p. 613), Hersley 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 $L_2$ to $L_1$, where

$L_1$ is: "you run the risk of losing $1000, with a probability of 1/100", and

$L_2$ is: "you can insure yourself for $10 against losing $1000";

whereas only 56.1% of subjects preferred $L_3$ to $L_1$, where

$L_3$ 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 evaluation of the actions with respect to a reference point which can be manipulated (Tversky (1975), figs. 1, 2 and 3, and Kahneman
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 one to resist them). Secondly, even if the realism of certain situations the subjects are put in sometimes seems questionable, 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 the presence again of these effects violating the axioms of the theory and inaccurate perceptions of actions considered as lotteries.
This sensitivity of the shape of the utility function to the assessment method has subsequently been confirmed, and more convincingly demonstrated, by Hershey, Kurentner and Schoemaker (1982), MacCord and De Neufville (1982) and Kamarkar (1979). Hershey, Kurentner 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 (1982) confirm that the method exerts a great influence. Their study has two advantages compared with Hershey, Kurentner and Schoemaker's (1982). First, each of their 23 subjects was interviewed separately, and for each one a utility function for financial gains was assessed (from $0 to $100) 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 normativity of the axioms. MacCord and De Neufville (1982) (pp. 13, 17) 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 (1982) 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 (cf. MacCord and De Neufville, figs. 2 and 3). This last observation is perfectly compatible with the anchoring phenomena noticed in the preference-reversal problem.

Hershey, Kurentner and Schoemaker's results (1982) are difficult to compare with those of MacCord and De Neufville (1982), since they deal
respectively only with financial losses and financial gains. 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. The formal theory, in effect, leaves the nature of the consequences on which the preferences are expressed completely uninterpreted. It follows 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), Kamarkar (1979), Handa (1977), plus the theories bringing in the notion of regret, e.g. Loomes and Sugden (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 note (1982) (p. 17) that the replies given to their questions were relatively stable over time... while continuing to violate utility theory. This fact that replies to questions in terms of "certainty equivalents" were stable implies in my opinion that a specific context (the question) "triggers" off stable heuristics for processing
information, and not that stable attitudes or preferences necessarily exist in the decision-maker's mind. The distinction may seem so subtle as to be non-existent. But it should be emphasised that the preferences expressed are probably stable only because they result from the repeated application of a single mechanism. I believe that this is the only way of explaining why systematic violations of principles like transitivity, with a great deal of normative attraction, should be predictable and reproducible (cf. Tversky (1969)). The attractiveness and elegance of the axioms of the theory make one believe that no decision-maker would ever consider consciously letting his preferences violate them in any important way at a given time.

5. DISCUSSION

Where does this leave the practice of decision-aid studies? It should be pointed out straightaway that none of the empirical studies just mentioned involves using the "art" necessary for assessing utility functions (see however MacCord and De Neufville's attempt (1982)). It is possible a priori to deny any value in these studies outside of their descriptions of individuals' intuitive behaviour in laboratory conditions. This criticism would seem partly justified and it would be a very good idea to establish written 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 would nevertheless seem so fundamental (and to a certain extent so intuitive, as in the context effect), that it would seem very unlikely that they could ever be totally excluded from decision-aid studies.

To be able to interpret the practice of real study, taking these facts into account, one is bound to admit that the assessment process in the strict sense would very often seem to mark the point where most of the work has already been done. In general, the most difficult 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 risks. In fact, 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 \( u(x) = ax + b \), \( u(x) = a e^{cx} + b \), \( u(x) = (x + b)^c \), etc.).

Consequently, if one has managed to ascertain these general characteristics, the estimation itself of the coefficients 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. It is obviously unthinkable to test 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 in mainly 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 legitimate the decision-analysis model if it is purely descriptive. But the crux of the whole problem is that this model 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 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 conceivable one to help a decision-maker structure "rationally" his preferences.

Ultimately, the whole decision-analysis process depends to every great extent on the decision-maker's willingness to adopt such and such an attitude towards risk when comparing actions, described in terms of probability 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 se, but from the type of convention it offers to the decision-maker for structuring his preferences.
REFERENCES


Hirsch G., Jacquet-Lagrèze E. and Marchet J.C. (1978): Description d'un processus de décision - II: Illustration : Le cas de la raffinerie de Brest, Cahier du LAMSADE n° 17, Université de Paris-Dauphine.


Miller G.A. (1956) : The magical number seven, plus or minus two, some limits on our capacity for processing information, Psychological Review, Vol. 63, n° 2, pp. 81-97.


Raiffa H. (1968) : Decision analysis, Addison Wesley.


