

CAHIER DU LAMSADE

Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision
(Université de Paris-Dauphine)
Unité de Recherche Associée au CNRS n° 825

ARE JUDGMENTS ABOUT RELATIVE IMPORTANCE OF CRITERIA DEPENDENT OR INDEPENDENT OF THE SET OF ALTERNATIVES ? AN EXPERIMENTAL APPROACH

CAHIER N° 111

V. MOUSSEAU

mai 1992

Reçu : avril 1992.

CONTENTS

Résumé	i
Abstract	i
Acknowledgements	ii
Introduction	1
1. Previous related research.	2
2. Description of the experiment scheme.	3
2.1. Experiment design.	3
2.2. The questionning procedure.	4
2.3. Nature of the collected information.	4
3. Results.	5
3.1. Relation between intuitive weights and rankings.	5
3.2. Results concerning personal preferences about prizes.	6
3.3. Results concerning the student's ranking inferred from his weights	8
3.4. Results concerning the woman's weights infered form her ranking.	10
3.5. Effect of question's order on subject's answers.	11
4. Discussion.	
4.1. Descriptivist vs constructivist approach to multiple criteria decision aid.	13
4.2. Mismatch between theoretical and DM's interpretation of weights.	15
4.3. Implications for importance parameters' elicitation methods.	16
Conclusion.	17
Appendix	
Statistical tests	18
References	20
Questionnaire	22

**LES POINTS-DE VUES D'UN DECIDEUR SUR L'IMPORTANCE RELATIVE DES
CRITERES SONT-ILS INDEPENDANT OU CONTINGENT
A L'ENSEMBLE DES ACTIONS POTENTIELLES ?
UNE APPROCHE EXPERIMENTALE.**

Résumé

Cet article décrit un travail expérimental concernant la notion d'importance relative des critères en aide multicritère à la décision. Il est montré le poids intuitif et la décisivité d'un critère sont affectés par une modification de l'ensemble des actions potentielles. Les impacts de ce résultat expérimental sur la modélisation des préférences et les méthodes d'évaluation des paramètres d'importance sont discutés.

Mots-clés : Aide multicritère à la décision, Importance relative des critères, Expérience.

**ARE JUDGMENTS ABOUT RELATIVE IMPORTANCE OF CRITERIA
DEPENDENT OR INDEPENDENT OF THE SET OF ALTERNATIVES ?
AN EXPERIMENTAL APPROACH.**

Abstract

This paper describes an experimental work about relative importance of criteria in Multiple Criteria Decision Aid. It is shown that intuitive weights and decisivity of criteria change with a modification of the set of potential alternatives. Impacts on preference modelling and importance parameters elicitation methods are discussed.

Keywords: Multiple Criteria Decision Aid, Relative Importance of Criteria, Experiment.

Acknowledgements

I would like to thank Dominique Champ-Brunet, Dominique François and Margareth Jarvis for having corrected and improved the english version of the text. Patrice Perny should also be thanked for his fruitfull comments on the draft of this paper.

Lastly I am gratefull to all subjects for their kind participation to the questionnaire.

Introduction

Multiple criteria preference modeling requires that one obtains from the decision-maker(s) (DM) inter-criteria preference information to discriminate between pareto-optimal alternatives. This information is modeled through importance parameters that vary across models : hierarchy of criteria, importance coefficients or weights, veto thresholds, aspiration levels, etc.

From a theoretical point of view, the values assigned to such parameters are meaningless as long as the aggregation rule in which they are used is not specified¹ : The use of different aggregation procedures on the same data set and with the same values for the importance parameters may induce preference reversals. Moreover, in some aggregation procedures (such as the weighted sum, partial utility function aggregation) the importance parameters are linked with the numerical representation of each criterion, while in others (ELECTRE, PROMETHEE) these variables are independent of the scale on criteria.

From a practical point of view, many papers have discussed the problem of the evaluation of such information; some of them have proposed importance parameters' evaluation methods. Several methods use a set of alternatives to put questions to the DM. These alternatives may be real (i.e. corresponding to a concrete decision) or fictitious (i.e. without any link to a concrete decision). Fictive alternatives are chosen to put questions that provide a particular information; they may be either realistic fictive alternatives (that might correspond to a concrete decision) or unrealistic fictive alternatives (alternatives that could not exist in the real world situation, whose performances do not respect the existing statistical links).

In this paper, we will present an experiment that aims at enhancing the following problem : Does a modification of the set of alternatives influence the DM's judgments about relative importance of criteria? This issue is crucial for the above-mentioned

¹ Consequently, every method aiming at collecting preferential information about relative importance of criteria should provide a good match between the collected information and its use in the aggregation of preferences. Unfortunately, numerous methods proposed in the literature ignore the above theoretical consideration.

methods. The answer to this question directly influences the conception of such methods. A positive answer would require that these methods provide information that is consistent with the set of real alternatives. In particular, any method that put questions using unrealistic fictive alternatives (i.e. alternatives whose evaluations are extreme) would become unacceptable.

1. PREVIOUS RELATED RESEARCH.

Numerous related works have been carried out within the framework of multiple attribute utility theory (MAUT). These studies are based on the theoretical finding that MAUT requires a specific relationship between the weight parameters and the range of attribute evaluations. In other words, weights of MAUT are scaling constants and are to be modified when a change of the encoding of criteria occurs (see [Hobbs 80]).

These papers empirically investigate the modifications in the elicited weights due to a change in attribute range. Most authors reported that the elicited weights changed in the correct direction but found the amount of weight modification to be insufficient (c.f. [Stewart & Ely 84], [Beattie & Baron 91], [Von Nitzsch & Weber 90]). [Von Nitsch & Weber 90] studied this phenomenon using different weight elicitation methods. Note that all the above-mentioned studies referred to MAUT and usually upheld the concept of true weight (cf. § 3.1.).

[Goldstein 90] carried out experimental work without reference to any aggregation procedure (only single peakedness of preference orders was hypothesized) and showed that intuitive elicited weights vary with a change in the set of potential alternatives and should be interpreted in relation to this particular set. Our work is similar to this approach and does not hypothesize any aggregation procedure underlying DM's preferences.

2. DESCRIPTION OF THE EXPERIMENT SCHEME.

2.1. Experiment design.²

The concrete framework of the experiment is the following : Subjects are told that the organizer of a tombola wants to investigate gamblers' preferences about the prizes to win. He hesitates between two types of prizes :

- credits in a bookstore.
- free cinema access cards.

In order that the subjects feel really involved in the decision, they are told at the beginning of the questionnaire that one of them will be selected and will win the prize that he has chosen.

Two series of composite prizes are successively proposed : the prizes are made up of two components, a credit in a bookstore and a free cinema access card. They differ according to the amount of the credit and the duration of the card. The two series of prizes are described hereafter :

		Duration of the free cinema access card	Amount of the credit in the bookstore
1 st Series	A	18 months	1400 french francs
	B	15 months	1800 french francs
	C	12 months	2200 french francs
	D	9 months	2600 french francs
	E	6 months	3000 french francs
	F	3 months	3400 french francs
2 nd Series	U	18 months	0 french francs
	V	15 months	1000 french francs
	W	12 months	2000 french francs
	X	9 months	3000 french francs
	Y	6 months	4000 french francs
	Z	3 months	5000 french francs

² this experiment was inspired by [Goldstein 90]

Table 1 : Sets of proposed alternatives.

2.2. The questioning procedure.

Each subject individually filled in the questionnaire (see appendix). He was asked to fulfill the three following tasks for the two series of prizes :

- A* Rank the 6 prizes and assign intuitive numerical values to weights (values must add up to ten). Weights are defined without any relation with an aggregation procedure.
- B* Infer the ranking of a 24 year old student from his given weights. The weights of the student are the same in the two series of prizes. These weights are chosen among four different weighting systems (8-2, 6-4, 4-6, 2-8) that appear with an equal frequency in the questionnaires.
- C* Infer the weights of an active 30 year old woman from the ranking of prizes that she gave. For each subject, the ranking of each series places the prizes in the same order (for example : BCADEF and VWUXYZ). Therefore criteria have the same decisivity (cf. 2.3.) in the two proposed rankings.

2.3. Nature of the collected information.

They are two different types of collected information :

- Weighting systems corresponding to a ranking of prizes (tasks *A* and *C*). Let us denote w_b^1 and w_b^2 as the weight assigned to criterion "book" in the first and second series of prizes, respectively. Let w_c^1 and w_c^2 be the weight assigned to criterion "cinema" in the first and second series of prizes, respectively.³
- Rankings of prizes corresponding to a weighting system (tasks *A* and *B*). It is

³ In this experiment, the values assigned to the intuitive weights do not refer to any aggregation procedure. Moreover, it is not possible to determine to which aggregation rule the subjects referred when assigning values to such intuitive weights. Therefore, no underlying aggregation model is hypothesized and our aim is not to analyse the collected information within a particular framework.

necessary to define a way of using this information in numerical terms. Let us denote n_b^1 and n_b^2 as the number of elementary permutations (permutation of two consecutive elements) that transform O (corresponding to the first and the second series of prizes, respectively into a lexicographic order in favor of criterion "book". As the two series contain six prizes, $n_b^i \in [0,15]$. Let us define $d_b^1(O)$ and $d_b^2(O)$ as the decisivity index of criterion "book" for the ranking O and for the 1st and second series of prizes respectively as follows :

$$d_b^i(O) = \frac{15 - n_b^i(O)}{15} \in [0,1] \quad (i \in \{1,2\})$$

Examples :

$$d_b^1(A \succ B \succ C \succ D \succ E \succ F) = \frac{15 - 15}{15} = 0$$

$$d_b^1(F \succ E \succ D \succ C \succ B \succ A) = \frac{15 - 0}{15} = 1$$

If subjects answer the questionnaire so that $w_b^1 \neq w_b^2$ and $d_b^1 \neq d_b^2$, then we will conclude that judgments about relative importance of criteria are dependent on the particular set of potential alternatives considered.

3. RESULTS.

The 124 subjects that filled in the questionnaire were students and teachers. The average age was 24 and they took an average time of 14'30" to fill in the questionnaire. Subjects profiles were relatively similar; however, the aim of the experiment is to carry out an intra-subject analysis. No inter-subject interpretations will be made : our purpose is not to compare answers across subjects but to test if judgments of each subject about relative importance of criteria are affected by a modification of the set of proposed alternatives.

3.1. Relation between intuitive weights and rankings.

In such an experiment, one should check for an internal validity in the subject's answers. In particular, a necessary condition for the data not to be suspicious is to find a relatively high correlation between weights and decisivities given by the same subject.

The correlations coefficients between d_b^i and w_b^i ($i=1,2$) are reported in table 2.

	Series of prizes A-F	Series of prizes U-Z
Personnal preferences	0.84	0.81
inferred student's preferences	0.91	0.89
inferred woman's preferences	0.75	0.70

Table 2 : Correlation coefficients between d_b^i and w_b^i ($i=1,2$).

The values of these coefficients are high except for those concerning the inferred woman's preferences. These high values seem intuitive and show a good internal consistency in the answers. It allows us to state that the obtained data are exploitable.⁴ However, it would be meaningless to base any interpretation on these high values. Indeed, basing any interpretation on such correlations hypothesizes that all subjects are using the same mental procedure to link a ranking to weights. Practical experience seems to contradict this hypothesis. In other words, the nature of the information contained in the values assigned to intuitive weights seems to vary across subjects.

3.2. Results concerning personal preferences about prizes.

Table 3 presents the whole subjects' responses concerning their personal preferences about prizes proposed. A majority of subjects answered such that the weights and the ranking's decisivity of criterion "book" is greater in the second series of prize than in the first one.

	$d_b^1 < d_b^2$	$d_b^1 = d_b^2$	$d_b^1 > d_b^2$	
$w_b^1 < w_b^2$	45	4	3	52
$w_b^1 = w_b^2$	29	34	3	66
$w_b^1 > w_b^2$	1	2	3	6
	75	40	9	124

Table 3 : Answers concerning personal preferences.

⁴ However, data concerning the inferred woman's preferences seems slightly suspicious.

The two following figures illustrate graphically the values assigned by subjects to w_b^1 and w_b^2 (figure 1) and values of d_b^1 and d_b^2 computed from the rankings (figure 2). In these two figures, the size of points is proportional to the number of subjects that gave this particular answer.

In both figures, the scatterplots are significantly located over the bisectrix. This means that subjects gave a greater weight (and a greater decisivity in their rankings) for criterion "book" in the first than in the second series of prizes.

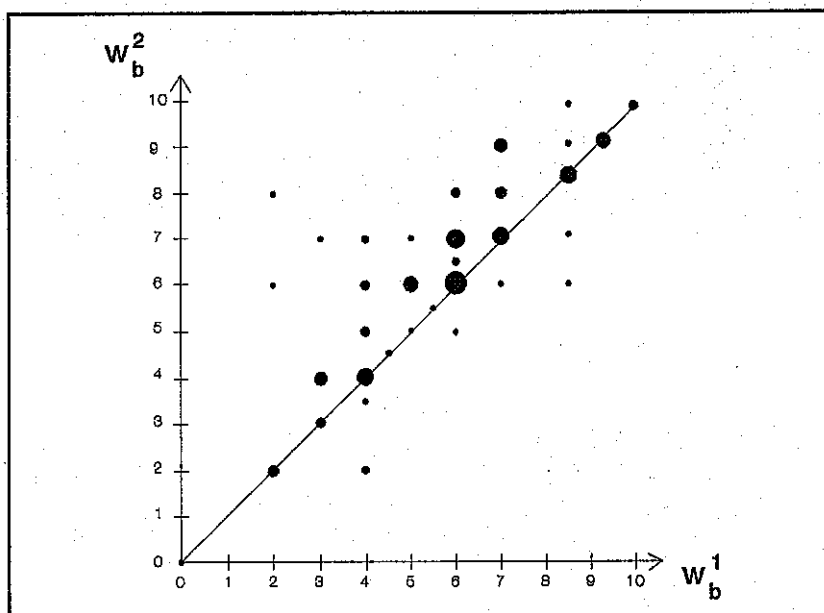


Figure 1 : Values assigned by subjects to w_b^1 and w_b^2

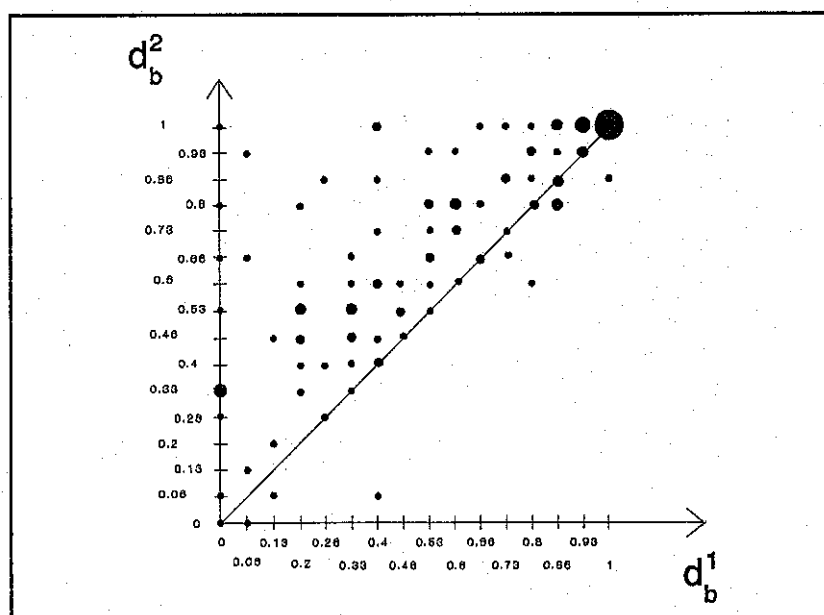


Figure 2 : Values d_b^1 and d_b^2 computed from subjects' rankings.

Table 4 presents the mean values and the standard deviations of subjects' intuitive weights and ranking's decisivity for criterion "book". Though these measures aggregate judgments among subjects, it appears that the mean values are such that $\bar{w}_b^1 < \bar{w}_b^2$ and $\bar{d}_b^1 < \bar{d}_b^2$. A T student test lead to reject $\bar{w}_b^1 = \bar{w}_b^2$ and $\bar{d}_b^1 = \bar{d}_b^2$ with $\alpha=0.05$ (see appendix).

	w_b^1	d_b^1	w_b^2	d_b^2
mean values	5.75	0.58	6.32	0.72
standard deviations	2.00	0.33	2.01	0.27

Table 4 : Mean values of personal preferences.

3.3. Results concerning the student's ranking inferred from his weights.

Subjects inferred, for each series of prizes, the ranking of a 24 year old student according to the weights that he gave (task *B*). We computed from these rankings the two decisivity indeces d_b^1 and d_b^2 . The following table gives the number of subjects that inferred rankings such that $d_b^1 > d_b^2$ (or $d_b^1 = d_b^2$ or $d_b^1 < d_b^2$) according to the student's weights :

	$w_b^1 = w_b^2 = 0.2$	$w_b^1 = w_b^2 = 0.4$	$w_b^1 = w_b^2 = 0.6$	$w_b^1 = w_b^2 = 0.8$	total
$d_b^1 < d_b^2$	15	24	15	4	58
$d_b^1 = d_b^2$	17	4	6	22	49
$d_b^1 > d_b^2$	2	3	7	5	17
total	34	31	28	31	124

Table 5 : Inferred student's preferences.

Figure 3 illustrates graphically the values of d_b^1 and d_b^2 computed from the rankings inferred from the student's weights. The size of points is proportional to the number of subjects that gave this particular answer.

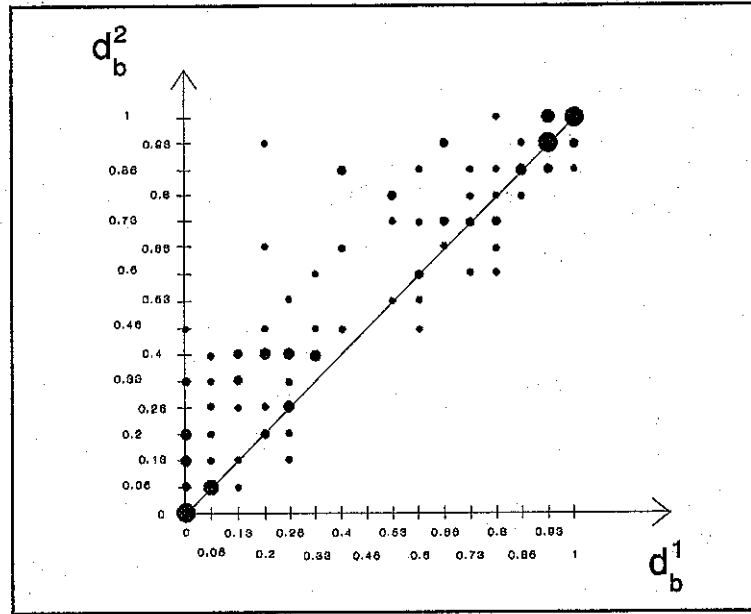


Figure 3 : Values of d_b^1 and d_b^2 computed from the ranking that subjects inferred from the student's weights.

The scatterplot is significantly located above the bisectrix. The mean values $\bar{d}_b^1$ and $\bar{d}_b^2$ were proved to be such that $\bar{d}_b^1 < \bar{d}_b^2$. A t student test lead to reject $\bar{d}_b^1 = \bar{d}_b^2$ with $\alpha=0.05$ (see appendix). Moreover, the scatterplot may be divided into two groups of points which correspond to two lexicographic preferential profiles. A large number of subjects seem to have interpreted the given weights in terms of a lexicographic aggregation.

	d_b^1	d_b^2
Mean values	0.45	0.54
Standard deviations	0.37	0.34

Table 6 : Mean values of inferred student's preferences

3.4. Results concerning the woman's weights inferred from her ranking.

Subjects inferred the weights of a young woman according to the rankings of prizes that she gave (task C). In each questionnaire, the two preference orders reflect the same preferential profile and have thus the same decisivity. The 5 different rankings that were uniformly distributed in the questionnaires are given in table 7 :

	Series A-F	Series U-Z	decisivity
1	B>C>A>D>E>F	V>W>U>X>Y>Z	0.1333
2	C>D>B>A>E>F	W>X>V>U>Y>Z	0.3333
3	C>D>E>B>F>A	W>X>Y>V>Z>U	0.5333
4	D>C>E>F>B>A	X>W>Y>Z>V>U	0.6667
5	E>D>F>C>B>A	Y>X>Z>W>V>U	0.8667

Table 7 : Proposed woman's rankings.

Table 8 gives the number of subjects that inferred weights such that $w_b^1 > w_b^2$ (or $w_b^1 = w_b^2$ or $w_b^1 < w_b^2$) according to the woman's rankings :

	1	2	3	4	5	total
$w_b^1 < w_b^2$	5	9	8	9	7	38
$w_b^1 = w_b^2$	15	8	11	10	11	55
$w_b^1 > w_b^2$	5	8	8	6	4	31
total	25	25	27	25	22	124

Table 8 : Inferred woman's preferences.

Figure 4 illustrates graphically the values of w_b^1 and w_b^2 that subjects inferred from the woman's ranking. The size of points is proportional to the number of subjects that gave this particular answer.

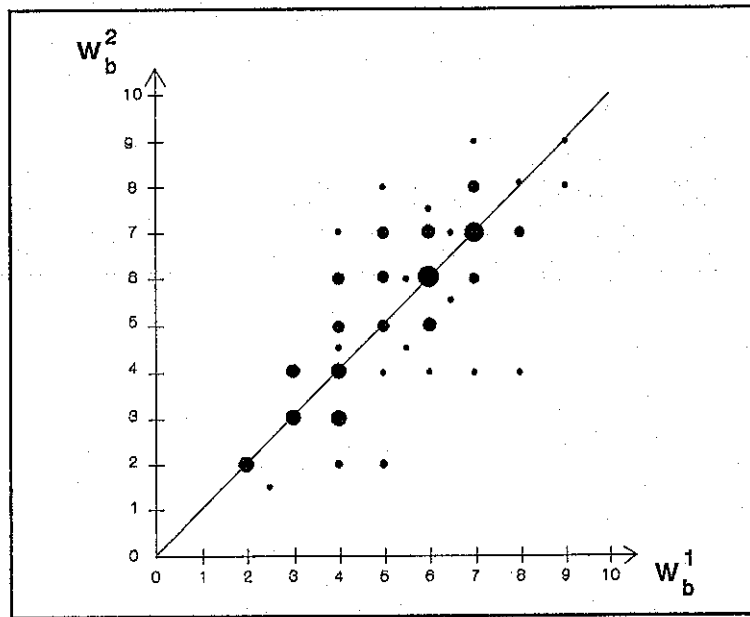


Figure 4 : Values w_b^1 and w_b^2 inferred by subjects from the woman's ranking.

It is difficult to draw clear conclusions from these results. It seems that subjects have difficulties in analysing the importance of criteria that underly a preference ranking. This interpretation is strenghtened by the relatively low correlation between d_b^i and w_b^i ($i=1,2$) that express a weak internal consistency in the subjects' answers (cf. § 3.1.).

3.5. Effects of question's order on subject's answers.

The obtained data enable us to state that subjects' answers were significantly influenced by the order in which questions were posed in the questionnaires. In half of the questionnaires (61/124), the questions concerning the prizes A→F were presented first while in the other half (63/124), these questions were presented after those concerning prizes U→Z. Table 9 presents the mean values and standard deviations of the answers when questions about prizes A→F were presented first and when they were presented at the end.

		Questions about prizes A→F presented first	Questions about prizes U→Z presented first
personal preferences	$\bar{w}_b^1, (s_{w_b^1})$	5.43 , (1.90)	6.05 , (2.10)
	$\bar{w}_b^2, (s_{w_b^2})$	6.01 , (2.03)	6.60 , (1.96)
	$\bar{d}_b^1, (s_{d_b^1})$	0.52 , (0.35)	0.66 , (0.31)
	$\bar{d}_b^2, (s_{d_b^2})$	0.68 , (0.28)	0.77 , (0.24)
inferred student's preferences	$\bar{d}_b^1, (s_{d_b^1})$	0.51 , (0.37)	0.41 , (0.35)
	$\bar{d}_b^2, (s_{d_b^2})$	0.59 , (0.31)	0.48 , (0.36)
inferred woman's preferences	$\bar{w}_b^1, (s_{w_b^1})$	5.16 , (1.66)	5.11 , (1.66)
	$\bar{w}_b^2, (s_{w_b^2})$	5.14 , (1.81)	5.23 , (1.86)

Table 9 : Mean values of the answers according to the order of the questions.

The data presented in table 9 show that there exists a difference between the mean values of the elicited parameters⁵ when questions about prizes A→F were presented first or after those concerning prizes U→Z. Parameters concerning personal preferences were higher when questions about prizes U→Z were presented first while parameters concerning inferred student's preferences were higher in the opposite case: there is no systematic direction of change of the parameter's values. No direct interpretation of this direction may be given. However, the results exhibit *question order effect* which has already been reported in the literature (cf. [Fishhoff et al. 89]).

⁵ The inferred woman's weights seem to be invariant through the two samples. However, as mentioned in § 3.1., these data were slightly suspicious.

4. DISCUSSION.

4.1. Descriptivist vs Constructivist approach to multiple criteria decision aid (MCDA).

The main result of this experimental work is that judgments of relative importance of criteria are linked with the evaluations of the potential alternatives in consideration. However, the consequences of such a result on multiple criteria decision aid may differ according to the epistemologic standpoint we consider. Roughly, two main epistemologic approaches are opposed on underlying assumptions about the evaluation of importance of criteria. These two streams disagree about the existence of true weights : the central issue of this debate relies on the existence of an objective reality⁶.

The first stream, *the descriptive approach*, refers to a stable information about relative importance of criteria that is pre-existent to the modelling process. It aims at capturing the existing values of the true weights. In this framework, the importance parameter evaluation methods are designed to elicit values that estimate these true weights. However, the obtained values are approximations of the reality that is to be described. If the elicitation method avoid biases, the estimated weights should be a correct approximation of the true weights. [Beattie and Baron 91] argue that there is "a distinction between true and estimated weights and that it is possible that the subjects' true weights remain constant at all times, but become distorted in the elicitation process". Within this approach, the role of the elicited weights is to fit to the pre-existing DM's preferences and accurately to describe the decision situation. The model using these weights is used to predict the DM's choices.

The second stream, *the constructivist approach*, rejects the existence of true weights representing an objective reality. This approach bases this rejection on several arguments. Firstly, the values assigned to importance parameters are meaningless as long as the aggregation procedure in which they are used is not specified : the same values for weights induce different preference structures according to the aggregation rule that is used. Therefore the concept of true weights becomes meaningless. Secondly, in certain aggregation procedures (e.g. MAUT), a modification of the encoding of the scale of criteria implies a modification of the weights in order to keep the preference structure

⁶ in a platonician sense.

unchanged. Finally, numerous experimental studies [Fichhoff et al 89], [Tversky et al 88], Von Nitzsch & Weber 90] have showed the lability of weights : framing effects, violation of procedure invariance, range effects have been reported in the literature as inducing a lability in the elicited weights. In this framework, weights are considered as appropriate values that must be interpreted through the aggregation procedure used. As the DM's preferences are not supposed to be entirely pre-defined, the elicited weights do not refer to true values. The role of the importance parameter elicitation methods is to construct weights that are acceptable for the DM and consistent with the aggregation procedure that is to be used.

What is the role of experimental work and behavioral science studies in these two approaches ?

In the descriptive approach, experiments are viewed as a way to stress the mismatch observed phenomena and model predictions (for example : preference reversals, biases in the estimation of the pre-existing true weights,...) . The results of these experiments are used to build new models that take into account the effects that have been pointed out in the experiments and to explain the observed preferences. For instance, preference reversals that have been studied in numerous works cannot be reported in multiple criteria utility theory. Several choice theories have been proposed to avoid the description failure of preference reversal (cf. [Tversky et al. 88], [Goldstein & Einhorn 87], [Loomes & Sugden 83], [Holt 86]).

The constructivist approach uses these experimental studies in a slightly different way. Within this framework, decision aid is to be prescriptive : the model does not attempt to be an accurate description of the real world but aims at providing to the DM recommendations that are consistent with the input parameter and the aggregation procedure (no predictive ability is required from the model). Nevertheless behavioral science studies may be useful to prescriptive decision aid. Rejecting the concept of true weights does not imply that experimental observations are suspicious or misleading. For the models to be useful, it should not be disconnected from the DM's way of thinking. For example, importance parameter elicitation methods should take into account the way DMs analyse the notion of importance of criteria (cf. [Belton & Gear 1984]).

4.2. Mismatch between theoretical and DM's interpretation of weights.

The present work has proved that intuitive values assigned by DMs to weights are varying with the set of potential alternatives. Similar modifications appear on the decisivity inferred from intuitive rankings of the alternatives. An analysis of the results shows that the intuitive weight of a criterion increases with the extent of the evaluations of the alternatives. This empirical assertion states that the notion of relative importance of criteria is understood by DMs as contingent to the set of alternatives. In other words, DM's understanding of this notion simultaneously integrates the discriminating power of the criterion with respect to the set of alternatives and the role of this particular criterion in the construction of overall preferences. This interpretation is consistent with the experimental observations reported in [Mousseau 91].

On the other hand, MCDA methods require preference information to infer overall preferences from the evaluations of alternatives on the criteria. From a theoretical point of view, preference information may be divided into two categories. *Intra-criterion preference information* is needed to build partial preferences on a single criterion from the evaluations of alternatives on this criterion. This information refers to the discriminating power on the criterion. It is frequently modeled through indifference and preference thresholds. *Inter-criteria preference information* is used to transform the partial preferences into overall preferences. It refers to the relative importance of criteria and defines the role played by each criterion in the aggregation of preferences. This information is often taken into account through weights or importance parameters.

This theoretical distinction between intra-criterion and inter-criteria preference information seems to mismatch with the DM's understanding of weights. In the aggregation procedure weights refer to inter-criteria preference information while DM's intuitive understanding of these weights includes inter- and intra-criteria preference information. This implies that the part of the information expressed in the intuitive weights concerning intra-criterion preferences is redundant and will be taken into account twice in the decision modelling process : once when modelling criteria (eventually through thresholds) and once when evaluating weights. This assertion implies that one should avoid direct rating methods to evaluate weight parameters; the information elicited in such methods does not fit to the concept of inter-criteria

preference information.⁷

This conclusion is an argument for the methods that elicit simultaneously inter-criteria and intra-criterion parameters (cf. [Martel & Nadeau 88], [Kiss et al 91]). In such cases, the preference parameter elicitation procedure matches quite well to the DM's reasoning.

4.3. Implications for importance parameter elicitation methods.

Numerous authors have proposed methods aiming at eliciting importance parameters from information obtained from the DM. These methods may be divided into two main groups : *direct methods* directly elicit importance parameters and do not refer to any aggregation procedure. *Indirect methods* are linked to a particular aggregation procedure and induce importance parameters' values from indirect information (frequently pairwise comparison of alternatives) using the aggregation rule. Most indirect methods obtain information through question about alternatives (pairwise comparisons, rankings, evaluations,...).

The present study has direct implications for indirect methods : as the values assigned to weights by DMs vary with the set of proposed alternatives, the DM's answers to questions about alternatives will be influenced by the particular alternatives to be compared (ranked or evaluated). The choice of the alternatives used in the questions will have an impact on the elicited values.

In order to fit to the DM's reasoning, these alternatives should correspond to an acceptable solution. Such alternatives should either be real (i.e. existing in the real world situation) or fictive but realistic (i.e. with medium evaluations on each criterion, these evaluation should conform to the existing statistical links between criteria).

The technique proposed by [Keeney & Raiffa 76], for example, is based on pairwise comparison of alternatives whose evaluations are extreme. This method is to be prohibited ; the DM will not be able to perceive the proposed alternatives and will have difficulty in comparing them ; though his answers will be suspicious or erroneous. Similarly, techniques that require pairwise comparisons of lotteries should be avoided. We propose to use real or realistic alternatives in the questions to be posed to the DM

⁷ However, the absence of link between weights elicited by such methods and the aggregation procedure used in the model is a sufficient argument to avoid these direct weighting methods.

(cf. for example [Roy et al 86], [Jacquet-Lagrèze & Siskos 82] and [Martel & Nadeau 88]).

However, the realistic aspect of the fictive alternatives to be presented to the DM may differ according to the problem statement. Choosing the best alternative, sorting alternatives into predefined categories, ranking alternatives (c.f. [Roy 85]) are different problem statements in which the DM have a different understanding of an acceptable solution. In a choice problem, the presented alternatives should correspond to "good" alternatives (i.e. having "desirable" evaluations on criteria) that could possibly be chosen. Otherwise, the DM would focus his attention on the marginal aspect of the decision process and it could lead the decision process in an erroneous direction. On the contrary, in a ranking problem, the evaluation of the presented alternatives should be distributed along the criteria scales.

Conclusion.

The present experimental work focussed on the empirical link between the notion of relative importance of criteria and the set of potential alternatives. It showed that a change of the set of alternatives produces a significant modification in the elicited values for importance parameters. The interpretation of such experimental work was shown to be different in a descriptive or constructivist approach to MCDA. The empirical observation enhanced the mismatch between the formal distinction inter/intra-criteria preference information and the DM's reasoning. We discussed the implications of this study for importance parameters' elicitation methods and showed that indirect methods posing questions dealing with alternatives should be very carefully designed. The alternatives intervening in the questions must be chosen within the DM's perceptive field and must be consistent with the statement of the decision problem.

Further research should be undertaken in order to check if lability in the weights still occur when other changes appear in the set of alternatives (addition or deletion of one or several alternatives).

APENDIX

STATISTICAL TESTS

equality of $\bar{w}_b^1$ and $\bar{w}_b^2$ for personal preferences.

as a preliminary, lets check for the equality of the variances $\sigma_{w_b^1}^2$ and $\sigma_{w_b^2}^2$.

$$\text{test} \begin{cases} H_0 : \sigma_{w_b^1}^2 = \sigma_{w_b^2}^2 \\ H_1 : \sigma_{w_b^1}^2 \neq \sigma_{w_b^2}^2 \end{cases}$$

we study the parameter $F = \frac{S_{w_b^2}^2}{S_{w_b^1}^2}$ a F-distribution with 123, 123 freedom degrees.

with $\alpha=0.05$, the critical value is 1.49. as $\frac{S_{w_b^2}^2}{S_{w_b^1}^2} = \frac{2.01^2}{2.00^2} = 1.01 < 1.49$ we accept H_0 . Let $s_{w_b} = 2$

Let's carry out the student's test to study : $\begin{cases} H_0 : \bar{w}_b^1 = \bar{w}_b^2 \\ H_1 : \bar{w}_b^1 \neq \bar{w}_b^2 \end{cases}$

$t = \frac{\bar{w}_b^2 - \bar{w}_b^1}{s_{w_b} \sqrt{\frac{1}{124} + \frac{1}{124}}}$ follows a student distribution with 124+124-2 degrees of freedom.

With $\alpha = 0.05$ the critical value is $t_\alpha = 1.96$. In our sample, $t = 2.216 > 1.96$ therefore we reject H_0 . $\bar{w}_b^2$ is significantly different from $\bar{w}_b^1$.

equality of $\bar{d}_b^1$ and $\bar{d}_b^2$ for personal preferences.

as a preliminary, lets check for the equality of the variances $\sigma_{d_b^1}^2$ and $\sigma_{d_b^2}^2$.

$$\text{test} \begin{cases} H_0 : \sigma_{d_b^1}^2 = \sigma_{d_b^2}^2 \\ H_1 : \sigma_{d_b^1}^2 \neq \sigma_{d_b^2}^2 \end{cases}$$

we study the parameter $F = \frac{S_{d_b^2}^2}{S_{d_b^1}^2}$ a F-distribution with 123, 123 freedom degrees.

with $\alpha=0.05$, the critical value is 1.49. as $\frac{S_{d_b^2}^2}{S_{d_b^1}^2} = \frac{0.33^2}{0.27^2} = 1.49$ it is difficult to conclude. we

may accept to carry out the mean comparison test ($n=124$). Let $s_{d_b} = 0.3$

Let's carry out the student's test to study : $\begin{cases} H_0 : \bar{d}_b^1 = \bar{d}_b^2 \\ H_1 : \bar{d}_b^1 \neq \bar{d}_b^2 \end{cases}$

$$t = \frac{\bar{d}_b^2 - \bar{d}_b^1}{s_{d_b} \cdot \sqrt{\frac{1}{124} + \frac{1}{124}}} \text{ follows a student distribution with } 124+124-2 \text{ degrees of freedom.}$$

With $\alpha = 0.05$ the critical value is $t_{\alpha} = 1.96$. In our sample, $t = 3.76 > 1.96$ therefore we reject H_0 . $\bar{d}_b^2$ is significantly different from $\bar{d}_b^1$.

equality of $\bar{d}_b^1$ and $\bar{d}_b^2$ for the inferred student's preferences.

as a preliminary, lets check for the equality of the variances $\sigma_{d_b^1}^2$ and $\sigma_{d_b^2}^2$.

$$\text{test} \begin{cases} H_0 : \sigma_{d_b^1}^2 = \sigma_{d_b^2}^2 \\ H_1 : \sigma_{d_b^1}^2 \neq \sigma_{d_b^2}^2 \end{cases}$$

we study the parameter $F = \frac{s_{d_b^2}^2}{s_{d_b^1}^2}$ a F-distribution with 123, 123 freedom degrees.

with $\alpha=0.05$, the critical value is 1.49. as $\frac{s_{d_b^2}^2}{s_{d_b^1}^2} = \frac{0.37^2}{0.34^2} = 1.18$ we accept H_0 . Let $s_{d_b} = 0.355$

Let's carry out the student's test to study : $\begin{cases} H_0 : \bar{d}_b^1 = \bar{d}_b^2 \\ H_1 : \bar{d}_b^1 \neq \bar{d}_b^2 \end{cases}$

$$t = \frac{\bar{d}_b^2 - \bar{d}_b^1}{s_{d_b} \cdot \sqrt{\frac{1}{124} + \frac{1}{124}}} \text{ follows a student distribution with } 124+124-2 \text{ degrees of freedom.}$$

With $\alpha = 0.05$ the critical value is $t_{\alpha} = 1.96$. In our sample, $t = 3.76 > 1.99$ therefore we reject H_0 . $\bar{d}_b^2$ is significantly different from $\bar{d}_b^1$.

REFERENCES

- J. Beattie & J. Baron** (1991) "*Investigating the effect of stimulus range on attribute weight*", Journal of Experimental Psychology : Human Perception and Performance, Vol. 17, n°2, pp. 571-585.
- V. Belton & T. Gear** (1984) "*A series of experiments into the use of pairwise comparison techniques to evaluate criteria weights*", in Y.Y. Haimes & V. Chankong eds., Proceedings, Cleveland, Ohio.
- B. Fishhoff, P. Slovic & S. Lichtenstein** (1989) "*Knowing what you want : Measuring labile values*" in Bell-Raiffa-Tversky (eds.), Decision making : Descriptive, Normative and Prescriptive Interactions, Cambridge University Press, pp. 398-421.
- W.M. Goldstein** (1990) "*Judgments of relative importance in decision making : Global vs. local interpretations of subjective weights*", Organizational Behavior and Human Decision Processes, Vol. 47, pp. 313-336.
- W.M. Goldstein & H.J.J. Einhorn** (1987) "*Expression theory and the reversal phenomena*", Psychological Review, Vol. 94, pp. 236-254.
- B.F. Hobbs** (1980) "*A comparison of weighting methods in power plant siting*", Decision Science, Vol. 11, n°4, pp. 725-737.
- C.A. Holt** (1986) "*Preference reversals and the independence axiom*", American Economic Review, Vol. 76, pp. 508-515.
- E. Jacquet-Lagrèze & J. Siskos** (1982) "*Assessing a set of additive utility functions for multicriteria decision making, The UTA method*", EJOR, vol. 10, n°2, pp 151-164.
- R.L. Keeney & H. Raiffa** (1976) "*Decision with multiple objectives : Preferences and value tradeoffs*", Wiley.
- L. Kiss, J-M. Martel & R. Nadeau** (1991) "*ELECCALC - Un DSS pour modéliser les préférences d'un décideur*", WP 91-38, FSA, Université Laval, Québec.
- G. Loomes & R. Sugden** (1983) "*A rationale for preference reversal*", American Economic Review, Vol. 73, pp. 215-233.
- J-M. Martel & R. Nadeau** (1988) "*Modélisation des préférences révélées avec ELECTRE II : ELECCALC, une approche interactive*", Document de Travail, GRADE, Université Laval, Québec.

- V. Mousseau** (1991) "*Empirical analysis of the notion of relative importance of criteria : some experiments*", Unpublished doctoral dissertation, Lamsade, Université Paris-Dauphine (in French).
- B. Roy** (1992) "*Decision science or decision aid science*", EJOR (to appear).
- B. Roy** (1985) "*Méthodologie Multicritère d'Aide à la Décision*", Economica.
- T.R. Stewart & D.W. Ely** (1984) "*Range sensitivity : a necessary condition and a test for the validity of weights*", Report n°270, Center for Research on Judgment and Policy, Institute of Cognitive Science, University of Colorado, Boulder.
- A. Tversky, S. Sattah & P. Slovic** (1988) "*Contingent weighting in judgment and choice*", Psychological Review, Vol. 95, pp. 371-384.
- R. Von Nitzsch & M. Weber** (1990) "*The effect of attribute range on weights in multiattribute utility theory*", WP, RWTH, Aachen.
- D. Von Winterfeld & Ward Edwards** (1986) "*Decision analysis and behavioral research*", Cambridge University Press.

QUESTIONNAIRE

Informations on the interviewe (confidential) :

Name :
Tel. Number:
Age :
Sex :
study level :

The organizer of a tombola wants to determine which type of prizes might interest the future participants. He hesitates between two types of prizes :

- Credits in a bokstore.
- Free cinema access card.

In order to fix the proportion of each of these type of prizes, the organizer wants to investigate the opinion of the public to which the tombola is intended.

In the following questionnaire, the different prizes differ according to the amount of the credit in the bookstore and the duration of the free cinema access card.

In the framework of this investigation, one of the interviewe will be randomly selected and will win a part of the prize for which has expressed preference. Good luck and do not forget to put down your name and telephone number.

Remarks :

This questionnaire does not aim at testing any consistency. There is no good or bad answer. The objective is not to judge your answers.

Any remark, commentary or suggestion concerning the questionnaire, difficulties in understanding it, problems in giving answers are welcome.

The prizes are the following :

- Prize A :** 1400 french Francs credit in a bookstore.
18 months of free access in all parisian cinema.
- Prize B :** 1800 french Francs credit in a bookstore.
15 months of free access in all parisian cinema.
- Prize C :** 2200 french Francs credit in a bookstore.
12 months of free access in all parisian cinema.
- Prize D :** 2600 french Francs credit in a bookstore.
9 months of free access in all parisian cinema.
- Prize E :** 3000 french Francs credit in a bookstore.
6 months of free access in all parisian cinema.
- Prize F :** 3400 french Francs credit in a bookstore.
3 months of free access in all parisian cinema.

Rank these six prizes according to your preferences.

For example : $C > D > B > A > E > F$

.....	>		>		>		>		>	
-------	---	-------	---	-------	---	-------	---	-------	---	-------

Give two weights or importance coefficients that add up to 10 in order to express the relative importance that you grant, in the above mentioned prize, to the credits in bookstores and free cinema access card :

Importance coefficients :	
Credits in bookstores	:
Free cinema access card	:

Now, you are to determine the preference ranking of a undergraduate 24 year old student according to the importance coefficients that he gave for each aspect of the prizes (credit in bookstores and free cinema access card) :

Here are the weights that he gave :

Credit in bookstores :

Free cinema access card :

According to this information that the student gave, rank the 6 prizes according to the student preferences :

Note for example : $C > D > B > A > E > F$

.....	>		>		>		>		>	
-------	---	-------	---	-------	---	-------	---	-------	---	-------

Reminder :

Prize A : 1400 french Francs credit in a bookstore.
18 months of free access in all parisian cinema.

Prize B : 1800 french Francs credit in a bookstore.
15 months of free access in all parisian cinema.

Prize C : 2200 french Francs credit in a bookstore.
12 months of free access in all parisian cinema.

Prize D : 2600 french Francs credit in a bookstore.
9 months of free access in all parisian cinema.

Prize E : 3000 french Francs credit in a bookstore.
6 months of free access in all parisian cinema.

Prize F : 3400 french Francs credit in a bookstore.
3 months of free access in all parisian cinema.

A 30 year old working woman gave us her preference ranking over the prizes. Try to determine (according to this only ranking), the importance coefficient (or weights) that this woman should fix for both aspects of the prizes (cinema and book) :

Woman's preference ranking : > > > > >

Importance coefficients :	
Credits in bookstores	:
Free cinema access card	:

Reminder :

- Prize A :** 1400 french Francs credit in a bookstore.
 18 months of free access in all parisian cinema.
- Prize B :** 1800 french Francs credit in a bookstore.
 15 months of free access in all parisian cinema.
- Prize C :** 2200 french Francs credit in a bookstore.
 12 months of free access in all parisian cinema.
- Prize D :** 2600 french Francs credit in a bookstore.
 9 months of free access in all parisian cinema.
- Prize E :** 3000 french Francs credit in a bookstore.
 6 months of free access in all parisian cinema.
- Prize F :** 3400 french Francs credit in a bookstore.
 3 months of free access in all parisian cinema.

We will now focuss on 6 other prizes (U, V, W, X, Y, Z) :

Prize U : 0 french Francs credit in a bookstore.
18 months of free access in all parisian cinema.

Prize V : 1000 french Francs credit in a bookstore.
15 months of free access in all parisian cinema.

Prize W : 2000 french Francs credit in a bookstore.
12 months of free access in all parisian cinema.

Prize X : 3000 french Francs credit in a bookstore.
9 months of free access in all parisian cinema.

Prize Y : 4000 french Francs credit in a bookstore.
6 months of free access in all parisian cinema.

Prize Z : 5000 french Francs credit in a bookstore.
3 months of free access in all parisian cinema.

Rank these six prizes according to your preferences.

For example : $W > X > V > U > Y > Z$

.....	>		>		>		>		>	
-------	---	-------	---	-------	---	-------	---	-------	---	-------

Give two weights or importance coefficients that add up to 10 in order to express the relative importance that you grant, in the above mentioned prize, to the credits in bookstores and free cinema access card :

Importance coefficients :	
Credits in bookstores	:
Free cinema access card	:

Now, you are to determine the preference ranking of a undergraduate 24 year old student according to the importance coefficients that he gave for each aspect of the prizes (credit in bookstores and free cinema access card) :

Here are the weights that he gave :

Credit in bookstores :

Free cinema access card :

According to this information that the student gave, rank the 6 prizes according to the student preferences :

Note for example : $W > X > V > U > Y > Z$

.....	>		>		>		>		>	
-------	---	-------	---	-------	---	-------	---	-------	---	-------

Reminder :

Prize U : 0 french Francs credit in a bookstore.
18 months of free access in all parisian cinema.

Prize V : 1000 french Francs credit in a bookstore.
15 months of free access in all parisian cinema.

Prize W : 2000 french Francs credit in a bookstore.
12 months of free access in all parisian cinema.

Prize X : 3000 french Francs credit in a bookstore.
9 months of free access in all parisian cinema.

Prize Y : 4000 french Francs credit in a bookstore.
6 months of free access in all parisian cinema.

Prize Z : 5000 french Francs credit in a bookstore.
3 months of free access in all parisian cinema.

A 30 year old working woman gave us her preference ranking over the prizes. Try to determine (according to this only ranking), the importance coefficient (or weights) that this woman should fix for both aspects of the prizes (cinema and book) :

Woman's preference ranking : > > > > >

Importance coefficients :

Credits in bookstores :

Free cinema access card :

Reminder :

Prize U : 0 french Francs credit in a bookstore.
 18 months of free access in all parisian cinema.

Prize V : 1000 french Francs credit in a bookstore.
 15 months of free access in all parisian cinema.

Prize W : 2000 french Francs credit in a bookstore.
 12 months of free access in all parisian cinema.

Prize X : 3000 french Francs credit in a bookstore.
 9 months of free access in all parisian cinema.

Prize Y : 4000 french Francs credit in a bookstore.
 6 months of free access in all parisian cinema.

Prize Z : 5000 french Francs credit in a bookstore.
 3 months of free access in all parisian cinema.

Please write down how long did it take you to fill in the questionnaire :

Approximate time :

If you have any remarks or comments, please write them down here :

Remarks :

Thank you for participating to this questionnaire. you will be notified of the winner of the drawing.