

## Het bepalen van utiliteiten. Een literatuuroverzicht.

Ben Wilbrink 3 mei 1976.

Gardiner, P.C., & Edwards, W. Public values: multi attribute utility measurement for social decision making. In Kaplan, M.F., & Schwartz, S. (Editors) Human judgment and decision processes. London: Academic Press, 1975

Een belangrijke bijdrage tot praktische oplossingen voor het probleem van toekenning van utiliteiten, al eerder uitvoerig besproken in een memo van 2 maart 1976, waarnaar verwezen.

Vlek, C.A.J., & Wagenaar, W.A. Oordelen en beslissen in onzekerheid. In Michon, J. A., Eijkman, E. G. J., & De Klerk, L. F. W. (redactie) Handboek der psychonomie. Deventer: Van Loghum Slaterus, 1976.

Over nutsaxioma's:

Een aantrekkelijker verzameling axioma's (dan die van Von Neumann & Morgenstern of Savage 1954, b.w.) is voorgesteld door Luce en Krantz, (Conditional expected utility. *Econometrica*, 1971, 39, 253-271); minder technisch uiteengezet in Krantz, Luce, Suppes en Tversky Foundations of measurement. Volume 1, additive and polynomial representations. New York: Academic Press, 1971). In hun theorie van het voorwaardelijk verwachte nut is de verzameling van relevante omstandigheden afhankelijk van het alternatief dat wordt overwogen; nut wordt gedefinieerd over handelwijzen i.p.v. consequenties; kans en nut worden m.b.v. conjuncte meetprocedures simultaan vastgesteld. In het eindige geval kunnen echter Luce en Krantz' (1971) en Savage's (1954) theorieën in elkaar worden vertaald.

Alternatieve beslissingscriteria.

Er moet op worden gewezen dat beslissers rationeel kunnen zijn terwijl zij een ander doel nastreven dan het maximaliseren van verwachte waarde of nut. (is dat zo? Vergelijk opmerkingen van Petersen & Novick 1976) Dit treedt bijvoorbeeld op wanneer de waarschijnlijkheden verbonden met de diverse uitkomsten niet in aanmerking genomen (kunnen) worden. Aan Coombs, Dawes en Tversky (1970, hfdst 5) ontleen we voor dat geval de volgende mogelijkheden;

(a) Maximaxcriterium. kies het alternatief dat onder de gunstigste

omstandigheden de grootste winst oplevert.

(b ) maximincriterium. kies het alternatief dat onder de ongunstigste omstandigheden de grootste winst oplevert.

(c) pessimismeoptimisme criterium: bepaal voor ieder alternatief een gewogen som van de gunstigste en ongunstigste resultaten en kies het alternatief met de hoogste som.

(d) minimax spijtmethode: bereken voor iedere omstandigheid hoeveel spijt u van ieder van de mogelijke keuzen zou hebben, waarbij spijt is gedefinieerd door het verschil tussen de winst bij de beste keuze, en de actuele winst. Kies het alternatief waarbij de grootste spijt zo klein mogelijk is.

(e) Hoogste gemiddelde waarde: middel per alternatief de waarden van de mogelijke consequenties, en kies het alternatief met het hoogste gemiddelde.

Hull, J., Moore, P. G.<sup>5</sup> & Thomas. H. Utility and its measurement. Journal of the Royal Statistical Society, Series A., 1973, 136, 226-247.

'This paper (...) discusses methods of quantifying and measuring a decision maker's preference pattern between alternative outcomes arising from the different courses of action open to him.' Alzo is het niet direct duidelijk of het gaat om descriptieve of normatieve besliskunde (zie bijv. Coombs Dawes & Tversky 1970 blz 114).

"In a general decisionmaking situation preferences are of two types: a) Direct preferences (e.g. I prefer outcome A to outcome B). b) Attitudes to risk (e.g. I would rather play safe and accept outcome C for certain than adopt a course of action giving me a 30 percent chance of outcome A and a 70 per cent chance of outcome B)."

Een beetje een mystieke opmerking; de rest van de tekst zal moeten leren wat de auteurs hier mee bedoelen. Er wordt wel enigszins gezinspeeld op het gevoelsmatig moeilijk liggende punt dat in beslissingsproblemen met onzekere uitkomsten utiliteiten en kansen 'op één hoop gegooid worden' (zie de uiteenzetting van Lindley in zijn Decision making').

Afzonderlijk besproken worden Unidimensionele utiliteiten, Multidimensionele utiliteit, en groeps-utiliteit. Een indeling die ik ook maar over moet nemen om de overzichtelijkheid een beetje in de hand te houden.

## UNIDIMENSIONAL UTILITY.

Methods requiring minimal assumptions.

Voor een volledig overzicht ('all the important methods') wordt verwezen naar het artikel van Fishburn, P.C. Methods of estimating additive utilities.

Management Science, 1967, 13, 435-453. Kennelijk op dat artikel steunend gaan de auteurs verder: There are four main groups:

- a) Direct rating methods
- b) Standard gamble methods
- c) Midpoint methods
- d) Ordered metric methods.

Direct rating.

This is conceptually the simplest method of measurement and is of great practical importance where the outcomes are discrete and do not readily lend themselves to any other sort of quantification. Torgerson (1958, chapter 4) discusses a number of variations on this method. (zal ik afzonderlijk bespreken). "However, direct rating has serious drawbacks when there are other scales of measurement describing the outcomes." (bijv. geldbedragen) "In situations like this, where the outcomes are already ranked by another scale of measurement and a relatively precise risk measure is required, standard gamble methods must be employed."

Standard gamble methods.

"Suggested by Von Neumann and Morgenstern (1947) and developed into the Basic Reference Lottery Ticket methods by Raiffa (1968) these involve establishing the certainty equivalents of a series of simple gambles.

"Consider for example the problem of measuring an investor's utility for quantities of money between £0 and £10,000. Standard gamble methods provide two possible forms for the questioning:

(i) 'Suppose you have two alternative investments: either you can accept a known certain return  $X$ , or you can take a chance on an investment which might be successful and give a return of £10,000 and might fail giving no return. How high would your chance of success have to be before you would take the risky option?'

(ii) 'Suppose you have two alternative investments; either you can choose a safe investment with a fixed return or you can choose an investment which you know has a chance  $P$  of succeeding and yielding a return of £10,000 and  $1-P$  of failing and yielding no return. How high does the return from the safe investment have to be before you would choose it in preference to the risky alternative?'

In (i) X would be systematically varied between £0 and £10,000. In (ii) P would be given different values between 0 and 1. In both cases a relationship is set up between a series of probabilities  $p_1, p_2, \dots, p_n$  and a series of returns  $x_1, \dots, x_n$  such that the decision maker is indifferent between  $x_i$  and a gamble giving him a probability  $p_i$  of £10,000 and  $1-p_i$  of nothing. Values may then be arbitrarily assigned to  $u(£0)$  and  $u(£10,000)$  (as utilities are only unique up to a positive linear transformation) and the expected utility relationship

$$u(x_i) = p_i u(£10,000) + (1-p_i) u(£0)$$

may then be used to obtain points on the utility curve.

A number of studies aimed at evaluating the utility of individual executives for money have been carried out using standard gamble methods (see section 6 this article). Some of these have varied the method so that the standard gambles are not always on the best and worst outcomes but sometimes on two intermediate ones. This does not create any new conceptual difficulties and is a good idea if it makes the analyst's questions more meaningful to the decision maker."

Commentaar. Het is duidelijk dat standard gamble methods een speciaal soort rating methoden zijn. Een optimale uitvoering van deze methoden kan dan ook gebaat zijn bij algemene aanbevelingen die gelden voor rating methoden. Een zo'n aanbeveling is om geen gebruik te maken van extreme waarden waarmee vergelijkingen gemaakt moeten worden, een aanbeveling waar sommige experimentators blijkens de laatst geciteerde zin ook rekening mee hielden. Het kan waarschijnlijk alleen maar goed wanneer dergelijke aanbevelingen iets meer systematisch door mij verzameld zouden worden. Een eerste bron daarvoor is bijv. Hamblin, Social attitudes: magnitude measurement and theory, in Blalock (ed) Measurement in the social sciences 1974, hoofdstuk 3, i.h.b. blz. 64 65).

"In some situations neither direct rating nor standard gamble methods are appropriate because the decision maker is unable to appreciate the quantitative concepts involved. Indeed the average person does not naturally think in terms of a brand of butter being 7/10 of the way up a scale between two other brands or of having a 1/5 chance of making a certain profit, for this reason the third and fourth groups of methods mentioned above are important.

Midpointmethods.

There are two midpoint methods. One involves only the concept of a 50 : 50 gamble; the other involves only the concept of one outcome being rated half

way between two other outcomes. In both the decision maker is asked to bisect successively the interval between the best and worst possible outcomes (i.e. first the interval is divided into two, then each half is divided into two, etc.). The questions asked in the methods take one of the following forms:

(i) What in your opinion is the outcome half way between outcomes A and B, and

(ii) which outcome would you consider equivalent to a 50 : 50 gamble on outcomes A and B.

Both methods are restricted to situations where the variable is continuous and are really just variations on direct rating and standard gamble methods respectively. Torgerson (1958, chapters 5 and 6) discusses method (i) in the general context of all fractionation and equisection methods.

Commentaar. Voor de eerste methode zullen we t.z.t. Torgerson bekijken. De tweede methode is de voorkeursaanpak van Bayesianen als Novick en Jackson. Daarvoor zal ik t.z.t. hun tekst eens bekijken. De aanpak van Novick c.s. is te meer belangrijk, omdat het computer orakel CADA nu is uitgebreid met een afdeling waarin het bepalen van utiliteiten voorgeprogrammeerd is. Mededeling daarover in Novick, M. R., & Petersen, N. S. Towards equalizing educational and employment opportunity *Journal of Educational Measurement* 1976, 13, 7788:(zal in het vervolg nog besproken worden).

Ordered metric methods.

"These methods (the name is due to Coombs, 1950) apply only when the utility is a function of a discrete variable. The outcomes are first ordered and then difference between the outcomes are ordered. The latter can be achieved in two ways:

(i) By asking whether the difference between outcomes A and B is greater than the difference between the outcomes C and D.

(ii) By asking whether a 50 ; 50 gamble on A and D is preferred to a 50 : 50 gamble on B and C.

Numerical assignments to the utility of each outcome must then be made in such a way that they are consistent with the two sets of rankings. One possibility is to set  $u(\text{outcome ranked last})=0$  and then to set  $u(\text{outcome ranked } k)u(\text{outcome ranked } k+1)=n$  where this difference is the Nth one in the second

ranking.

Commentaar. Niet verwezen wordt naar relevante literatuur. Een aardig artikel is bijv. Long, J.F., & Wilken, P. H. A fully nonmetric unfolding technique: interval values from ordinal data. In Blalock (ed) Measurement in the social sciences 1974, hoofdstuk 2. Te zijner tijd te bespreken dus.

Methods where assumptions are made.

"Some of the methods described (in de voorgaande blz.) become extremely tedious when the utility function is required for a wide range of outcomes and this has prompted many writers to make simplifying assumptions. These assumption usually reduce the measurement problem from one of obtaining a whole curve to one of determining a few parameters in a functional form.

Een koortsachtige theoretische activiteit m.b.t. deze methoden heeft plaats gevonden op het gebied van portfolio investment selection, waarvan de auteurs een kort overzicht geven. Hoewel, uiteindelijk is het oordeel nogal sobertjes:

"Instead of presenting the decisionmaker with a whole series of Raiffa type gambles, it should eventually only be necessary to ask a few questions phrased in terms of his (the decision maker's, b.w.) objectives, riskaversion, etc. However, for this to be possible it is necessary for the decision maker to acquire an understanding of the underlying concepts such as expected compound returns, decreasing riskaversion, etc. and a great deal more work needs to be done in the future in order to develop methods for making these concepts more meaningful to him."

#### MULTIDIMENSIONAL UTILITY.

"The use of more than one variable to describe outcomes is a matter of convenience. However, when measurement is considered, new practical problems arise simply because the number of possible outcomes has increased by at least one order of magnitude. Consider, for example, using Raiffa's method to evaluate a corporate decision maker's utility function for asset position, market share and turnover. This would involve specifying a number of different levels, determining the equivalent standard gambles on the best and worst outcomes, i.e. it would involve asking the corporate decision maker at least one question for every point on a threedimensional grid. This sort of approach is clearly out of the question. Some assumptions must be made about the nature of the utility function in order to make measurements a practical possibility. In this section we discuss separately four different assumptions:

- a) a linear utility function.
- b) an additive utility function.
- c) a lexicographic utility function.
- d) a utility function with some other specific property.

Linear utility functions.

"Linearity is the simplest assumption applicable when each dimension lends itself to direct quantification on a scale of measurement (e.g. £, minutes, inches, share of market, etc.). It states that

$$u(x_1, x_2, \dots, x_n) = \sum a_i x_i,$$

where  $u$  is the utility function,  $x_i$  is the value of the variable corresponding to the  $i$ th dimension and  $a_i$  are constants ( $i = 1, 2, \dots, n$ ). Linearity implies that there is a constant rate of tradeoff between one dimension and another (i.e.  $a_i$  units of variable  $x_i$  are always equivalent to  $a_j$  units of variable  $x_j$ ).

In the example given above involving the corporate decision maker, this means that he would be willing to make a series of statements of the form:

No matter what the value of my asset position, market share and turnover, I would always consider a 1 percentage unit increase in market share and a £100,000 increase in assets equally valuable.

The measurement of a linear utility function involves determining a number of such tradeoffs. If there are  $N$  dimensions, then only  $N-1$  tradeoffs need to be evaluated. Unless  $N$  is very large the measurement process is not therefore very tedious once linearity has been established. It is interesting to note that in spite of its restrictive properties the linear utility assumption is often made implicitly in studies in economics and operational research. Cost benefit analyses, for example, almost invariably assume that all costs and benefits, whatever the units they are measured in, can be traded off at fixed rates for money. These assumptions are unlikely to be perfectly true and the extent of their validity should always be investigated."

Additive utility functions.

"Additivity is a slightly more general assumption than linearity. It states that

$$u(x_1, x_2, \dots, x_n) = \sum u_i(x_i),$$

where  $u_i$  is a function of the variable  $x$  corresponding to the  $i$ th dimension.

There are a number of ways of viewing this assumption. Tests for its validity have been constructed by Adams and Fagot (1959, Behav. Sci. 4,110), Luce and Tukey (1964, J Math. Psychol., 1, 1 27), and Fishburn (1965, Operat. Res. 13, 28~45; 1965, Man. Sci., 11,792801; 1966 Operat. Res. 14, 1143-1148)

Basically it implies that the factors are valuewise independent or that the utility of the whole equals the sum of the utilities of the parts. In terms of tradeoffs it requires that the rate of tradeoff between two variables depends only on the values of those variables and not on that of the others. Wanneer alle factoren binair zijn (slechts twee waarden aan kunnen nemen) is het meetprobleem eenvoudig op te lossen omdat alle unidimensionele methoden daarvoor in principe beschikbaar zijn. Het lijkt me dat Cronbach deze idee in zijn hoofd had bij het schrijven van 'Equity in selection Where psychometrics and political philosophy meet. Journal of Educational Measurement 1976, 13, 3141.

"Unfortunately, measuring the general additive utility is considerably more complicated than this. In equation 1 (hierboven)  $u$ , can clearly be interpreted as the unidimensional utility function when all the variables except  $x_i$  are kept fixed (i.e. when  $x$  is the only variable describing the outcomes). However, it is not sufficient simply to determine the  $u_i$ 's individually by the unidimensional methods described in section 2. For each  $u_i$  would then only be determined up to a linear transformation, and linear transformations of each  $u_i$ , do not give necessarily give a linear transformation of  $\sum u_i$ . It is clear that the dimensions must be rated in importance relatively to each other in some way. (zoals bijv. Gardiner en Edwards in hun al besproken artikel deden) (Verwezen wordt naar Turban and Metersky, Management Science 1971, voor een voorbeeld).

In some situations (e.g. the asset position, the market turnover share turnover example) rating dimensions in importance may not be a very natural procedure for the decision maker. Establishing tradeoffs between the dimensions may be more efficient. Fishburn (1967) lists a number of methods for this. They fall into two general categories:

- (i) Methods where one  $u_i$  is assumed to be known and is used to determine another  $u_i$  (known as scaling methods).
- (ii) Methods where two  $u_i$ 's are determined simultaneously (known as flight of stairs methods).

In (i) questioning is generally aimed at trading off increases in one dimension with decreases in another. In the corporate decision maker example

given above, suppose that the u. corresponding to assets has already been determined. Questioning would take the form:

If initially you have the assets of A and a market share of B, to what value would your market share have to be increased in order to compensate for a decrease in the level of your assets to C.

Either A and B are kept fixed and C is varied, or B en C, are kept fixed and A is varied. In the first case an 'indifference curve' is obtained; In the second case a 'transformation curve' is obtained. In both cases at least part of the u. function corresponding to market share is known at the end of the process.

Voor bespreking van de tweede soort methoden, zie het artikel op blz. 238; toepasbaarheid van die methoden is me niet geheel duidelijk.

Lexicographic utilities.

"Lexicographic utilities describe situations where one factor is of overwhelming importance and in a choice between two outcomes a second factor is taken into account only when there is a tie on the first. Measurement problems are more akin to those arising from unidimensional situations and do not appear to present any special extra difficulties."

Other utility structures.

"The only alternative structure not already mentioned is due to Keeney (1968, 1972 (de laatste in Management Science: The basic assumption is that tradeoffs between gambles on only one of the variables are independent of the values of the other variables. This is slightly less strict than additivity and Keeney shows that in the twodimensional case it leads to the functional form:

$$u(x_1, x_2) = u_1(x_1) + u_2(x_2) + K u_1(x_1) u_2(x_2)."$$

## GROUP UTILITY

Stemprocedures zijn besproken door Fishburn (1971 Behavioral Science, t.z.t. te bespreken). Onderhandelingsprocedures vormen een andere categorie, waarbij forse theoretische problemen bestaan waar het op beschrijven aan komt, problemen die heel summier door de auteurs besproken worden. "In practical studies, however, where the objective is prescription rather than description, these considerations rarely pose any great problems. An analyst normally tries

to determine in two separate and distinct stages a group probability distribution and a group utility function and then maximizes expected utility. For, although this does not ensure Pareto optimality for the group, it does make the problem manageable and give it the right logical structure. The measurement of group utility, in spite of the theoretical consideration mentioned above, is therefore an important practical problem. Spetzler (1968 IEEE Trans. Systems Sci. Cybern., SSC4,279300) has made by far the most important contribution in this area. He considered the problem of prescribing a risk policy for making capital investment decisions in a corporation and adopted the following general procedure:

a) Present a number of executives in the company with a series of hypothetical investment decisions and from their answers plot utility functions (one for each executive).

b) Agree with the help of further questioning on the form of functional form to fit the utility plots (the same functional form but with different parameters being used for each executive).

c) Reinterview, allowing each executive to spend a great deal of time on a few questions and from the results determine, for each executive, the parameters  $k_i$  in the function.

d) Present the results to top management and attempt to iron out in discussion with them the best parameters for a corporate risk policy.

Spetzler found that at all levels of investment the final corporate utility function was far less conservative than the average for the executives interviewed. (This can be viewed as a particular example of the 'risky shift phenomenon' which has been well documented by psychologists.)

Al met al lijkt dit artikel op het eerste gezicht een hoop informatie te verschaffen, maar bij nader inzien is het toch een tamelijk oppervlakkige babbel, die mogelijk wat helderder te maken is door literatuur waarnaar verwezen wordt nog eens te bestuderen. Zoals terloops al aangetekend werd, is er nogal wat literatuur die, hoewel direct relevant, niet genoemd wordt. Ook literatuur van recenter datum dus, waarvoor de auteurs geen verwijt gemaakt kan worden. Onder de laatste categorie valt bijv. Zeleny (ed.) Multiple Criteria Decision Making 1975 Springer Verlag, dat naar het zich laat aanzien nog een uitgebreide bespreking in dit stuk zal krijgen.

Coombs, C. H., Dawes, R. M., & Tversky, A. Mathematical Psychology. Englewood Cliffs, N.J.: PrenticeHall, 1970. (Hoofdstuk 5: Individual decision making). Interessant is de tweedeling in beslissingsproblemen die gemaakt wordt:

"This chapter deals with two types of decision problems: decisions with incomplete knowledge and decisions with unsure preferences. Both types of decision problems are accompanied by uncertainty. In the former type, however, the uncertainty concerns the future state of the world, whereas in the latter it concerns the decision maker's own state of mind."

Op blz. 114 wordt een korte maar heldere uiteenzetting gegeven over het onderscheid en de ononderscheidbaarheid a.h.w. van de beschrijvende en de voorschrijvende aanpak van beslissingsprocessen.

Decisions with incomplete knowledge: theories of risky choice.

"Decisions with incomplete knowledge, where one does not know for sure which state of the world will, in fact, obtain, are typically represented in the form of a payoff matrix."

Merk op de koppeling van payoff matrix representatie aan dit type beslissingen. Zou je kunnen zeggen dat Cronbach & Gleser's studie vooral dit soort beslissingen behelst? Er wordt op gewezen dat niet altijd een duidelijke representatie in payoff matrix vorm te vinden is.

"there are many ways of structuring or representing a decision problem in payoff matrix form. The art of finding the 'right way' of structuring a decision problem contributes a great deal to its successful solution. An appropriate representation of a decision problem in a payoff matrix form, therefore, is the result of an adequate formulation rather than a substitute for one. It is nevertheless a useful analytic tool."

"In decision making under risk, it is assumed that the individual can evaluate the likelihoods of the various states of nature. More specifically, his beliefs about the likelihoods of the relevant states can be expressed by some (possibly subjective) probability distribution. Risky choices are essentially gambles whose outcomes are determined jointly by the choice of the individual and the result of some specified random process."

Een interessante aantekening is: "More than 100 years later, the Bernoullian notion of a subjective scale (voor utiliteit nl., b.w.) became the cornerstone of quantitative psychophysics founded by G.T. Fechner."

"... the decreasing marginal utility hypothesis according to which the utility function is concave, or negatively accelerated. (The similarity to Weber's law is not accidental. In fact, the logarithmic utility function proposed by Bernoulli was the one put forth by Fechner as the form of the general

psychophysical law.) Interessant, omdat technieken voor het meten van utiliteiten gevonden kunnen worden in literatuur over methoden gebruikt bij psychofysisch onderzoek (bijv. Torgerson 1958).

Na bespreking van de verklarende kracht van het principe van de verwachte utiliteit:

"It is important to realize that we have not shown that people insure their property or gamble on horses in order to maximize some utility function. All we have shown is that although these phenomena are incompatible with the expected value principle, they can be accounted for, if an appropriate utility function is introduced.

Opm. Heeft dat misschien iets te maken met het feit dat de utiliteitsfunctie in deze gevallen een hypothetische constructie is van een aantal psychologen, en niet een functie die door de betreffende personen zelf is opgesteld? Voor het laatste, het construeren van utiliteitsfuncties, kunnen een aantal verschillende technieken aangeboden worden. Technieken, die niet noodzakelijk allen zullen resulteren in dezelfde utiliteitsfuncties, waarmee de volgende vragen gesteld worden:

- hoe vergelijken verschillende methoden voor het kiezen van utiliteitsfuncties zich met elkaar?

- Op welke dimensies zijn dergelijke vergelijkingen relevant?

- Wat moeten we precies met de aldus verkregen utiliteitsfuncties aan? Moeten ze dienen als verklarende principes voor gedrag dat ook zónder al deze expliciteringen plaats vindt ( het voorbeeld van verzekeringen en kleine gokkerijen)? Of gaat het er juist om te laten zien dat er discrepanties zijn tussen keuzen die naïef gemaakt worden en keuzen die langs deze geëxpliciteerde weg gemaakt worden, discrepanties bovendien die belangrijk genoeg zijn om de extra moeite van vaststelling van utiliteiten e.d. te rechtvaardigen,

- die door de betrokken personen worden ervaren als te zijn belangrijke verbeteringen in keuzeresultaat van de expliciete methode vergeleken met de naïeve

- en hoe valideer je dan weer dat gevoel van de betrokken personen dat ze er met die expliciete besliskunde inderdaad op vooruit gaan?

- kortom, hoe kom je er achter of een stukje besliskunde dat in situaties met harde criteria (guldens etc.) mogelijk goed functioneert, dat ook doet in situaties waar alles wat vager wordt, persoonlijke satisfacties een rol spelen etcetera.

"In spite of the intuitive mark appeal of this rationale, the application of the expected utility principle to essentially unique (niet herhaalbare) choice

situations requires an independent justification. Modern utility theory provides such a justification in the form of an axiomatic foundation of the expected utility principle."

Opm. Cronbach en Gleser 1965 hadden er ook al moeite mee: bij institutionele selectie kun je spreken van zich steeds herhalende identieke beslissingsproblemen. Daarvoor zou je het principe dan zonder meer toe kunnen passen. Voor individuele selectie modellen geldt dat de keuze eenmalig is, en dan zitten we voor een situatie waar het citaat uit Coombs. e.s. op slaat. Alleen, het probleem was al door Von Neumann en Morgenstern opgelost, d.w.z. al lang voor de eerste editie (1957) van Cronbach en Gleser's studie.

### Modern utility theory.

"Modern utility theory was first developed by von Neumann and Morgenstern as an appendix to their famous Theory of games and economic behavior (1947), The theory consists of a set of axioms about preferences among gambles. The basic result of the theory is summarized by a theorem stating that if an individual's preferences satisfy the specified axioms then his behavior can be described, or rationalized, as the maximization of his expected utility. Because the axioms can be regarded as maxims of rational behavior, they provide a normative justification for the expected utility principle."

Een toestand om eens even rustig over na te denken. Over dergelijke axiomata valt inhoudelijk ook best nog wat te twisten (o.a. Suppes 1974 over de axioma's van Savage).

"It is important to realize that utility theory was developed as a prescriptive theory, justified on the basis of normative considerations alone. The close interrelationships between normative and descriptive considerations, however, suggest that utility theory may also be used as a psychological theory of decision making under risk."

### the subjective expected utility model.

"Another criticism of the expected utility theory revolves around the concept of probability. The theory is formulated in terms of gambles whose numerical probabilities are assumed to be known in advance. Such knowledge, however, is missing in most applications. Can the theory be generalized to situations where no a priori knowledge of (numerical) probabilities is available? The answer is positive, provided some consistency requirements are fulfilled. Savage (1954) has developed an axiomatic theory leading to simultaneous measurement of utility and subjective probability. We do not wish to present this theory here, but do wish to show how the relation 'more probable than' between events can

be defined in terms of preferences between gambles." Zie hun tekst. Savage has further shown that his axioms are sufficient to establish the existence of a uniquely additive subjective probability function  $s$  and an interval scale utility function  $u$  such that

- 1)  $x \succ y$  if and only if  $u(x) > u(y)$  and
- 2)  $u(x, E, y) = s(E)u(x) + (1-s(E))u(y)$ ,

where  $(x, E, y)$  denotes the gamble where  $x$  is obtained if  $E$  occurs and  $y$  otherwise.

"A more recent attempt to test utility theory and to measure utility and subjective probability simultaneously, was conducted by Tversky (1967) *Journal of Experimental Psychology*). Consider a set of gambles of the form  $(x, p)$  in which one wins (or loses)  $\$x$  if  $p$  occurs etcetera etcetera, zie de tekst blz. 134-137. Interessante items in die bespreking: een methode voor het toetsen van de aanname van onafhankelijkheid van utiliteit en subjectieve waarschijnlijkheid; onderzoek van de veronderstelling dat utiliteit een machtsfunctie is. "From a methodological viewpoint it was shown that the application of additivity analysis to a specified class of gambling experiments yields a simple simultaneous construction of utility and subjective probability scales. From a substantive viewpoint the data support the multiplicative relation between utility and subjective probability embodied in the subjective expected utility principle. Moreover, the results indicate that (within the range of payoffs investigated) the utility of money can be described as a power function of money with different exponents for positive and negative outcomes."

special preference models.

The subjective expected utility model is not the only theory put forth to account for decisions under risk. More than half a century ago, an economist, Irving Fisher (1906), proposed that people base their choices among gambles as well as on the expectations. The variance of a gamble  $G$ , denoted  $V(G)$ , is given by the formula  $V(G) = E(G^2) - E(G)^2$ , where  $G^2$  is the gamble obtained from  $G$  by squaring its outcomes. In particular, if  $G$  has two outcomes  $x$  and  $y$  obtained with probabilities  $p$  and  $1-p$ , respectively, then its variance equals  $p(1-p)(x-y)^2$ .

$$\text{namelijk } V(G) = x^2 p + y^2 (1-p) - (xp + y(1-p))^2 = p(1-p)(x-y)^2.$$

The variance is the most common measure of the dispersion of the outcomes, and it seems conceivable that different attitudes toward risk be reflected by

preferences for different amounts of variance. Volgens de auteurs kunnen sommige soorten van variantie voorkeuren wel opgevangen worden door een geschikte keuze van utiliteitsfuncties. Zij geven echter geen procedure aan, noch een literatuur verwijzing (wordt dus een beetje moeilijk misschien).

Het is erg jammer, maar deze tekst gaat op deze voorkeur modellen niet verder in.

Alternatives to utility theory can be developed by first characterizing each gamble in terms of its objectively defined attributes and then formulating decision rules based on these attributes. Thus, instead of choosing among gambles according to their subjective expected utilities, one can choose among them according to some linear combination of their expectations and variances, for example.

Onduidelijk; waarschijnlijk gaat het toch om verwachtingen en varianties van utiliteiten, anders lijkt me deze aanpak wel heel erg beperkt bruikbaar. Uitzoeken dus nog.

The major difficulty with this approach lies in the need to specify in advance the relevant attributes, or dimensions, whose combination accounts for the data. On the other hand, this approach might be a great deal simpler than utility theory in that only a few linear weights, instead of an entire utility function, would be required to explain decisions under risk. (Alleen descriptief interessant dus? Of zouden die descripties bruikbaar zijn om in een soort iteratieve procedure de beslissing nader tot andere en betere gedachten, zeg meer coherente gedachten, te brengen?)

The most general theory of this type is Coombs' ideal point model (1964) and the unfolding method associated with it (see Sec. 3.4). Variantie in uitkomsten speelt in deze modellen een belangrijke rol: "In this model each gamble with monetary outcomes is represented as a probability distribution over the possible outcomes. The gamble obtained by tossing a fair coin 100 times, with heads winning and tails losing, for example, is represented by a binomial distribution with  $p=0,5$  and  $n=100$ . For any expectation level each individual is assumed to have an ideal distribution, that is, a gamble that he likes best. In choosing among gambles, the individual selects the one that is 'closest' to his ideal gamble, in some welldefined sense. Although this model has not been treated in full generality, several studies have demonstrated systematic preferences for some specific aspects of gambles.

Het lijkt er op dat deze modellen niet toegesneden zijn op aanpak van eenmalige keuzeproblemen (zoals studiekeuze door de student).

If Coombs' unfolding theory is applied to variance preference (1variantiel

is dan kennelijk een relevante 'dimensie'), it implies that each individual has an ideal point somewhere along the variance continuum that corresponds to the amount of variance he prefers most. Moreover, his preference for variance decreases monotonically on both sides of the ideal point. Put differently, an individual's preference order, with respect to a given dimension, can be described by a singlepeaked preference function over that dimension. (Think of the amount of sugar you prefer in your coffee or of the room temperature you prefer as cases in point.).

Tenslotte, de theorie rond achievement motivation (bijv. Atkinson 1964), levert een interessante aanpak van beslissingsproblemen waarbij kansen op succes niet zozeer van toeval afhankelijk zijn, als wel van eigen inspanning. Een stukje differentiële psychologie wordt ingebracht door de veronderstelling dat verschillen in motivatie (achievement motivatie en motivatie om mislukkingen te vermijden) leiden tot verschillende 'optimale' keuzen (?) voor verschillende personen in dezelfde situatie. In eerste instantie een beschrijvende theorie dus, die echter in het oog gehouden moet worden omdat bijvoorbeeld waar het gaat om selectieprocedures voor het hoger onderwijs voor bepaalde selectiemodellen wel eens zou kunnen gelden dat ze minder opgaan in de mate waarin Atkinson's theorie van toepassing is. (individuele beslissingen van aspirant studenten zijn daar natuurlijk bijzonder vatbaar voor, maar mogelijk ook de voorkeuren die, zeg, hoogleraren hebben voor bepaalde deelgroepen uit de aankomende studentengroepering). Hoewel een en ander in de verste verte nog niet hard uit te werken lijkt, is een stukje filosofie en discussie van belang om andere modellen enigszins te relativiseren, en mogelijk ook om nog extra en nieuwe alternatieven op te sporen waar het gaat om selectiemodellen waaruit te kiezen valt.

simplification principles. Decisions under 'ignorance'

Een belangrijke aanname bij alle voorgaande modellen is dat alle beschikbare alternatieven en alle 'states' en 'outcomes' geëvalueerd kunnen worden. Vaak is dat niet realistisch, en moet naar vereenvoudiging gezocht worden. Allereerst hebben we dan een categorie modellen voor beslissingen waarbij verondersteld wordt dat niets bekend is over waarschijnlijkheden van 'states of nature' bijv. Daarnaast is het belangrijkste bezwaar tegen domweg toepassen van deze criteria al aangegeven. Meestal zullen deze modellen dus domweg irrationeel zijn omdat zij ook met het kleine beetje informatie dat we meestal toch wel hebben, geen rekening houden. (vgl. de opmerking van Vlek en Wagenaar of blz. 1 van dit overzicht). decision criteria that do not depend on likelihood considerations:

1) maximin criterion; select an alternative that maximizes the minimum payoff.

It is a very conservative criterion and hence recommendable in those contexts where extreme caution is desired. Maar in medische context, bijv., moeilijk toepasbaar omdat risico's voor alle alternatieven vaak ernstig zijn, en omdat waarschijnlijkheden vrijwel altijd te schatten zijn, (b.w.). Toepasbaar 'when the state of nature is determined by a hostile opponent'.

2) maximax criterion: choose an alternative whose highest value is maximal. 'optimal if the best that could happen always should happen'.

3) pessimism optimism criterion: een procedure die een weging is van de beide eerste. (Zie voor het wegingsvoorschrift de tekst).

4) principle of insufficient reason. Kent aan alle 'states of nature' een even grote waarschijnlijkheid toe. Omdat echter het aantal states of nature tamelijk willekeurig is, zijn deze waarschijnlijkheden, en daarmee de resultaten van de procedure, eveneens willekeurig.

5) minimax regret (zie beschrijving blz.1) (nee, doe dat niet: die beschrijving is duister). For each state  $s(j)$  let  $v(j)$  denote the highest value attainable under that state. That is,  $v(j) = \max_i v(ij)$ . To each payoff matrix, associate a regret matrix with entries  $r(ij) = v(j) - v(ij)$ . Thus each entry of the regret matrix is a measure of the difference between the payoff that is actually obtained and the payoff that could have been obtained had if the true state of nature had been known in advance. To each alternative assign its minimax maximal regret value and select an alternative whose maximal regret is minimal. Thus the minimax regret criterion is similar to the maximin criterion in that it concentrates on the worst possible case, but the worst is defined by the maximal regret rather than by the minimal value.

the satisficing principle.

Een heel andere aanpak is die van Simon (1957).

The essential simplification built into Simon's model can be described as the replacement of the maximization principle by a satisficing principle. According to this principle, the outcomes are first classified as 'satisfactory' or 'unsatisfactory' with respect to each of the relevant attributes, and the first alternative that satisfies one's level of aspiration along each attribute is selected." "Because, according to Simon's view, each complex decision problem is reformulated by the individual, the model must characterize the subjective process of defining the problem. To do so, a theory describing the cognitive processes involved in the reformulation of a decision problem, is required. This requirement has resulted in a shift in emphasis from motivation to cognition as the psychological basis of a theory of choice."

"The General Problem Solver has been quite succesfull in simulating human behavior in simple learning tasks as well as in some complex decision problems. (In one of the more interesting applications, Clarkson 1962 employed the general problem solver to investigate the decision process of investment trust officers in a middlesize bank).

revision of subjective probabilities.

Tenslotte "we discuss the problem of how people revise their subjective probabilities in the light of new information.

"Do people revise their subjective probabilities in accordance with Bayes' rule? Ward Edwards and his associates at the university of Michigan have investigated this question in great detail."

Belangrijk resultaat daarvan:

"People are conservative probability estimators in the sense that their estimates are considerably less extreme than those calculated from Bayes' rule."

Of, in de woorden van Edwards, Lindman & Phillips 1965:

"Men are incapable of extracting all of the certainty from information that Bayes theorem indicates in that information."

Op de een of andere manier moet hier een aantal opzienbarende lessen uit te peuren zijn voor waar het gaat om afzetten van rationele beslissingsmodellen voor selectie tegenover a priori ideeën van betrokken personen over de mogelijkheden en onmogelijkheden van deze selectiemodellen. Ook voor gewone beoordelingsprocedures in het onderwijs zou dit verschijnsel van betekenis kunnen zijn (en, mits onderkend, kunnen leiden tot veel efficiënter doorstromingsbeleid).

probabilistic theories of choice.

constant utility model (Luce 1959) (strong utility models) random utility model (Coombs) intransitivity and the evaluation of multidimensional alternatives (Tversky 1969) Weer een nieuwe wereld, maar nogal aan de descriptieve kant lijkt het, dus een meer intensieve bespreking lijkt voorlopig niet nodig.