

het nemen van beslissingen. nieuwe literatuur.  
2-3-76.

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

outcome evaluation :

"Another way of putting the question to be answered in outcome evaluation is 'Suppose I could know for certain that this particular act would lead to this particular outcome, and suppose I then chose this act. How attractive would its outcome be to me, or to someone, or to some collection of people?'

Note that outcomes, not acts, are evaluated. We often think of values as being attached to acts. That is in a sense appropriate, since the whole purpose of obtaining values of outcomes and probabilities of states is to permit the simple computation that associates an expected value with an act. In the special case of no uncertainty (that is, some state has probability 1, or, more often, you treat it as though it had probability 1 even though you know that unlikely alternative states exist), each act has only one outcome, and so the value of the outcome is also the value of the act."

Over het evalueren van uitkomsten in dollars maken de auteurs de volgende interessante, voor de hand

liggende, maar voor mij in deze context toch nieuwe, opmerking:

"The remark 'That's so cheap it can't be any good' is true enough often enough to illustrate the phenomenon, and its frequent falsity illustrates how unsatisfactory the procedure is. Price depends more on the relation between supply and demand, than on value. Air is free but if some predatory corporation or country had a monopoly on it, wouldn't you be willing to pay any amount within your means for a year's supply?"

Over het succes van de hier gepropageerde aanpak: "Unpublished studies strongly argue that the simple ratingscale procedures described below produce results essentially the same as much more complicated procedures involving imaginary lotteries."

Simple Multi Attribute Rating Technique (SMART) in tien stappen:

1. Voor welke persoon, personen, organisatie moeten de utiliteiten gemaximaliseerd worden? Probeer de medewerking van deze personen of vertegenwoordigers van de organisatie te krijgen.

2. Beschrijf het beslissingsprobleem waarvoor de utiliteiten nodig zijn. "The same objects or acts may have different values, depending on context and purpose. In general, utility is a function of the evaluator, the entity being evaluated, and the purpose for which the evaluation is being made. The third argument of that

function is sometimes neglected." Deze opmerking lijkt ook van belang voor het uitvoeren van checks op de consistentie van de beslissingnemer: consistentie hoeft dan niet te betekenen dat aan dezelfde uitkomst altijd dezelfde utiliteit wordt toegekend omdat de aard van het beslissingsprobleem kan verschillen (als moderator functioneert) (b.w.).

beslissingen 2 2-3-76.

3. Beschrijf de te evalueren uitkomsten. Bedenk daarbij dat het onderscheid tussen uitkomsten en acties ambigu is: uitkomsten zijn vrijwel altijd gelegenheden tot verdere acties die weer tot een groot aantal andere uitkomsten kunnen leiden, etcetera. De auteurs lijken aan te bevelen als uitkomsten te beschrijven het soort zaken dat zich voor dat soort beslissingsprobleem het best daarvoor blijkt te lenen. De ervaring zal dan uitwijzen op welk punt in de zich eindeloos vertakkende boom van acties en uitkomsten gekapt kan worden.

"Often, in practice, it is sufficient to treat an action itself as an outcome. This amounts to treating the action as having an inevitable outcome that is, of assuming that uncertainty about outcomes is not involved in the evaluation of that action. Paradoxically, this is often a good technique when the outcome is utterly uncertain so uncertain that it is impractical or not worthwhile to explore all its possible consequences in detail and assign probabilities to each."

4. Beschrijf welke waarde dimensies aan de orde moeten komen. Uit de illustratieve uitwerking die de auteurs verderop geven blijkt dat hiermee bedoeld wordt de eigenschappen van de uitkomsten die van belang zijn bij het afwegen van voor en tegens van alternatieven.

Een andere manier van spreken: i.p.v. relevante eigenschappen kun je de doelen beschrijven die je wilt bereiken. Dergelijke doelen zijn vaak hiërarchisch t.o.v. elkaar geordend, maar de auteurs bevelen aan die ordening eerst te negeren en eenvoudig de relevante doelen op te sommen. Zijn dit er veel (zeg meer dan tien) dan kun je beter de minder belangrijke doelen schrappen, of sommige afzonderlijke doelen combineren tot één doel.

de dimensies van stap 4 worden nu gerangordend. "This ranking job, like step 4, can be performed either by an individual, by representatives of conflicting values acting separately, or by those representatives acting as a group. Our preferred technique is to try group processes (Delphi Technique ? b.w.) first, mostly to get the arguments on the table, and to make it more likely that the participants start from a common information base, and then to get separate judgments from each individual.

beslissingen 3 2-3-76.

6. Beoordeel de dimensies op belangrijkheid,

uitgaande van score 10 voor de minst belangrijke dimensie. Ga voor de tweede dimensie na hoeveel belangrijker die dimensie is dan de eerste dimensie.

Voor de derde dimensie, idem, maar ga ook na hoeveel belangrijker deze dimensie is dan de tweede dimensie. Als de laatste beoordeling een inconsistent resultaat oplevert, wijzig dan de beoordeling van de tweede dimensie zodanig dat de zaak wél klopt. Dergelijke veranderingen komen vaak voor, en zijn geen schande. Ga zo te werk bij alle volgende dimensies: vergelijk iedere nieuwe dimensie afzonderlijk met alle vorige dimensies.

Opmerking 1. De auteurs beschrijven hier geen technieken om een en ander rap en effiegent uit te voeren.

Opmerking 2. De afweging van 'belangrijkheid' lijkt nogal fors aan de intuïtieve kant te zijn. Ik ben erg geïnteresseerd in de argumenten waarom deze werkwijze de beslissingsprocedure niet zóveel vertroebelt dat het hele zaakje rampetamp wordt. Overweging is misschien dat de ingebouwde vaagheid weinig kwaad kan zolang de beoordelaar maar dezelfde werkwijze volgt door de hele reeks dimensies heen. Maar ook dit argument zou ik toch graag hard gemaakt willen zien.

Voor iedere dimensie hebben we dan een gewicht, dat de belangrijkheid aangeeft.

7. Deze ruwe gewichten worden gefatsoeneerd : tel ze op, deel ieder gewicht door dat totaal, en vermenigvuldig voor ieder gewicht met 100. (je kunt het laatste natuurlijk ook achterwege laten, als dat makkelijker is).

8. De tekst is hier enigszins aan de onduidelijke kant. De stap bestaat uit twee deelstappen:

8a). Iedere dimensie kent zo zijn eigen schaal: dat kan zijn centimeters lengte, graden temperatuur, mate van aantrekkelijkheid, veel of weinig, aantallen, guldens etc. Geef nu op de schaal voor iedere dimensie afzonderlijk de punten aan waarvoor je bereid bent te zeggen dat het schaalpunt geen enkele waarde heeft, resp. zijn maximale waarde ' bereikt. (minimum plausible value, resp. maximum plausible value.) Minimum resp. maximum utiliteit dus, en niet minimum of maximum bereikbare schaalwaarden.

Het blijft een beetje van de onduidelijke voorlopig. Maar goed, hebben we die punten aangegeven, dan verbinden we ze door een rechte lijn, wat onze 'utiliteits-kromme' is. Een kromme die in dit geval recht is, wanneer we de auteurs willen volgen:

"Strong reasons argue for the straightline procedure whenever the underlying dimension is conditionally monotonic. These reasons essentially are that such straight lines will produce magnificent approximations to the true value functions after aggregation over dimensions; correlations in excess of .99 are typical."

In gevallen waarin toch de nodige twijfels bestaan, kunnen natuurlijk ook kromlijnige utiliteits curven gehanteerd worden, zoals o.a. ook door de auteurs gedaan is in hun case study.

8b). Aan iedere uitkomst wordt voor iedere dimensie afzonderlijk een schaalwaarde toegekend door meting, schatting, beoordeling of iets dergelijks. Omdat de utiliteits krommen al bekend zijn, en bijv. uitgetekend in diagrammen van schaal van de dimensie tegen schaal van de utiliteiten (de laatste van 0 tot 100, of als de gewichten gestandaardiseerd zijn op 0 tot 1, dan van 0 tot 1), kan de utiliteit van een bepaalde uitkomst op een bepaalde dimensie zonder meer afgelezen worden.

beslissingen 4 2-3-76.

Op dit punt geven de auteurs een korte toelichting op het begrip 'belangrijk' zoals dat in stap 6 aan de orde was.

'Formally, judgments at step 6 should be designed so that when the output of step 7 (or of step 6, which differs only by a linear transformation) is multiplied by the output of step 8, equal numerical distances between these products on different dimensions correspond to equal changes in desirability.'

Het voorbeeld dat ze geven loopt ongeveer als volgt: er zijn twee mogelijke uitkomsten, en twee dimensies die

van belang zijn, waarbij de tweede dimensie 'twee keer zo belangrijk' is als de eerste.

Stel uitkomst A en B hebben een zodanige plaats op de eerste dimensie dat die overeenkomt met de utiliteiten 50, respectievelijk 10. Idem op de tweede dimensie A utiliteit 70 en B utiliteit 90. De utiliteit voor A is dan  $50 + 2 \cdot 70 = 190$ , voor B:  $10 + 2 \cdot 90 = 190$ , zodat beide dan ook gelijk in utiliteit zijn.

"It is important that judges understand this concept as they perform both steps 6 and 8."

Voor dat betere begrip moeten we dan wel wat uitgebreider voorbeelden en uitlegmateriaal construeren dan Gardiner en Edwards hier doen.

9. Bereken voor iedere mogelijke uitkomst de utiliteit, zoals in het juist n.a.v. stap 8 gegeven voorbeeld.

10. "Decide". Kies de actie die tot de grootste utiliteit leidt.

Opmerking: meestal zullen we vóór stap 10 de verre van eenvoudige analyse moeten plegen van de diverse 'states of nature' die zich kunnen voordoen met de bijbehorende waarschijnlijkheden.

Aanvulling bij stap 4: (uit Edwards & Guttentag 1975)  
"Goals, for this purpose, should not be stated as target numbers (e.g., reduction of drinking driver

accidents by 50%), but rather as dimensions (reduction of drinking driver accidents),

aanvulling bij 8b: (uit Edwards & Guttentag 1975)  
"A linear transformation is almost always adequate for this purpose: errors produced by linear approximations to monotonic non linear functions are likely to be unimportant relative to test-retest unreliability, inter respondent differences, and the like."

beslissingen 5 2-3-76

Onafhankelijkheidseigenschappen.

De aanname bij een en ander is wel dat de dimensies onafhankelijk van elkaar zijn.

Value independence is a strong assumption, not easily met and not satisfied. Fortunately, in the presence of even modest amounts of measurement error, quite substantial amounts of deviation from value independence will make little difference to the ultimate number of utilities and even less to the rank ordering of the utility values (utiliteiten van acties, of zo men wil, uitkomsten). (For recent discussions of the robustness of linear models, on which this assertion depends, see Dawes & Corrigan, Psychol. Bull. 1974) A frequently satisfied condition that makes the assumption of value independence very unlikely to cause trouble is conditional monotonicity; that is, the additive approximation will work well if, for each dimension, either more is preferable to less, or less

is preferable to more throughout the range of the dimension that is involved in the evaluation, for all available values of the other dimensions. When the assumption of value independence is unacceptable even as approximation, much more complicated models and elicitation procedures that take value independence into account are available." Het vervelende is alleen dat de auteurs hiervoor geen literatuurverwijzing geven.

'Value independence' betekent dat de plaats die een actie of uitkomst is toegekend op een dimensie niet beïnvloed wordt door de plaats die het op andere dimensies is toegekend of toegekend zal worden, of beter: het verschil in utiliteit tussen twee punten op een bepaalde dimensie wordt er niet door beïnvloed op welke plaatsen op andere dimensies een bepaalde actie of uitkomst beoordeeld wordt.

De auteurs besteden er verder niet zoveel aandacht meer aan, maar een meer gedetailleerde uitwerking en een groter aantal voorbeelden zijn ongetwijfeld dxiiáx nodig om duidelijk te maken om welk soort samenhang het hier gaat, en zeker om deze onafhankelijkheid af te grenzen van een ander soort onafhankelijkheid: 'environmental independence'.

"A trickier issue than value independence is what might be called environmental independence. The traffic congestion caused by a coastal development is extremely likely to be positively correlated with the number of people served by the development. Yet these two

dimensions may be value independent; the correlation simply means that programs with both little traffic congestion and many people served are unlikely to be presented for evaluation. Violations of environmental independence can lead to double counting. If two value dimensions are perfectly environmentally correlated, only one need be included in the evaluation process. If both are included, care must be taken, to insure that the aggregate importance weight given to both together properly captures their joint importance. For example, if number of people served and traffic congestion were perfectly environmentally correlated and measured on the same scale after rescaling, if they had equal weights, and if one entered with positive and the other with negative sign into the aggregation, the implication would be that they exactly neutralized each other, so that any feasible combination of these two variables would be equivalent in value to any other feasible combination.

beslissingen 6 2-3-76.

The decision maker is unlikely to feel that way but may have trouble adjusting his importance weights to reflect his true feelings. His life could be simplified by redefining the two dimensions into one, for example, number of people served, taking into consideration all that it entails with respect to traffic.

The problem is trickier if the environmental correlation is high but not perfect. But the solution remains the same: Try, whenever possible, to define or

redefine value dimensions so as to keep environmental correlations among them low. When that cannot be done, check on the implications of importance weights and location measures assigned to environmentally correlated dimensions to make sure that their aggregate weight properly reflects their aggregate importance."

Dan volgt nog een bespreking van het verschijnsel dat dergelijke environmental correlations in het algemeen inderdaad aanwezig zijn, en bovendien negatief zijn. Dat zou in veel gevallen geweten kunnen worden aan vóórselectie waarbij onwaarschijnlijke (of gedomineerde) acties, voorstellen of plannen geëlimineerd zijn. Het verschijnsel is dan analoog aan wat zich voordoet bij selectie van studenten op meerdere variabelen (alle variabelen positief gescoord, dat wil zeggen hoe hoger score op de variabele  $x$ , hoe beter studieprognose): ná een dergelijke selectie moet je niet verbaasd zijn in de geselecteerde groep negatieve correlaties tussen die selectie variabelen te vinden. De laatste opmerking is van mij, en biedt misschien een aanknopingspunt om het verschijnsel van environmental correlation te verhelderen. (Vergelijk Science artikel van Dawes, 1973 geloof ik). Raiffa (1968) maakt ook nog wat onduidelijke opmerkingen over deze problematiek. Onduidelijk, omdat hij niet lijkt te onderscheiden tussen value independence en environmental independence, tussen outcomes en utiliteit van de outcomes op de dimensies. Maar goed, hij stelt voor om bij afhankelijkheid van dimensies te zoeken naar 'verklarende dimensies', die dan wél weer

onderling onafhankelijk zouden zijn. Zonder echter daarover meer informatie direct bij de hand te hebben (en ook hij verwijst niet naar relevante plaatsen in de literatuur) kom ik er direct niet veel verder mee. Het zo maar gaan doen van factor analyses op dergelijk materiaal zou toch wel een beetje al te grijs zijn.