Decisions involving multiple objectives: alternatives to SMART 4

Introduction

Although SMART is a relatively simple method for supporting decision-makers who are faced with problems involving multiple objectives, some of the judgments required by the method can still be quite demanding. In this chapter we consider a number of alternative methods that are designed to make judgments about multi-objective decisions easier. As we shall see, SMARTER simplifies the decision process by using linear value functions and an approximation method to estimate the decision-maker's swing weights. The Even Swaps approach avoids the need to estimate scores or weights altogether. Finally, the analytic hierarchy process (AHP) and MACBETH allow decision-makers to express their preferences using words rather than numbers.

SMARTER

The assessment of value functions and swing weights in SMART can sometimes be a difficult task, and decision-makers may not always be confident about the numbers that they are providing for the decision model. As a result, the model may not accurately reflect the decision-maker's true preferences. Because of this, Edwards and Barron[1] have argued for 'the strategy of heroic approximation'. Underlying this strategy is the idea that, while a very simple decision-making model may only approximate the real decision problem, it is less likely to involve errors in the values elicited from a decision-maker because of the simpler judgments it involves. Consistent with this strategy, Edwards and Barron have suggested a simplified form of SMART that they call SMARTER (SMART Exploiting Ranks).

SMARTER differs from SMART in two ways. First, value functions are normally assumed to be linear. Thus, the assessment of a value function for office floor area over the range 400–1500 ft^2, for example, would involve giving 400 ft^2 a value of 0 and 1500 ft^2 a value of 100, as before, and then simply drawing a straight line, rather than a curve, between these two points on a diagram like Figure 3.3 in Chapter 3. Clearly, this approximation becomes more inaccurate as the curvature of the 'true' value function increases, so, to guard against poor approximations, Edwards and Barron recommend that preliminary checks be made.

For example, we would ask the decision-maker to think about small increases in office floor area. Specifically, would a 100 ft^2 increase in floor area be more attractive if it fell near the bottom of the scale (e.g. 400–500 ft^2), in the middle (e.g. 1000–1100 ft^2) or near the top (e.g. 1400–1500 ft^2), or would it not matter where the increase occurred? If it does not matter, then a linear approximation can be used. Suppose, however, that the decision-maker says that the increase at the lower end of the scale is most appealing, while an increase at the top end of the scale would be least useful. We could then ask how much more desirable the improvement at the bottom is compared with the improvement at the top. As a rule of thumb, if the ratio is less than 2:1, then Edwards and Barron suggest that the linear approximation is probably safe, otherwise we should fall back on methods such as bisection (see Chapter 3) to obtain the value function.

The second difference between SMART and SMARTER relates to the elicitation o the swing weights. Recall that in the office location problem in Chapter 3 the decision maker was asked to compare and evaluate swings from the worst to the best position on the different attributes. For example, a swing from the worst to the best position fo 'office visibility' was considered to be 80% as important as a swing from the worst t the best position for 'closeness to customers'. In SMARTER we still have to compar swings, but the process is made easier by simply asking the decision-maker to *ran* the swings in order of importance. This avoids the need to estimate a number t represent their relative importance. SMARTER then uses what are known as 'ran order centroid', or ROC, weights to convert these rankings into a set of approximat weights.

While a set of equations, or tables, is needed to obtain the ROC weights, the bas idea is easy to understand. Suppose that the office location decision had involved ju two attributes – 'closeness to customers' and 'visibility' – and that the decision-mak had considered the swing in 'closeness to customers' to be more important than th swing in 'visibility'. We know that, after normalization, the two weights will sum t 100. As the swing in 'closeness to customers' is more important, its normalized weig must fall between just over 50 and almost 100. This suggests an approximate weig of 75, and this is indeed what the ROC equations would give us. Clearly, the RO weight for 'visibility' would be 25.

Table 4.1 shows the ROC weights for decision problems involving up to seve attributes (see Edwards and Barron[1] for more details). In the 'original' office locatio problem, the decision-maker would have simply ranked the importance of the swin for the six attributes, as shown below; this would have yielded the ROC weigh

pter 3. Clearly, this
ue' value function
arron recommend

small increases in
be more attractive
dle (e.g. 1000–1100
where the increase
be used. Suppose,
r end of the scale is
be least useful. We
ottom is compared
s less than 2:1, then
ably safe, otherwise
to obtain the value

to the elicitation of
apter 3 the decision-
to the best positions
the best position for
g from the worst to
till have to compare
cision-maker to *rank*
timate a number to
are known as 'rank
a set of approximate

C weights, the basic
on had involved just
t the decision-maker
e important than the
weights will sum to
ts normalized weight
approximate weight
us. Clearly, the ROC

volving up to seven
riginal' office location
portance of the swings
led the ROC weights

Table 4.1 – Rank order centroid (ROC) weights

Rank	Number of attributes 2	3	4	5	6	7
1	75.0	61.1	52.1	45.7	40.8	37.0
2	25.0	27.8	27.1	25.7	24.2	22.8
3		11.1	14.6	15.7	15.8	15.6
4			6.3	9.0	10.3	10.9
5				4.0	6.1	7.3
6					2.8	4.4
7						2.0

indicated, and these could then have been used to obtain the aggregate benefits of the offices in the normal way (the original normalized SMART weights are also shown below for comparison):

Rank of swing	Attribute	ROC weight	SMART weight
1	Closeness to customers	40.8	32.0
2	Visibility	24.2	26.0
3	Image	15.8	23.0
4	Size	10.3	10.0
5	Comfort	6.1	6.0
6	Car parking	2.8	3.0
		100.0	100.0

How good are the ROC weights as approximations to the weights that might have been obtained in SMART? Edwards and Barron report the results of extensive simulations suggesting that SMART and SMARTER will agree on which option has the highest aggregate benefits in 75–87% of cases. Even when they did not agree, the options identified as having the highest aggregate benefits tended to have very similar scores, suggesting that an option that was 'not too bad' was being picked by SMARTER.

All of this suggests that SMARTER is a technique that is well worth employing. However, we should note some reservations about the method. First, in problems where it has been necessary to separate costs from benefits, you might obtain a different efficient frontier if you use SMARTER rather than SMART. This means we should be very careful before we exclude dominated options from further consideration. In particular, if you were to employ the method suggested by Edwards and Newman for selecting an option from the efficient frontier, then SMART and SMARTER may well suggest that different options should be chosen. This is because the assessment of the worth of a value point to the decision-maker is based on the normalized weights, and differences between the SMART and ROC weights can lead to large discrepancies in this assessment.

These discrepancies become less important if we recall that the main purpose of a decision analysis model is not to tell us what to do in a mechanistic fashion but to yield insights and understanding about the problem in hand. However, this raises another concern about SMARTER, which Edwards and Barron acknowledge. By simplifying the decision-maker's judgmental task, we may be encouraging only a superficial consideration of the problem and hence precluding the very insights that we hope to obtain. Analysts sometimes find that these insights only emerge when the decision-maker is forced to grapple with more demanding judgmental tasks that require deeper thinking about the issues.

Finally, the ROC weights themselves raise a number of concerns. The method through which they are derived involves some sophisticated mathematics, which means that they will lack transparency to most decision-makers. To be told that you implied weight for an attribute is 15.8, without understanding why this is the case, i likely to reduce your sense of ownership of the model that is purporting to represen your decision problem. This may reduce the model's credibility. Furthermore, Belto and Stewart[2] point out that the ratio of the ROC weights between the most and leas important attributes is generally very high. For example, in a seven-attribute prob lem, this ratio is 37/2 = 18.5 (see Table 4.1). This makes the relative importance of th lowest-ranked attribute so low that, in practice, it would probably be discarded fro the analysis.

Both of these problems can be mitigated to some extent by using an alternati weight-approximation method. Several methods exist, but Roberts and Goodwi have recommended the much simpler rank-sum method for problems that invol more than two or three attributes. Rank-sum weights are easily calculated and hen are more transparent. Suppose that three attributes have been ranked. The sum of the ranks will be 1 + 2 + 3 = 6. The least important attribute is therefore assigned a ra of 1/6, the second-ranked attribute a rank of 2/6 and the highest-ranked attribute rank of 3/6 (i.e. the weights are 0.167, 0.333 and 0.5). For four attributes, the weigh will be 0.1, 0.2, 0.3 and 0.4, and so on.

Even Swaps

As we saw in Chapter 3, trade-offs are one of the most difficult judgments to ma when faced with decisions involving multiple objectives. When choosing a holid how many extra hours are you prepared to fly in order to sunbathe on a beach wh the climate is typically 5 degrees warmer? When choosing a job, will the $8000 extra salary offered by a position 1000 miles away compensate you for the inconvenience of moving house and the end of the social life you are currently enjoying? As we have seen, SMART and SMARTER use swing weights to represent these trade-offs. Hammond *et al.*[4] have proposed a radically different approach, which they call Even Swaps. In this approach, decision-makers are asked to consider directly how much gain in one attribute they would need to receive in order to compensate them for a loss in another attribute – a so-called swap. As we demonstrate below, Even Swaps uses these swaps progressively to reduce the size of the decision problem until only one option remains.

To illustrate Even Swaps, consider the following decision faced by a manufacturer who has to choose a components supplier from abroad. The choice will be based on four objectives. The manufacturer wants to minimize the annual purchase costs of the components, minimize the average delivery time, minimize the average percentage of defective components in each delivery and receive the best after-sales service from the supplier. The table below shows how the suppliers perform against these objectives (in Even Swaps, this is known as a consequences table):

Supplier location	Annual purchase cost ($)	Average delivery time (days)	Average % defective	After-sales service
Canada	140 000	3	2	Med
Mexico	80 000	4	4	Good
Japan	190 000	5	6	Med
S. Korea	90 000	5	3	Good
China	70 000	8	5	Med
India	75 000	7	4	Poor

To apply Even Swaps we proceed as follows:

(1) *Identify any options that can be eliminated because they are dominated.* If one option performs better than another on all of the attributes, then it is said to dominate the other option. Dominance can also occur if one option performs better than another option on some of the attributes and performs just as well on the remaining attributes. We can usually spot dominance more easily if the options are ranked from best to worst on each attribute, as shown below (1 = best, 6 = worst):

Supplier location	Annual purchase cost ($)	Average delivery time (days)	Average % defective	After-sales service
Canada	5	1	1	3
Mexico	3	2	3	1
Japan	6	3	6	3
S. Korea	4	3	2	1
China	1	6	5	3
India	2	5	3	6

If we study the ranks carefully, we can see that the Japanese supplier perform worse than the Canadian supplier on all of the attributes except after-sales servic where they tie. This means that the Japanese supplier can be eliminated from th decision.

(2) *Identify any options that can be eliminated because they are practically dominated.* If w compare the South Korean supplier with the Mexican supplier, we see that th Mexican supplier is either better than or at least as good as the South Korean c all attributes except average percentage defective. However, here the Korea supplier is only slightly better than the Mexican supplier (offering an averag of 3% defectives rather than 4%). The manufacturer judges that this small adva tage does not compensate for the $10 000 extra cost of the Korean supplier ar the extra day's delivery time. The Mexican supplier is therefore said 'practicall to dominate the Korean supplier, and the latter can be eliminated from furth consideration. Our decision has now been reduced to the one below:

Supplier location	Annual purchase cost ($)	Average delivery time (days)	Average % defective	After-sale service
Canada	140 000	3	2	Med
Mexico	80 000	4	4	Good
China	70 000	8	5	Med
India	75 000	7	4	Poor

(3) *Perform Even Swaps so that attributes can be eliminated.* Suppose that the percent defective rate of all of the suppliers was exactly the same at, say, 5%. This attrib would then be irrelevant to the decision because, whichever supplier we chose, would end up with the same percentage defective rate. Even Swaps uses this i to simplify the problem further by eliminating attributes. Consider the Canad supplier's average percentage defective rate of 2%. Suppose that this rose to What compensation would the manufacturer require on another attribute so t

the attraction of the Canadian supplier remained unchanged? The manufacturer feels that a reduction in the annual cost of the Canadian supplier to $80 000 would be sufficient compensation for this large deterioration in quality. Thus, in our table, we can replace the Canadian supplier's performance on the attributes with an equally attractive performance, as shown below:

Supplier location	Annual purchase cost ($)	Average delivery time (days)	Average % defective	After-sales service
Canada	80 000 ~~140000~~	3	5 ~~2~~	Med

Suppose that the Mexican supplier's defective rate also went up to 5%. The manufacturer says he would require a reduction in the Mexican supplier's annual costs down to $60 000 to compensate him for this 1% increase in defectives if this supplier is to remain equally attractive to him. He would require the same reduction in costs from the Indian supplier if its defective rate also increased from 4 to 5%. The table below shows these changes:

Supplier location	Annual purchase cost ($)	Average delivery time (days)	Average % defective	After-sales service
Canada	80 000 ~~140000~~	3	5 ~~2~~	Med
Mexico	60 000 ~~80000~~	4	5 ~~4~~	Good
China	70 000	8	5	Med
India	55 000 ~~75000~~	7	5 ~~4~~	Poor

We can now see that the average percentage of defectives is identical for all of the suppliers, so this attribute can now be eliminated. We also note that the Chinese supplier is now dominated by the Mexican supplier, so we can also remove China from our decision. Our considerably simplified decision is shown below:

Supplier location	Annual purchase cost ($)	Average delivery time (days)	After-sales service
Canada	80 000	3	Med
Mexico	60 000	4	Good
India	55 000	7	Poor

We now aim to see how the table would change if the quality of the after-sales service of all the suppliers was changed to 'Good'. The manufacturer judges that, if the quality of the Canadian supplier's after-sales service were improved to this level, then he would be prepared to accept an increase in the average delivery time to 5 days. If the Indian supplier's after-sales service improved from its current 'Poor' level to 'Good', he would accept an increase in the average delivery time to 12 days. Our new table is shown below:

Supplier location	Annual purchase cost ($)	Average delivery time (days)	After-sales service
Canada	80 000	5	Good
Mexico	60 000	4	Good
India	55 000	12	Good

We see that after-sales service no longer discriminates between the suppliers, and also that Canada is now dominated by Mexico. The new version of the decision problem becomes:

Supplier location	Annual purchase cost ($)	Average delivery time (days)
Mexico	60 000	4
India	55 000	12

Finally, the manufacturer indicates that he would require a reduction in the Mexican supplier's annual cost to $30 000 if this supplier's delivery time increased to 12 days. The final table is shown below:

Supplier location	Annual purchase cost ($)	Average delivery time (days)
Mexico	30 000 ~~60000~~	12 ~~4~~
India	55 000	12

Clearly, the Mexican supplier is dominant, and hence this supplier should be chosen.

Even Swaps versus SMART

What are the advantages and disadvantages of using Even Swaps in a decision rather than SMART?

The relative strengths of Even Swaps

(1) *Avoids need to assign scores.* In Even Swaps, the decision process is applied directly to the information in the consequences table. In the above example, the decision-maker dealt directly with costs, delivery times and levels of after-sales service. This avoided the need to assign scores to represent the performance of the options on each attribute on a 0–100 scale, which is a requirement of SMART.

(2) *Avoids need to determine swing weights.* In SMART, decision-makers often have difficulties in understanding the true meaning of the swing weights. Even Swaps avoids this by asking decision-makers to make the sort of trade-offs with which they are likely to be familiar. It has been argued that, because Even Swaps deals with concrete changes in objectives, the trade-offs are easier to think about and are more understandable.[5] Similarly, the idea of dominance, which is central to Even Swaps, is also likely to be well understood by decision-makers.

(3) *Even Swaps may be closer to a natural decision process.* Research suggests that decision-makers who do not have access to decision-aiding technologies often go through a process of progressively simplifying the decision problem until they reach a decision. As we have seen, this principle is also inherent in the Even Swaps procedure. For example, we saw in Chapter 2 that, when faced with complex decisions involving many options and attributes, people often initially use a non-compensatory procedure, such as elimination by aspects, to reduce the number of options that they have to consider to a manageable level.[6] They then apply more cognitively demanding compensatory decision strategies to the remaining options. Other research shows that decision-makers actively restructure decision problems until one alternative is seen to be dominant.[7] This may involve operations such as collapsing two or more attributes into a more comprehensive one, emphasizing an attribute or adding new attributes to the problem representation that will bolster one alternative. Thus, although both SMART and Even Swaps are compensatory methods, the principle of progressive simplification used by Even Swaps has some consistency with the results of psychological studies of unaided choice. By contrast, the SMART approach, which preserves all the alternatives within a choice set, is less close to descriptions of unaided choice.

The relative limitations of Even Swaps

(1) *It is relatively hard to apply without practice.* There is some evidence that people can find Even Swaps harder to use than SMART, especially when they have not had much practice in applying it. In particular, the identification of dominance or practical dominance can be demanding and time consuming if done manually.

Also, for sizeable decision problems, the decision-maker will need to make a very large number of swaps, so that the effort and time required will be substantial.[8,9] In one study that compared Even Swaps with SMART,[10] people made a number of errors when applying the technique. For example, they became confused when bigger values on one attribute were better (e.g. a larger market share) while bigger values on another attribute were worse (e.g. higher costs) and made compensations in the wrong direction. Another study[11] found that decision-makers sometimes tried to make links between the swaps that were made and what they thought was likely to occur in practice. In our supplier example, experience might have taught the decision-maker that better after-sales service is usually associated with a lower delivery time *and* fewer defectives. He therefore might attempt to make a swap where higher costs are compensated for not only by better after-sales service but also by improvements in average delivery time and average number of defectives. Thus, the increase in costs is compensated for many times over. There is also the problem of ensuring that all of the swaps are consistent with each other, but checking consistency is not a simple process in Even Swaps. All of this points to the need for software to support applications of Even Swaps, especially for larger problems. One package, SMART SWAPS,[12] provides guidance by suggesting potential swaps that might be attractive to users based on their initial swaps. It is also designed to trap any errors that the decision-maker might make when applying the swaps.

(2) *The output of the process is less informative.* Even if decision-makers can apply Even Swaps without errors, the output of the process provides less information than SMART. In our supplier example, it only told us which supplier to choose, and we therefore have no idea of how close the other suppliers were to being the best choice. In SMART we would have a list of all the options and their associated scores, so that any option that came a close second could be identified. It is also difficult to perform sensitivity analysis in Even Swaps. To do this, we would have to return to earlier stages in the process and apply different swaps to examine their effects. However, without sensitivity analysis, we cannot tell how robust the recommended choice is to slight changes in the swaps that we made. Decision analysis models can play a valuable role in enabling managers to explore decision problems and in helping them to explain the rationale behind choosing a particular course of action. These benefits will be less easy to achieve when Even Swaps have been used to make a decision.

(3) *Use of Even Swaps may not have a neutral effect on choice.* There is some evidence that using Even Swaps may not have a neutral effect on the alternative that the decision-maker chooses at the end of the process. First, as we have seen, Even Swaps deliberately creates tables where all of the alternatives perform equally well on particular attributes (e.g. in our choice of supplier we created a table where the average percentage defectives of all the suppliers was the same). Norma

decision theory suggests that people should then simply ignore these attributes and base their choice on how the options perform on the remaining attributes. Research suggests that this might not be the case – the attributes where performance is equal can still have an effect on how the remaining attributes are perceived.[13,14] Under some conditions, these equal performances may make the alternatives look more similar. This dilutes the effect of the attributes where they perform differently. Under other conditions, these equal performances emphasize the importance of the different performances on the remaining attributes.[15–19] In Even Swaps this may have an impact on the identification of practical dominance. For example, in our supplier example, creating equal performances on some of the attributes may have the effect of either diminishing or exaggerating the importance of a three-day difference in the delivery time of two suppliers in the decision-maker's eyes. If the importance is diminished, then this may lead to a decision that practical dominance applies – a decision that results purely from the effect of the eliminated (and now supposedly irrelevant) attributes. Indeed, this raises the possibility that the final choice of an alternative may depend on the order in which the attributes are eliminated during the Even Swaps process.

Second, another study[20] found that the seemingly equivalent preference assessment procedures of *choice* (e.g. 'choose between a store's own-brand cola at 40 cents and a 55 cent Coke') and *matching* (e.g. 'imagine if a store's own-brand cola costs 40 cents: at what price would a Coke be attractive to you?') generate systematically different estimates of a consumer's price-quality trade-offs. This finding of a lack of 'procedural invariance' illustrated the prominence effect, in that people were more likely to prefer the alternative that was superior on the more important attribute in a straightforward choice than in a matching task. This suggests that the more important attribute is more salient in choice than in matching. The process of defining an even swap is essentially a matching task, and thus the use of this process may affect decision-making.

The analytic hierarchy process

The analytic hierarchy process (AHP), which was developed by Thomas Saaty when he was acting as an adviser to the US government, has been very widely applied to decision problems in areas such as economics and planning, energy policy, material handling and purchasing, project selection, microcomputer selection, budget allocations and forecasting.[21] Saaty developed a user-friendly computer package, called EXPERT CHOICE,[22] to support the method. Other software that supports the AHP includes HIPRE 3+ (HIerarchical PREference analysis).[23]

We will use the following problem to demonstrate the application of the AHP. A manager in a food processing company has to choose a new packaging machine to

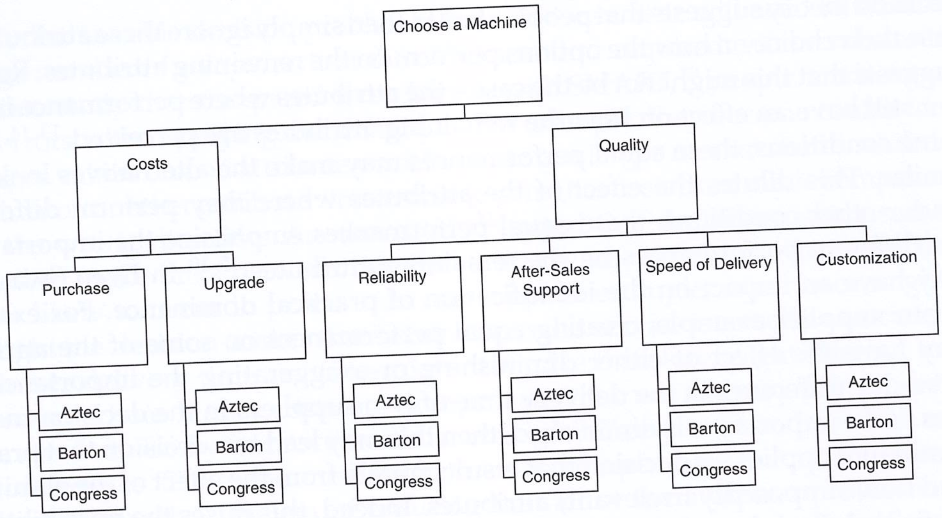

Figure 4.1 – A hierarchy for the packaging machine problem

replace the existing one which is wearing out. The manager has a limited budget for the purchase and has narrowed down the possible options to three: (i) the Aztec, (ii) the Barton and (iii) the Congress. However, the decision is still proving to be difficult because of the variety of attributes associated with the machines, such as the purchase price, reputation for reliability and the quality of after-sales support provided by the different manufacturers.

To apply the AHP, we proceed as follows:

(1) *Set up the decision hierarchy.* This is similar to a value tree in SMART, but the main difference is that the alternative courses of action also appear on the hierarchy at its lowest level. Figure 4.1 shows the decision hierarchy for the packaging machine problem. At the top of the tree is a statement of the general objective of the decision, in our case 'Choose a Machine'. The 'general' attributes associated with the decision problem ('Costs' and 'Quality') are then set out below this. As shown, these attributes can be broken down into more detail at the next level. For example, within 'Quality' the manager wishes to consider the attributes 'Reliability', 'After-Sales Support', 'Speed of Delivery' and 'Customization' (this is the extent to which the manufacturer is able to adapt the machine for the specific requirements of the food company). If necessary, this process of breaking down attributes continues until all the essential criteria for making the decision have been specified. Finally, the alternative courses of action are added to the hierarchy, below each of the lowest level attributes.

(2) *Make pairwise comparisons of attributes and alternatives.* This is used to determine the relative importance of attributes, and also to compare how well the options perform on the different attributes. For example, how much more important is the initial purchase price than the cost of upgrading the machine at a later date? Is the Aztec strongly preferred to the Barton for the quality of after-sales support?

Following each 'split' in the hierarchy, the importance of each attribute is compared, in turn, with every other attribute immediately below that 'split'. Thus, the importance of 'Costs' and the importance of 'Quality' are first compared. Then the four 'Quality' attributes are compared with each other for importance, and so on. Note that the comparisons are pairwise, so that, if there are four attributes, A, B, C and D, we need to make six comparisons: A with B, A with C, A with D, B with C, B with D and, finally, C with D.

Saaty recommends that these pairwise comparisons be carried out using verbal responses. For example, the manager is asked to consider whether 'Costs' and 'Quality' are of equal importance or whether one is more important than the other. The manager indicates that 'Costs' are more important, so he is then asked if costs are:

weakly more important?	(3)
strongly more important?	(5)
very strongly more important?	(7)
extremely more important?	(9)

The method then converts the response to the number shown in brackets. For example, if 'Costs' are 'strongly more important' than 'Quality', then they are assumed to be 5 times more important. Note that intermediate responses are allowed if the decision-maker prefers these (e.g. 'between weakly and strongly more important, which would be converted to a '4'). Also, if decision-makers prefer not to use verbal responses, then they can either make direct numerical inputs, on a scale from 1 ('equally important') to 9, or they can use a graphical facility in EXPERT CHOICE to make these inputs.

Each set of comparisons can be represented in a table (or matrix). From the 'Costs' versus 'Quality' comparison, we obtain Table 4.2.

Table 4.2 – Comparing the importance of 'Costs' and 'Quality'

	Costs	Quality
Costs	1	5
Quality		1

Table 4.3 – Comparing the importance of the 'Quality' attributes

	Reliability	After-Sales Support	Speed of Delivery	Customization
Reliability	1	4	5	4
After-Sales Support		1	3	1/2
Speed of Delivery			1	1/3
Customization				1

Similarly, for the four 'Quality' attributes, the manager's judgments lead the values in Table 4.3. The numbers in the tables represent how much mo important the 'row' attribute is compared with the 'column' attribute. For examp 'Reliability' is four times more important than 'After-Sales Support'. Fractio values therefore indicate that the 'column' attribute is most important. example, 'Speed of Delivery' is only 1/3 as important as 'Customization'. N that only 1s appear on the diagonal of the tables, as each attribute must have eq importance with itself. A similar table is obtained from the manager's compari of the importance of 'Purchase' and 'Upgrade' costs.

Finally, the same process is used to compare the manager's relative *prefere* for the machines with respect to each of the lower-level attributes. For example will be asked to consider the purchase costs of the machines and asked whethe *terms of purchase costs*, the Aztec and Barton are 'equally preferred'. If he indic that the Barton is preferred, he will then be asked whether it is 'weakly prefer 'strongly preferred' or 'extremely strongly preferred' (with intermediate respo allowed). This leads to the values in Table 4.4, which shows, for example, tha Aztec is twice as preferable as the Congress on purchase cost.

This process is repeated, yielding a table for each of the lowest-level attri to represent the manager's preferences for the machines in terms of that attri

(3) *Transform the comparisons into weights and check the consistency of the decision-m comparisons.* After each table has been obtained, the AHP converts it into of weights, which are then automatically normalized to sum to 1. A nu

Table 4.4 – Comparing the machines on 'Purchase Cost'

	Aztec	Barton	Congress
Aztec	1	1/3	2
Barton		1	6
Congress			1

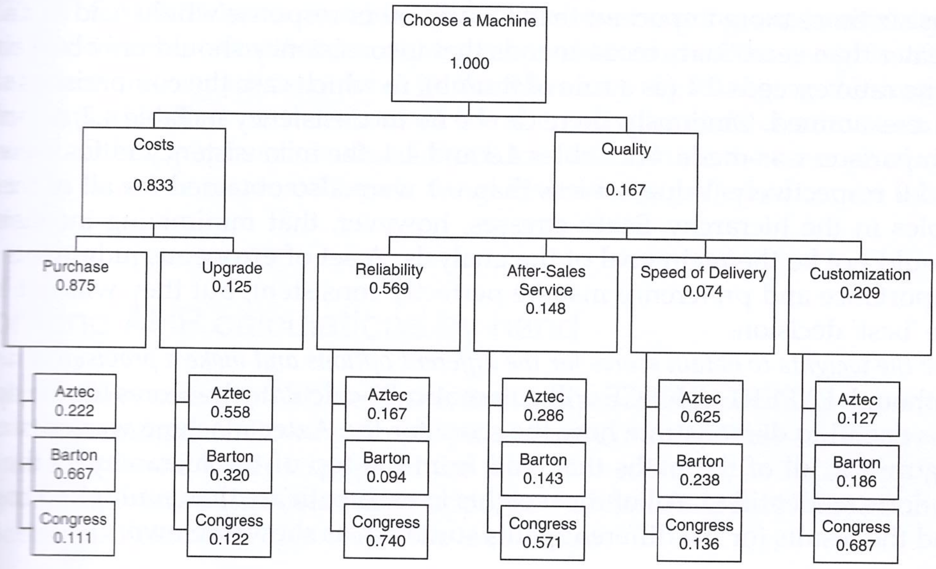

Figure 4.2 – Weights for the packaging machine problem

of conversion methods are possible. Saaty recommends a mathematical approach based on eigenvalues (see Saaty[24] for details of this method). Because this involves a relatively complex mathematical procedure, software such as EXPERT CHOICE is usually needed to perform the calculations. However, later on we will show a simple method for approximating the weights used in the AHP.

Figure 4.2 shows the weights obtained from all the tables in the hierarchy using EXPERT CHOICE. For Table 4.2, where 'Costs' were considered to be 5 times more important than 'Quality', the derivation of the weights is clear (a 5:1 ratio yields weights of 5/6 and 1/6, i.e. 0.833 and 0.167). The derivation is less transparent for the larger tables. For example, for Table 4.3 the weights are 'Reliability' 0.569, 'After-Sales Service' 0.148, 'Speed of Delivery' 0.074 and 'Customization' 0.209, suggesting that the decision-maker considered 'Reliability' to be by far the most important of the 'Quality' attributes.

Along with the weights, the AHP also yields an inconsistency ratio. This is produced automatically by EXPERT CHOICE (see Saaty[24] for details of the method of calculation), but later on we will show a way of getting a good approximation to this ratio by hand calculation. The ratio is designed to alert the decision-maker to any inconsistencies in the comparisons that have been made, with a value of zero indicating perfect consistency. For example, suppose a decision-maker's responses imply that attribute A is twice as important as B, while B is judged to be three times as important as C. To be perfectly consistent, the decision-maker should judge that

A is six times more important than C. Any other response would lead to an index greater than zero. Saaty recommends that inconsistency should only be a concern if the ratio exceeds 0.1 (as a rule of thumb), in which case the comparisons should be re-examined. Obviously, there can be no inconsistency in Table 4.2, as only one comparison was made. For Tables 4.3 and 4.4, the inconsistency ratios were 0.059 and 0 respectively. Values of less than 0.1 were also obtained for all of the other tables in the hierarchy. Saaty stresses, however, that minimizing inconsistency should not be the main goal of the analysis. A set of erroneous judgments about importance and preference may be perfectly consistent, but they will not lead to the 'best' decision.

(4) *Use the weights to obtain scores for the different options and make a provisional decision.* Although EXPERT CHOICE will automatically calculate the scores for the options, it is useful to demonstrate how the score for the Aztec machine was obtained. In Figure 4.2, all of the paths that lead from the top of the hierarchy to the Aztec option are identified. All of the weights in each path are then multiplied together, and the results for the different paths summed, as shown below:

$$\begin{aligned}\text{Score for Aztec} = {} & 0.833 \times 0.875 \times 0.222 \\ & + 0.833 \times 0.125 \times 0.558 \\ & + 0.167 \times 0.569 \times 0.167 \\ & + 0.167 \times 0.148 \times 0.286 \\ & + 0.167 \times 0.074 \times 0.625 \\ & + 0.167 \times 0.209 \times 0.127 = 0.255\end{aligned}$$

Note that the Aztec scores well on attributes that are considered to be relatively unimportant, such as 'Upgrade Costs' (which carries only 0.125 of the 0.833 weight allocated to costs) and 'Speed of Delivery' (which carries only 0.074 of the weight allocated to 'Quality', which itself is relatively unimportant). It scores less well on the more important attributes, so its overall score is relatively low. The scores for all three machines are shown below:

Aztec	0.255
Barton	0.541
Congress	0.204

This clearly suggests that the Barton should be purchased.

(5) *Perform sensitivity analysis.* As in any decision model, it is important to examine how sensitive the preferred course of action is to changes in the judgments made by the decision-maker. Many of these judgments will be 'rough and ready', and the decision-maker may be unsure about exactly what judgments to input

EXPERT CHOICE has a number of facilities for carrying out sensitivity analysis. In dynamic sensitivity analysis, a bar chart shows the weights attached to attributes at a particular level in the hierarchy. By changing the lengths of these bars, the effect on the scores of the alternative courses of action can be examined. Other graphs allow decision-makers to examine the amount of change that can be made to an attribute's weight before the preferred course of action changes.

Performing AHP calculations by hand

If you do not have access to AHP software, then it is possible to obtain approximations of the weights using the following simple procedure. Consider Table 4.3. We first enter the numbers into the lower triangle of the table. For example, as Reliability is four times more important than After-Sales Support, After-Sales Support must be only 1/4 as important as Reliability. This yields the table below:

	Reliability	After-Sales Support	Speed of Delivery	Customization
Reliability	1	4	5	4
After-Sales Support	1/4	1	3	1/2
Speed of Delivery	1/5	1/3	1	1/3
Customization	1/4	2	3	1

Next, we sum the columns of the table and then divide each number in the table by the total of its column. For example, the total of the Reliability column is 1.7. This means that the four values in the Reliability column become 0.588, 0.147, 0.118 and 0.147. The table below shows all the results. Finally, we average the numbers in each row. These averages, which are also shown in the table, can now be used as approximate weights for the four attributes. Similar calculations can be applied to the other tables in the hierarchy.

	Reliability	After-Sales Support	Speed of Delivery	Customization	Average of row
Reliability	0.588	0.545	0.417	0.686	**0.559**
After-Sales Support	0.147	0.136	0.250	0.086	**0.155**
Speed of Delivery	0.118	0.045	0.083	0.057	**0.076**
Customization	0.147	0.273	0.250	0.171	**0.210**

It is also possible to calculate an approximation to the inconsistency ratio by using the following procedure. This may seem involved, but it is easily implemented on a spreadsheet. We will demonstrate the process on Table 4.3:

Step 1: Fill in the lower triangle of the table, as before. Then write the weight for each attribute (or option) at the top of each column. The results are shown below:.

	Reliability	After-Sales Support	Speed of Delivery	Customization
Weights	**0.559**	**0.155**	**0.076**	**0.210**
Reliability	1	4	5	4
After-Sales Support	1/4	1	3	1/2
Speed of Delivery	1/5	1/3	1	1/3
Customization	1/4	2	3	1

Step 2: Multiply the weight at the top of each column by each of the numbers in that column. Then sum each row of the resulting table:

	Reliability	After-Sales Support	Speed of Delivery	Customization	Sums
Reliability	0.559	0.620	0.380	0.840	**2.399**
After-Sales Support	0.140	0.155	0.228	0.105	**0.628**
Speed of Delivery	0.112	0.052	0.076	0.070	**0.309**
Customization	0.140	0.310	0.228	0.210	**0.888**

Step 3: Divide each of these sums by the weight for that attribute (or option). Then average the resulting ratios.

	Sums	Weight	Ratio
Reliability	2.399	0.559	**4.291**
After-Sales Support	0.628	0.155	**4.056**
Speed of Delivery	0.309	0.076	**4.078**
Customization	0.888	0.210	**4.221**
		Average ratio	**4.161**

Step 4: An inconsistency index can be calculated using the following formula:

$$\text{Inconsistency index} = \frac{\text{average ratio from step 3} - n}{n - 1}$$

where n is the number of rows in the table we are investigating. In our case, this is 4, so we have

$$\text{Inconsistency index} = \frac{4.161 - 4}{4 - 1} = 0.054$$

Note that, if our table had been perfectly consistent, the average ratio from step 3 would have been 4.0, so our inconsistency index would have had a value of zero.

Step 5: Divide the inconsistency index by the appropriate value from Table 4.5 to obtain the inconsistency ratio. The values in this table were generated by Saaty to estimate the inconsistency indices for random tables. Our inconsistency ratio is therefore $0.054/0.90 = 0.06$. As this is below 0.1, we should have no concerns about inconsistency in this table. Note that this is very close to the 0.059 value produced by EXPERT CHOICE.

Table 4.5 – Random indices for checking the consistency of a table

	n								
	2	3	4	5	6	7	8	9	10
Random index	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

The axioms of the AHP

The AHP is based on four axioms:[25]

(1) The *reciprocal axiom* states that, if A and B are options or attributes in the decision hierarchy and A is n times more preferable (or more important or more likely) than B, then B must be $1/n$th as preferable (or important or likely) as A. For example, if Reliability is four times more important that After-Sales Support, then After-Sales Support must be only 1/4 as important as Reliability.

(2) The *homogeneity axiom* states that the elements being compared should not differ by extreme amounts on a criterion. For example, this axiom would be violated if

A were 24 times more important than B. This axiom is reflected in the range of the AHP verbal scale, which runs from 1/9 to 9. As we discuss below, this axiom can be relaxed if this is judged to be absolutely necessary.

(3) The *synthesis axiom* states that judgments about the importance of elements in a hierarchy do not depend on the elements below them. For example, in our hierarchy, judgment about the relative importance of Reliability and After-Sales Support does not depend on the packaging machines that are available. Thus, the relative importance would be the same even if a different set of machines were on offer. This axiom may be violated in many practical applications. For example, suppose that we state that Reliability is four times more important than After-Sales Support and then discover that all of the available machines have extremely high and similar levels of Reliability that far exceed the minimum acceptable level. However, they differ to a considerable extent in the quality of After-Sales Support offered. In this case we may wish to change our mind and judge After-Sales Support as being more important in our choice between the machines. To guard against this danger, it is recommended that a 'bottom-up' approach be applied when evaluating the elements in an AHP hierarchy (i.e. we should start with the alternative courses of action and work upwards). By comparing the machines' performances on Reliability and After-Sales Support first, we would learn about their similarities in reliability, and this would inform our judgment when we came to compare the importance of these two attributes. Alternatively, the analytical network process (ANP)[26] provides a formal approach to this problem, but at the cost of greater mathematical complexity.

(4) The *expectation* axiom states that decision-makers should make sure that their ideas are adequately represented in the decision model. This is similar to the concept of requisite decision modeling in SMART. If the decision-maker's intuitively preferred option differs from the best option suggested by the model, then this indicates that the model should be investigated to identify the reason for the discrepancy. Perhaps the hierarchy is incomplete or the relative importance of attributes is not independent of the options (see the synthesis axiom above). Alternatively, the investigation might reveal that the decision-maker's intuition is at fault because he or she is unable to comprehend a complex decision problem in its entirety.

The AHP versus SMART

It can be seen that the AHP is fundamentally different to SMART in many respects. We next consider the relative strengths of the AHP and then consider the main criticisms that have been made of the technique.

The relative strengths of the AHP

(1) *Simplicity of pairwise comparisons*. The use of pairwise comparisons means that the decision-maker can focus, in turn, on each small part of the problem. Only two attributes or options have to be considered at any one time, so that the decision-maker's judgmental task is simplified. Verbal comparisons are also likely to be preferred by decision-makers who have difficulty in expressing their judgments numerically.

(2) *Redundancy allows consistency to be checked*. The AHP requires more comparisons to be made by the decision-maker than are needed to establish a set of weights. For example, if a decision-maker indicates that attribute A is twice as important as B, and B, in turn, is four times as important as C, then it can be inferred that A is eight times more important than C. However, by also asking the decision-maker to compare A with C, it is possible to check the consistency of the judgments. It is considered to be good practice in decision analysis to obtain an input to a decision model by asking for it in several ways and then asking the decision-maker to reflect on any inconsistencies in the judgments put forward. In the AHP this is carried out automatically.

(3) *Versatility*. The wide range of applications of the AHP is evidence of its versatility. In addition to judgments about importance and preference, the AHP also allows judgments about the relative likelihood of events to be made. This has allowed it to be applied to problems involving uncertainty, and also to be used in forecasting.[27–30] AHP models have also been used to construct scenarios by taking into account the likely behavior and relative importance of key actors and their interaction with political, technological, environmental, economic and social factors (Saaty,[24] p. 130).

Criticisms of the AHP

(1) *Conversion from verbal to numeric scale*. Decision-makers using the verbal method of comparison will have their judgments automatically converted to the numeric scale, but the correspondence between the two scales is based on untested assumptions. If you indicate that A is weakly more important than B, the AHP will assume that you consider A to be 3 times more important, but this may not be the case. In particular, several authors have argued that a multiplicative factor of 5 is too high to express the notion of 'strong' preference.[30]

(2) *Problems of the 1–9 scale*. Experimental work suggests that, when one attribute or option is 'extremely more important' than another, then ratios of 1 to 3 or 1 to 5 are more appropriate than the 1 to 9 ratio assumed by the AHP.[31] However, if a

decision-maker using the verbal scale does wish to incorporate very extreme ratios into the decision model, the restriction of pairwise comparisons with a 1–9 scale is bound to create inconsistencies. For example, if A is considered to be four times more important than B, and B is four times more important than C, then, to be consistent, A should be judged to be 16 times more important than C, but this is not possible with the AHP's verbal scale.

To avoid this problem, Forman and Gass[25] recommend that, when setting up the AHP hierarchy, the decision-maker should attempt to arrange the elements in clusters so that they do not differ in extreme ways (this would also ensure conformance with axiom 2). They argue that this is desirable, anyway, because judgments involving extreme differences are likely to be unreliable. However, where extreme judgments are required, one can avoid verbal judgments altogether and directly input the desired numerical ratios.

(3) *Meaningfulness of responses to questions.* Unlike SMART, weights are elicited in the AHP without reference to the scales on which attributes are measured. For example, a person using SMART to choose a house might be asked to compare the value of reducing the daily journey to work from 80 to 10 miles with the value of increasing the number of bedrooms in the house from 2 to 4. Implicit in this type of comparison is the notion of a trade-off or exchange: 70 fewer miles may be only half as valuable as two extra bedrooms. It can be shown that AHP questions, which simply ask for the relative importance of attributes without reference to their scales, imply weights that reflect the relative value of the *average* score of the options on the different criteria,[32] which is a difficult concept for decision-makers to grasp. This may mean that the questions are interpreted in different and possibly erroneous ways by decision-makers.[32,33]

(4) *New alternatives can reverse the rank of existing alternatives.* This issue, which is related to the last point, has attracted much attention.[34] Suppose that you are using the AHP to choose a location for a new sales office and the weights you obtained from the method give the following order of preference: 1 Albuquerque, 2 Boston, 3 Chicago. However, before making the decision, you discover that a site in Denver is also worth considering, so you repeat the AHP to include this new option. Even though you leave the relative importance of the attributes unchanged, the new analysis gives the following rankings: 1 Boston, 2 Albuquerque, 3 Denver, 4 Chicago, so the rank of Albuquerque and Boston has been reversed, which may not seem to be intuitively reasonable. If Albuquerque is better than Boston, then surely it is still better than Boston, irrespective of whether or not Denver is available.

These rank reversals cannot occur in SMART, but some analysts have argued that in some circumstances they are desirable.[25] Their arguments are based on what is referred to as dilution. Suppose that two people, Alan and Barbara, work in a sales office. Alan has excellent computing skills but is less good at selling than

Barbara. Barbara knows very little about computers. Alan is therefore rated as the most valued of the two employees because, if he is absent, there is no one to fix computer problems. Subsequently, a third person, Colin, joins the office. While Colin is not as knowledgeable as Alan about computers, he is still quite skilled. Now Barbara is regarded as more valuable than Alan (i.e. their ranks have been reversed). She is a better salesperson and Alan's computer knowledge is now less vital, given that he is no longer the only computer buff in the office. The value of his computing skills has been 'diluted'.

In spite of this example, some research suggests, anyway, that rank reversals occur rarely in applications of the AHP.[25] However, if the decision-maker does want to avoid any danger of such reversals, then it is now possible to choose to run the AHP in what is referred to as 'ideal mode' (the original mode is referred to as 'distributive mode'). The term 'ideal' refers to the fact that the weights of alternatives are assigned relative to the ideal (or the most preferred) alternative. For example, the best option on an attribute might have a weight of 0.6. If it is twice as preferable as the second best option, then this second option will have a weight of 0.3, and so on. Adding further options at a later stage will not change the rankings of these weights because all the weights are compared with a fixed value (the ideal). This is similar to SMART, where the scores given to options on an attribute are also assigned relative to the best performing option, which has a score of 100.

(5) *Number of comparisons required may be large*. While the redundancy built into the AHP is an advantage, it may also require a large number of judgments from the decision-maker. Consider, for example, the office location problem in Chapter 3, which involved seven alternatives and seven attributes (if we simplify the problem to include 'Total Costs' and only lower-level benefit attributes). This would involve 168 pairwise comparisons of importance or preference. In a study by Olson *et al.*,[33] this requirement to answer a large number of questions reduced the attraction of the AHP in the eyes of potential users, even though the questions themselves were considered to be easy.

To address this problem, EXPERT CHOICE has a facility for using a 'ratings' or 'absolute' approach where each of the alternatives is rated on a single scale that the decision-maker can define. For example, we might define a scale for the computer skills of job applicants as 'Excellent', 'Good', 'Average', 'Poor' and 'Very Poor'. The decision-maker can then determine numerical values to represent the 'intensity' of the verbal descriptions. For example, 'Excellent' may be assigned a score of 1, 'Good' a score of 0.7, and so on. Having formulated the scale, each applicant can be rated directly on this scale. This avoids the need to make pairwise comparisons between all of the applicants. For example, if we had ten job applicants, we would only have to make ten direct ratings, rather than 45 pairwise judgments, for each attribute.

MACBETH

MACBETH (Measuring Attractiveness by a Categorical-Based Evaluation TecHnique), which was developed by Carlos Bana e Costa and Jean-Claude Vansnick,[35,36] is similar to the AHP in a number of ways. First, users are asked to compare only pairs of options or attributes at a time, and second, they express their preferences in terms of words rather than numbers. Like the AHP, this results in a table of comparisons that allows the method to inform decision-makers about the consistency of their pairwise judgments.

There are, however, a number of important differences between the methods. Whereas the AHP elicits a ratio for the relative importance or preference between elements of the hierarchy (e.g. Reliability is five times more important than After-Sales Service), MACBETH asks users to compare *differences* in attractiveness. For example, suppose that two packaging machines, the Aztec and the Barton, are compared for the attractiveness of the After-Sales Service they offer. The decision-makers would be asked to decide whether the difference in their attractiveness was 'Very Weak', 'Weak', 'Moderately Strong', 'Strong', 'Very Strong' or 'Extreme'. Alternatively, the decision-maker could indicate that the machines were equally attractive in their After-Sales Service. A similar process is used to obtain the swing weights for the attributes. Once the decision-maker has made these indications, MACBETH uses a mathematical algorithm to check their consistency and generate numerical scores and weights. When inconsistencies are discovered, the method indicates how they have arisen and suggests how greater consistency could be achieved.

Because MACBETH uses the additive value model (see Chapter 3), it is possible to integrate it with SMART. This integrative facility is available in at least one software product (HIVIEW 3). This is useful where decision-makers have problems in directly assigning the numerical scores required by SMART and are more comfortable in expressing their preferences in terms of words. In particular, some decision-makers may have problems in understanding the 0–100 scales used in SMART. For example, a score of zero simply indicates that an option is the worst performer on an attribute, not that it has no value. Similarly, because the zero is defined in this way, the scales used in SMART are interval scales, so an option scoring 50 on an attribute is not necessarily twice as preferable as an option scoring 25. The idea that it is the relative size of differences (or intervals) between scores that is meaningful may be difficult to convey. Macbeth addresses this by asking users to compare these differences in words.

In addition to these advantages, MACBETH includes extensive facilities for examining the robustness of decisions and their sensitivity to changes in the decision-maker's judgments. However, the mathematical algorithm underlying the method means that users have to rely on computer software to implement the technique, and this will reduce the transparency of the process through which the method produces its recommendations. Because of this, in any application of MACBETH, decision-makers should spend time reviewing the outputs of the method to ensure that they agree that their preferences are being accurately represented.

Summary

In this chapter we have reviewed some alternative methods to SMART. We saw that each of these alternatives contains attractive features such as automatic consistency checks on the decision-maker's judgments, avoidance of the need to specify weights or facilities allowing decision-makers to express preferences using words rather than numbers. However, none of the methods was clearly superior to the others on all counts. Because of this, there has been a convergence between some of the techniques in an attempt to embrace the best features of each. For example, verbal judgments can now be used alongside SMART via the MACBETH procedure. Similarly, some commercial software products now include facilities for both SMART and the AHP, allowing decision-makers to have considerable flexibility in the way they tackle different decision problems.

Exercises

Additional exercises can be found on the book's website

(1) The owner of a small business is unhappy with the service she has been receiving from her bank and has decided to move her account to a rival bank. Her decision on which bank to choose will be based not only on the estimated annual bank charges that each bank will levy but also on the following 'benefit attributes':

(a) the proximity of the local branch;
(b) whether the local branch has a small business adviser;
(c) the maximum automatic loan allowed;
(d) whether a telephone banking facility is offered.

The alternative banks are listed below, together with their estimated annual costs and the scores the business owner has allocated for each of the 'benefit attributes':

Bank	Estimated annual charge ($)	Proximity	Small business adviser	Maximum loan	Telephone facility
Central	3000	0	100	40	0
Northern	5000	100	100	80	0
Direct	2000	70	0	100	100
Royal	1000	30	0	0	100
Marks	4000	90	100	20	0

The business owner is then asked to imagine that she has her account with a hypothetical bank that had the lowest scores on all of the 'benefit attributes'. She is then asked to imagine that each attribute could be switched to its best possible value and asked to rank the attractiveness of these possible switches. Her ranks are given as:

Rank	Switch
1	Lowest maximum loan facility to highest
2	No telephone banking facility to existence of this facility
3	Non-availability of small business adviser to availability
4	Least close branch to closest branch

(a) SMARTER has been used to obtain scores to represent the aggregate benefits of the banks, and these are given below:

Bank	Aggregate score
Central	35.4
Northern	62.5
Direct	83.5
Royal	29.0
Marks	30.6

Show how the score of 35.4 for the Central Bank was determined.

(b) By taking into account the estimated annual charges of the banks, determine which banks lie on the efficient frontier. Explain the significance of the efficient frontier.

(c) SMARTER is based on the 'principle of heroic approximation'. Explain how this principle applies to your analysis of the businesswoman's problem, and discuss whether it is likely to be appropriate.

(2) Imagine that you have won a holiday for two people in a magazine competition and you have been given the choice of five holiday destinations. The names and details of these destinations are shown below:

Destination	Flying time (hours)	Typical sunshine (hours per day)	Time to walk to beach (minutes)	Size of place	No. of cultural attractions nearby	Night life
Alucia	4	8	0	Large town	3	Average
Bellonia	9	6	5	Village	1	Quiet
Canin	3	5	12	Isolated	0	Quiet
Dovrania	5	9	20	Large town	5	Lively
Estinet	2	5	2	Village	0	Average

Use the Even Swaps method to determine which destination you would choose on the basis of the information that has been supplied.

(3) A motorist is using the AHP to choose a new car from three possible models: an Arrow, a Bestmobile and a Commuter. The choice will be based on just two attributes, 'Cost' and 'Style'. The motorist considers that Cost is 'Weakly More Important' than Style.

When asked to compare the costs of the cars. The motorist makes the following statements: on cost, the Bestmobile is 'Weakly Preferred' to the Arrow, but the Arrow is 'Weakly Preferred' to the Commuter. Also, the Bestmobile is 'Extremely Preferred' to the Commuter.

On style, the Arrow is '*Very* Strongly Preferred' to the Bestmobile, but the Commuter is 'Weakly Preferred' to the Arrow. Also, the Commuter is 'Extremely Preferred' to the Bestmobile.

(a) Construct a hierarchy to represent the decision problem.
(b) Use an appropriate software package, or the approximation method, to calculate the weights for each table in the hierarchy, and hence determine which car should be purchased.
(c) Calculate the inconsistency ratios for the motorist's comparisons of the cars on (i) cost and (ii) style, and interpret your results (use either your software or the approximation method here).

(4) One of the criticisms of the AHP is that the introduction of new alternatives can change the ranking of existing alternatives. Under what circumstances, if any, is this likely to be reasonable?

(5) A manager is hoping to appoint a new assistant and decides to use the AHP to rank the applicants for the job. Then, as a check, she decides to repeat the process using SMART. She is surprised to find that the ranking of the applicants derived from the AHP differs significantly from the ranking suggested by the SMART analysis. Discuss why these differences might have arisen.

References

1. Edwards, W. and Barron, F.H. (1994) SMARTS and SMARTER: improved simple methods for multiattribute utility measurement, *Organizational Behavior and Human Decision Processes*, **60**, 306–325.
2. Belton, V. and Stewart, T.R. (2002) *Multiple Criteria Decision Analysis. An Integrated Approach*, Kluwer Academic Publishers, Norwell, MA.
3. Roberts, R. and Goodwin, P. (2002) Weight approximations in multi-attribute decision models, *Journal of Multi-Criteria Decision Analysis*, **11**, 291–203.
4. Hammond, J.S., Keeney, R.L. and Raiffa, H. (1998) Even Swaps: a rational method for making tradeoffs, *Harvard Business Review*, **76**, 137–150.
5. Kajanus, M., Ahola, J., Kurttila, M. and Pesonen, M. (2001) Application of Even Swaps for strategy selection in a rural enterprise, *Management Decision*, **39**, 394–402.
6. Payne, J. (1982) Contingent decision behaviour, *Psychological Bulletin*, **92**, 382–402.
7. Montgomery, H. (1983) Decision rules and the search for a dominance structure: towards a process model of decision making, in *Analyzing and Aiding Decision Processes*, ed. by Humphreys P.C. *et al.*, North Holland, Amsterdam, The Netherlands, pp. 343–369.
8. Mustajoki, J. and Hämäläinen, R.P. (2005) A preference programming approach to make the Even Swaps method even easier, *Decision Analysis*, **2**, 110–123.
9. Bouyssou, D. and Pirlot, M. (2005) Conjoint measurement tools for MCDM, in *Multiple Criteria Decision Analysis: State of the Art Surveys*, ed. by Figueira, J., Greco, S. and Erhgott, M., Springer, New York, NY, pp. 73–132.
10. Belton, V., Wright, G. and Montibeller, G. (2008) When is swapping better than weighting? An evaluation of the Even Swaps method in comparison with multi attribute value analysis, Working Paper, Department of Management Science, University of Stratchclyde.
11. Luo, C.-M. and Cheng, B.-W. (2006) Applying Even Swaps method to structurally enhance intuitive decision making, *Systemic Practice and Action Research*, **19**, 45–59.
12. Hämäläinen, R. (2003) Decisionarium – aiding decisions, negotiating and collecting opinions on the web, *Journal of Multi-Criteria Decision Analysis*, **12**, 101–110.
13. Dhar, R. and Sherman, S.J. (1996) The effect of common and unique features in consumer choice, *Journal of Consumer Research*, **23**(3), 192–203.
14. Chernev, A. (1997) The effect of common features on brand choice: moderating role of attribute importance, *Journal of Consumer Research*, **23**(4), 304–311.
15. Tversky, A. (1977) Features of similarity, *Psychological Review*, **84**, 327–352.
16. Nisbett, R.E., Zukier, H. and Lemley, R.E. (1981) The dilution effect: nondiagnostic information weakens the implications of diagnostic information, *Cognitive Psychology*, **13**, 248–277.
17. Mellers, B.A. and Biangini, K. (1994) Similarity and choice, *Psychological Review*, **101**, 505–518.
18. Meyer, R.J. and Eagle, T.C. (1982) Context-induced parameter instability in a disaggregate stochastic model of store choice, *Journal of Marketing Research*, **19**, 62–71.
19. Tversky, A. and Russo, J. (1969) Substitutability and similarity in binary choice, *Journal of Mathematical Psychology*, **6**, 1–12.

20. Carmon, Z. and Simonson, I. (1998) Price-quality tradeoffs in choice versus matching: new insights into the prominence effect, *Journal of Consumer Psychology*, **7**, 323–343.
21. Zahedi, F. (1986) The analytic hierarchy process – a survey of the method and its applications, *Interfaces*, **16**, 96–108.
22. Expert Choice is a product of Expert Choice, Inc., 4922 Ellsworth Avenue, Pittsburgh, PA 15213, USA.
23. Mustajoki, J. and Hämäläinen, R.P. (2000) Web-HIPRE: global decision support by value tree and AHP Analysis, *INFOR*, **38**, 208–220; see also www.hipre.hut.fi
24. Saaty, T.L. (1990) *The Analytic Hierarchy Process*, RWS Publications, Pittsburgh, PA.
25. Forman, E.H. and Gass, S.L. (2001) The analytic hierarchy process – an exposition, *Operations Research*, **49**, 469–486.
26. Saaty, T.L. (1996) *Decision Making with Dependence and Feedback*, RWS Publications, Pittsburgh, PA.
27. Saaty, T.L. and Vargas, L.G. (1991) *Prediction, Projection and Forecasting*, Kluwer Academic Publishers, Norwell, MA.
28. Ulengin, F. and Ulengin, B. (1994) Forecasting foreign exchange rates: a comparative evaluation of the AHP, *Omega*, **22**, 509–519.
29. Wolfe, C. and Flores, B. (1990) Judgmental adjustments of earnings forecasts, *Journal of Forecasting*, **9**, 389–405.
30. Belton, V. and Goodwin, P. (1996) Remarks on the application of the analytic hierarchy process to judgmental forecasting, *International Journal of Forecasting*, **12**, 155–161.
31. Belton, V. and Stewart, T.J. (2002) *Multiple Criteria Decision Analysis. An Integrated Approach*, Kluwer Academic Publishers, Norwell, MA.
32. Belton, V. (1986) A comparison of the analytic hierarchy process and simple multi-attribute value function, *European Journal of Operational Research*, **26**, 7–21.
33. Olson, D.L., Moshkovich, H.M., Schellenberger, R. and Mechitov, A.I. (1995) Consistency and accuracy in decision aids: experiments with four multiattribute systems, *Decision Sciences*, **26**, 723–748.
34. Belton, V. and Gear, T. (1983) On a shortcoming in Saaty's method of analytic hierarchies, *Omega*, **11**, 228–230.
35. Bana e Costa, C.A. and Vansnick, J.-C. (1997) Applications of the MACBETH approach in the framework of the additive-aggregation model, *Journal of Multi-Criteria Decision Analysis*, **6**, 107–114.
36. Bana e Costa, C.A. and Chagas, M.P. (2004) A career choice problem: an example of how to use MACBETH to build a quantitative value model based on qualitative value judgments, *European Journal of Operational Research*, **153**, 323–331.