

CHAPTER VIII-TN 25: AN INTEGRATED APPROACH TO MULTI-DIMENSIONAL EVALUATION AND COST-EFFECTIVENESS ANALYSIS

Based on a paper by W. Acar¹

ABSTRACT

This paper shows how to integrate cost-effectiveness analysis with multi-dimensional evaluation techniques for use at the operational level in government agencies. It is concerned with a procedure that can be utilized when only one of several projects is to be the recipient of funds.

Emphasis is on making effective use of the decision maker's intuition to disaggregate decisions. The aim is to consider all the different factors that should influence a decision while utilizing a framework that helps to ensure that different factors affecting the final decision are given the "right weight". The framework for project evaluation is a collection of rules that guide one in defining scores, selecting factors for consideration, etc. by using intuition more effectively than is possible when traditional scoring procedures are used.

The method proposed offers a means through which the decision maker can control the amount of influence that intangibles have on a decision when compared to particular economically measurable factors.

An example is given to show how the integrated approach can be used.

INTRODUCTION

If people are going to make judgments, it is imperative to take all possible steps to see that insight and intuition are used effectively. Yet, given the topic of the article, it is easy for an academic to see little value in it. One might criticize the "mathematics" or the psychological conceptualization as naive. But the real question is: Does the article present an approach that allows a manager to make better decisions?

Suggestions that propose that one should do research for five years to make a decision that must be made this day or made this week are not acceptable. Suggestions of spending \$5,000 on studies to select alternative ways to spend \$10,000 would also be unacceptable. These points deserve serious consideration before this article is criticized. Actually, the best constructive critique of this paper would be another paper offering a simpler, better method of using judgmental assessments than the method described here.

Regardless, the paper's prime purpose is to propose an evaluation technique that can be used when the staff of a business or government agency have the opportunity to decide how a program's objectives should be achieved when there are several projects that can be carried out and no "perfectly rational" quantitative method is available to select a best alternative. Such a situation arises with respect to recreation land acquisition and development, partly because of the difficulty of dealing with intangibles (see TN 34 on demand for recreation and price, and Coomber and Biswas 1972 and sources cited there on intangibles).

In the PPB (Planning Programming Budgeting) system or similar styles of government or business management practiced in North America, individual projects are seen as embedded components in a hierarchy of activities. The intended outputs of larger and encompassing programmes are decided on at the upper decision-making levels. At these levels apples and oranges are compared and decisions about how many (much) to consume (buy, provide) are made on political, ideological, etc. grounds. However, decisions to select individual projects to achieve a specific objective are made at the lower operational levels. At this level often there is

¹ TN 25 is based on a paper presented to the 1973 meeting of the Canadian Association of Administrative Sciences, see Acar (). It was revised for CORDS by J. Beaman and L. Belfry.

concern is with the comparison of truly comparable alternatives and there may even be guidelines implying comparison should be on the basis of cost-efficiency of different projects, or efficiency in some broader sense. One approach to efficient management has involved the use of project scoring to select the "best" project from among several that will achieve program objectives.

GOVERNMENTAL PROJECT EVALUATION AND MULTI-DIMENSIONAL SCORING

Often, when it is necessary in project evaluation to consider so-called intangible factors, factors that cannot be evaluated in economic terms, the various intangible factors considered relevant are assigned scores. The scores assigned to the different factors are added up and the projects evaluated on the basis of some total score. This process of scoring has often been referred to as the multi-dimensional scoring of projects.

Such multi-dimensional scoring practices are in wide use in academic, managerial and military spheres. Whether the alternatives to be scored are students' papers, business projects, or prospective employees, the procedure followed is essentially the same; one conjures up the factors or "dimensions" of an evaluation, scoring is done on each dimension, and then by some process a decision is made. But if there are two projects P(1) and P(2) to be scored along two criterion dimensions A and B, there are four possible cases. They are:

- a. P(1) is better than P(2) according to both A and B.
- b. P(2) is better than P(1) according to both A and B.
- c. P(1) is better than P(2) according to A but P(2) is better than P(1) according to B.
- d. P(2) is better than P(1) according to A but P(1) is better than P(2) according to B.

In the first two cases it is readily apparent which project would be preferred globally. In the last two cases a judgment is not readily possible because one must recognize that a project is the "best" because of only one criterion while acknowledging that another project is better in terms of a different criterion dimension. Consequently, the decision maker may arbitrarily elect to treat the two criteria as equally important (even if the one has questionable validity) and make a judgment in terms of some total score for the two factors.

A CARDINALIZATION TECHNIQUE FOR IMPROVING MULTI-DIMENSIONAL SCORING

As just implied, one reason most scoring procedures use numbers that can be added, subtracted or otherwise compared is the intuitive appeal of avoiding those situations in which a decision cannot logically be made. Whether the given quantifications are valid or not is another issue. For example, consider the cardinalization of the previously mentioned scorings by means of assigning score-values to each project proposal along the two dimensions. If case "d." is under consideration, one could, for example, be considering the situation shown in Table 1.

If the criteria are equally important, the respective "values" of the two projects can be reflected in their aggregated scores which are 15 and 10. But even if these really are the scores and the aggregate scores do not fall too close to each other, it still need not be clear which project truly scores better overall. Yet this type of scoring aggregation is widely used with various degrees of care.

The author claims that in order for the aggregate score values to be meaningful, a number of conditions have to be met. These include:

- Condition 1. The criteria dimensions along which scores are taken must be "relatively" independent of each other (to avoid redundancy of dimensions);
- Condition 2. The dimensions affecting the choice to be made should be relevant to the evaluation of alternatives being considered to achieve the objective of the program.

TABLE 1: PROJECT SCORES FOR TWO PROJECTS ON TWO CRITERION DIMENSIONS

	Project 1	Number 2
Criterion A	7	8
Criterion B	8	2
Aggregate Scores	15	10

TABLE 2: DETAILS FOR CALCULATING PROJECTS SCORES

Project	1	2	3	4
A. PROJECT SCORES				
Dimension A	1	4	2	3
Dimension B	4	2	3	1
Dimension C	3	2	4	1
B. PROJECT SCORES				
Dimension A	100	25	75	50
Dimension B	10	70	40	100
Dimension C	40	70	10	100
C. "NORMALIZED" PROJECT SCORES				
Dimension A	100	25	75	50
Dimension B	7.5	52.5	30	75
Dimension C	20	35	5	50
NORMALIZED PROJECT SCORE TOTALS	127.5	112.5	110.0	175.0

$$\text{TOTAL "NORMALIZED" PROJECT SCORE} = 127.5 + 112.5 + 110.0 + 175.0 = 525.5$$

Condition 1 implies that the criteria dimensions used should not involve a number of variables that are merely different names for the same phenomenon. In applied terms, an having counts of each bacterial species in a water sample to determine the degree of pollution is redundant when an aggregate count of certain types of bacteria measures the same thing. Similarly, from a parks management perspective, if a score is to reflect whether there is or is not historic development and to reflect the degree of that development, then it is more appropriate to include a single variable than to include a number of variables. Usually if, say, an historic building exists, artifacts, signs, habitation and other conditions "automatically" exist, so it is redundant to register their existence by having a great "checklist" of variables to use in "scoring". The aggregate score obtained by summing the scores on the individual variables can be a poor site selection criterion and a poor general indicator of the nature of the site.

Pursuing the earlier example, the existence of certain bacteria in an area designated for swimming is only relevant if these bacteria represent a hazard or are present when some hazard is likely. The total aggregate count of harmful bacteria may indicate the level of pollution to a manager of the area. One count, rather than counts of a variety of species, tells the planner whether or not the water is sanitary enough for swimming.

Turning to the actual operation of scoring, it is usual that for each specific criterion dimension the most favourable alternative is given the score of 100. Similarly the least favourable alternative is given a score of 0 or 1; the remaining projects or alternatives are scored between the two extremes. This effort is referred to as the "within variables scoring".

The problem with "within variables scoring" is that, with respect to each variable, the best alternative is not necessarily 100 compared to the lowest scoring project. Realistically, a 100 magnitude of difference in scoring probably does not occur very often. In their 1968 proposal to the California State Department of Parks and Recreation, Alfred Baxter and Associates (1965)

showed how simply this problem could be eliminated. Their solution was to give the least favourable project a score that reflected its "status" in relation to the alternative that scored 100 along the same dimension. This scoring method encourages one to use his insight about a relation to express the fact that, for example, along dimension i , the best alternative is "4 times" more preferable than the worst alternative. Such an insight would be expressed by giving the poor project a score of $100/4 = 25$. Now if on criterion j , the preferred alternative is 20 times better than that of the least preferred, the least preferred would receive a score of 5. (Because it is then assumed that the other projects take values at equal intervals between the best and the worst, the inclusion of a very bad project in the scoring will automatically reduce the value of intermediate projects relative to the best. An exceptionally good project will have the opposite effect. As a partial solution to this bias, not only the worst, but the median project (and even the 25th and 75th quartile projects) should be compared to the best, with intermediate projects taking values at equal intervals between.)

The flexibility introduced by not setting the lowest score at zero means in some way that "within variables scoring" is considered independently of the weighing of the variables. When "within variables scoring" has been accomplished, there remains the problem of assigning relative importance to each criterion dimension. Because each criterion dimension has a highest score of 100, assigning importance to the dimensions is independent of scoring within the dimension. What must be done is to compare the criteria that score 100, regardless of the project to which they apply and evaluate how variable X should relate to variable Y , variable Y to variable Z , etc. This is done until one has, say $X/Y = 4$, $Y/Z = 2$, $Z/W = 1$, $W/M = 1.1$. The series just given suggests relative weights of 8, 2, 1, 1, for X, Y, Z, W, M respectively.

The Alfred Baxter and Associates (1965) criterion dimensions for a cost-benefit scoring procedure to evaluate land for acquisition as a park can be used to illustrate in applied terms the import of the concepts reviewed to this point. Their criteria are:

1. Suitability for state acquisition;
2. "Ability" of the site to generate demand under a given development option;
3. Endangered future availability;
4. Quality and balance of the recreation experience provided;
5. Economic contribution to the surrounding a community;
6. Open-space values;
7. Appropriateness in meeting deficit in recreation facilities.

These criteria were given the following set of weights:
 $c_1 = 100$, $c_2 = .67$, $c_3 = .51$, $c_4 = .85$, $c_5 = .33$, $c_6 = .35$,
 $c_7 = .62$. Now, the merit of these consultants' handling of "within variables scoring" has already been acknowledged. But the adequacy or appropriateness of the above criterion dimensions with reference to conditions 1 and 2 cited earlier is dubious. The Baxter and Associates' criteria 3, 4, and 6 are meaningful. However, criterion 7 appears to be misplaced: there is no need for any project if there is no present or future deficit in recreation facilities recognized. As to criteria 2 and 5, the degree of relevance of these criteria should be dealt with by some type of economic analysis (as suggested later). Arbitrarily assigning them weights ignores the fact that they are two interrelated factors, which should be considered as part of a rigorous, totally quantitative cost-benefit analysis. Finally, suitability (stated as criterion 1) should not be used as a criterion because it refers either to ability to generate demand under some development option (criterion 2), or to such factors as land quality for development (criterion 4) or economic suitability (criterion 5). Its inclusion involves either redundancy or irrelevance and may even be a guess at

the score to be derived though it is part of the score. One can in fact see this variable as a "cooking" factor. It has the highest weight of all criterion dimensions is simply used to bias the final score.

INTEGRATING MULTI-DIMENSIONAL SCORING WITH COST EFFECTIVENESS ANALYSIS

Thus far, a clear distinction between how the variables that can be measured economically should be treated and how intangible variables should be treated has not been made. Put, as implied in the last section, this is where the Baxter and Associates' scoring method has shortcomings.

- A. There is a great temptation to over-weigh the mass of non-output-oriented (intangible) criteria with respect to "production efficiency".
- B. An adequate link between "suitability" and the classical cost-effectiveness, "C-E", approach is not specified: there is a need to tie the suitability of a project more directly to its output to cost or quantity to cost ratio, "Q/C?".

It is important to account for the non-quantifiable aspects by a scoring system, but it should not be balanced with quantitative analysis.

In short, the manager should realize that the ultimate aim of project evaluation is to produce output not maximize scores that using factors that may even relate to low output at high cost. Accordingly, the role of intangible scoring should be limited to the production of some index of "aggregate suitability". This "aggregate suitability" can be used to modify the output/cost ratio but it should presumably not dominate an evaluation *unless this is a clear and conscious choice of the manager responsible for project selection.*

The preceding discussion gives the rationale for the third condition for scoring in project selection:

Condition 3. Overall suitability of a project should be determined by using a multi-dimensional scoring method on "intangibles" to define a score that can be used to modify a C-E evaluation (1) to a fixed degree, and (2) in a well defined and understandable way.

To achieve "true" integration of the two kinds of results, the manager may, for example, calibrate his aggregate "intangible" score in such a way that for an average suitability score, the output to cost ratio, Q/C, becomes the only consideration.

Consider that a recreation facility or service manager's intuition can be used to decide what percentage of influence, "d", intangible factors should have on the tangible output/cost ratio for particular projects being evaluated. Then it is an easy matter to recalibrate the total scores to get an index, IS, that will have a value of 1 + d for the project with the lowest intangible total score and will have a value of 1 - d for the highest score. If d = .25 then IS will be .75 for the lowest score project and 1.25 for the highest scoring project. The transformed values are merely calculated using the following equation which is written in a way that makes it easy to see the rationale on which the equation is based. The equation is to compute IS(j), the index for the project j is:

$$IS(j) = (1.0 + (\text{score for project } j - (\text{best score} + \text{poorest score})/2) / (\text{best score} - \text{poorest score}) / 2$$

To emphasize why IS(j) was calculated it seems reasonable to call it the index of intangible project suitability.

Now, if one has the results of a rigorous cost efficiency analysis, then for each project, j, Q/C (j) is available. The crux of what is proposed here is that IS(j) can and should be used as a multiplier of Q/C (j) to produce a modified cost-efficiency ratio (to become the "cost-effectiveness ratio"):

$$MQ/C (j) = IS(j) (Q/C (j)).$$

Given that IS(j) has been calculated using the scoring approach suggested and that relevancy and redundancy of the variables used in calculating IS(j) have been appropriately considered, the manager ends up with a MQ/C (j) that is influenced by at most a known amount by intangibles. Assuming that Q/C (j) was calculated correctly, decision making should benefit from the very rational use of information on intangibles endorsed here.

APPLICATIONS OF THE METHOD OF CALCULATING THE INDEX OF INTANGIBLE SUITABILITY TO LAND ACQUISITION DECISION MAKING

Assume that four pieces of land are proposed for acquisition under, for instance, a Province's recreation program. Assume that three dimensions of intangible factors were felt to be relevant, and the scorings of 4 developments along them are given in Table 2 (Part A). Assume further that, along dimension A, the best proposal is deemed, by some evaluation committee, four times preferable to the worst and that this ratio is equal to ten for both dimensions B and C. This would lead to the three cardinalizations shown in Part B of Table 2 which are the "within variables scoring". To compare across criteria assume that the evaluation committee assigned the relative weights 1.00, .75 and .50 respectively to dimensions A, B and C. Applying these weights to transform the scores in part B of Table 2 one gets the results shown in part C of Table 2. The project scores are then easily derived by totaling across criteria in Part C of Table 2. The results are shown at the bottom of Table 2. As we can see, the spread of "suitability" values is from 110.0 to 175.0.

Assuming a spread of 10 percent is all that top management is willing to concede to considerations other than the project efficiency, the aggregate score must be recalibrated according to procedure introduced earlier to yield an index IS that ranges in value from .9 to 1.1 and which is centred around the mean value of the intangible scores for all projects considered.

Using the formula to calculate IS(j) with $d = .1$ results in:

$$IS(j) = 1 + ((\text{score for } j - (175 + 110)/2) / (175 - 110))$$

$$IS(j) = 1 + .00308 (j - 142.5)$$

So the index values are .95, .91, .90 and 1.1 for projects 1 to 4 respectively.

Given these values it is only a matter of multiplying and dividing to get the columns in Table 3. The modified project evaluation information is shown in the table.

TABLE 3: DATA AND FINAL COMPUTATION OF RESULTS TO OBTAIN THE MODIFIED COST-EFFECTIVENESS RATIO FOR FOUR PROJECTS

Project	1	2	3	4
Total Output* (physical units)	15	30	10	20
Total Cost* (\$,000 s)	20	30	10	15
Output-to-cost ratio Q/C	0.75	1.00	1.00	1.33
Suitability (IS)	0.95	0.91	0.90	1.10
Modified Cost- Effectiveness Ratio	0.71	0.91	0.90	1.46

* Outputs and costs are to be discounted to their present values.

Actually, in this example, project 4 would have been selected even if the IS(j)'s had all been equal (it is not being pushed upwards by the non-quantifiable aspects A, B, and C of the situation). However, if the realization of project 4 was barred because of some rigid constraint

(e.g. political infeasibility) the method permits identifying the next alternative which would be project 2 (two) in this case.

SUMMARY AND CONCLUSIONS

As pointed out, scoring alternatives according to several criterion dimensions by the means of an aggregate score is widely practiced. Given the faults which have been indicated in the procedures usually used, it is particularly unfortunate that the minor modifications suggested here are not regularly used. When it is recognized that the primary objective of an individual project is to produce tangible output rather than an intangible score the importance of separating the intangible analysis from the tangible analysis is not hard to recognize. In fact, since the author suggested the use of a scoring technique suggested by others, the only original contributions of the article are the suggestion that tangibles be treated separately and the presentation of a way to do this. The correct role of the index of intangible suitability, too commonly used as the sole scoring criterion, is to serve as a modifier of the Q/C ratio.

In a practical vein, if the reader basically accepts the approach suggested one may find it useful to have Steps 1 to 6 as a guide in doing his own analyses. The actions suggested follow logically but without some thought or experience with the technique suggested, one might end up doing computations or taking other actions that the following indicate as unnecessary.

Step 1: Work out the Q/C ratios and, if there are no intangible factors to consider, the analysis is complete. The project with the best Q/C ratio is selected to be carried out.

Step 2: Ask management to decide on the d value for the particular situation being considered. If the best project's Q/C ratio is more than $1 + d/(1 - d)$ times the second best there is no need to go any further in order to find the most suitable alternative. The higher Q/C ratio cannot be "pulled" down to make it of a lower score by any non-qualifiable influences being taken into account.

Step 3: If criteria other than mere efficiency are deemed relevant, then list the non-quantifiable factors that are deemed relevant. Make sure these are truly choice-dimensions (i.e. pure criteria). Remove those not obviously called for (irrelevant or redundant) and those better expressed as constraints.

Step 4: Establish the "within variables scoring" rationale.

Step 5: Establish the between dimension weights and use these and the special formula to get IS(j)'s.

Step 6: Score projects according to their modified cost-benefit ratios.

This will establish the scoring of the alternatives according to all criteria dimensions, output-quantifiable or not. But one consideration not raised previously is that, in practice, situations might arise in which the more cost-effective projects are not the largest ones. It may occur for example that in order to supply all user-days of recreation or in order to exhaust the budget, some of the lower scoring projects may have to be selected along with a best project. But, this is a problem of budget allocation and the issues that arise are too numerous to consider here.

In conclusion it seems important to reiterate that when decisions are to be made concerning the choice of projects for the allocation of funds, decision makers should examine all the factors involved -- whether they are quantifiable or non-quantifiable. At the same time the major objective of developing a facility must not be lost; output is what an agency wants. The effectiveness of the agency should be evaluated by that.

In addition, a certain amount of intuition about how the intangibles and non-quantifiable variables influence the output is desirable. In order to use this intuition effectively, a systematic

procedure is needed. This paper has demonstrated one way in which a systematic integration of the multi-dimensional scoring method into the common cost/effectiveness analysis can be made.

It is the last point that best typifies the type of analyses being suggested throughout. It is not really that it is never-ending, but rather that different kinds and depths of analyses may be used depending upon the problems being confronted. And, of course, initial analyses will suggest further questions. While unexplained variance, or more euphemistically, unknowns, will persist and call for judgment based on the experience of planners, it is certain that helpful guidance will stem from such quantitative study.