Chapter 3

Decision-making Concepts & Methodologies

Alternatives - Pros and Cons

Perhaps the most common 'formal' approach to making a choice among alternatives is to list the pros and cons of each alternative. A list of pros and cons is often embedded in a memorandum with a format shown below:

From:
To:
Issue: (i.e. problem or opportunity)
Alternative 1:
Pros:
Cons:
Alternative 2:
Pros:
Cons:

A meeting is then held to discuss the alternatives and their pros and cons.[1]

The meeting is relatively unstructured and the discussion jumps from topic to topic in a haphazard way. Some relevant points are discussed several times, some not at all. Frustration develops as "strong" members of the group dominate the discussion, often far beyond the level justified by their knowledge and experience. Shy members of the group may fail to speak at all, even though they may have important information to convey.

A decision is "made" either when it appears that no more progress is forthcoming, or when it is time to go to another meeting (or lunch), or time to go home for the day. This haphazard approach to decision-making is far too common. Even a small change in the decision process to a structured discussion of each alternative's pros and cons would be a vast improvement. Time would not be wasted by jumping back and forth, discussing some

[1] Decisions are often made without even this much preparation. A meeting may be held to make a decision without any serious attempt to identify alternatives or list their pros, and cons.

points over and over again, and inadvertently failing to discuss some issues at all.

However, even if we conducted a structured discussion, it is still not clear how to make the decision. Certainly it would be wrong to calculate the net number of pros over cons for each alternative and then select the alternative with the largest net number because the relative importance of the pros and cons differ. How then can one proceed?

Benjamin Franklin considered this problem over two hundred fifty years ago.

In a letter addressed to Priestly[2], Franklin explained how he analyzed his decisions:

> Dear Sir: In the affair of so much importance to you, where in you ask my advice, I cannot, for want of sufficient premises, advise you what to determine, but if you please I will tell you how. When those difficult cases occur, they are difficult, chiefly because while we have them under consideration, all the reasons pro and con are not present to the mind at the same time; but sometimes one set present themselves, and at other times another, the first being out of sight. Hence the various purposes or information that alternatively prevail, and the uncertainty that perplexes us. To get over this, my way is to divide a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con. Then, during three or four days consideration, I put down under the different heads short hints of the different motives, that at different times occur to me, for or against the measure. When I have thus got them all together in one view, I endeavor to estimate their respective weights; and when I find two, one on each side, that seem equal, I strike them both out. If I find a reason pro equal to some two reasons con, I strike out the three. If I judge some two reasons con, equal to three reasons pro, I strike out the five; and thus proceeding I find at length where the balance lies; and if, after a day or two of further consideration, nothing new that is of importance occurs on either side, I come to a determination accordingly. And, though the weight of the reasons cannot betaken with the precision of algebraic quantities, yet when each is thus considered, separately and comparatively, and the whole lies before me, I think I can judge better, and am less liable to make a rash step, and in fact I have found great advantage from this kind of equation, and what might be called moral or

[2] Letter from Benjamin Franklin to Joseph Priestly in 1772, is taken from: "Letter to Joseph Priestly", Benjamin Franklin Sampler, (1956).

A Fighter Aircraft Selection Problem

Attributes

Alternatives		Maximum	Ferry	Acquisition	Reliability	Maneuverability
	speed	range	payload	cost		
(Ai)	(Mach)	(NM)	(pounds)	($ x 10^7)	(high-low)	(high-low)
A1	2.0	1500	20000	5.5	average	very high
A2	2.5	2700	18000	6.5	low	average
A3	1.8	2000	21000	4.5	high	high
A4	2.2	1800	20000	5.0	average	average

Figure 1 – Fighter Aircraft Selection Problem

> prudential algebra. Wishing sincerely that you may determine for the best, I am ever, my dear friend, yours most affectionately.

Franklin's insights were far beyond his time. He sensed that the human brain is limited in the number of factors that can be kept in mind at any one time. Psychologists discovered this in the mid twentieth century. He knew enough not to rush an important decision, but to devote several days to the study of the pros and cons, attempting to make tradeoffs based on relative importance.[3] And he tried to develop a "moral or prudential" algebra, recognizing that some way of measuring qualitative as well as quantitative factors was necessary in the decision process. Not only was Franklin far ahead of his time, but his decision process is superior to that used by the vast majority of today's decision makers who fool themselves into thinking that they can make a good decision after a few hours of unstructured discussion and some "hard" thought.

---

[3] Relative comparisons are a key component of the Analytic Hierarchy Process (AHP), developed two hundred fifty years later.

Although Franklin was able to reduce the number of factors (pros and cons) under consideration by a process of cancellation, he could not complete a decision analysis because of his inability to “measure” the remaining factors. We have had significant advancements in the meaning and use of numbers and measurement since Franklin’s time. However, our use of numbers and measurement can sometimes be misleading or even *wrong*.

Consider the Fighter Aircraft Selection Problem illustrated in Figure 1.

Attributes[4] are shown for four fighter alternatives. How would you decide which alternative is best? Of course, the answer depends on many things not specified, such as who your enemies are, how far away are they, what planes do they have or are expected to have, and so on. But even if you knew all of these things, how would you decide? How would you ‘weight’ the importance of each attribute? Can you use the ‘data’ without applying judgment? Is a plane that can fly 2.5 Mach only 2.5/2.2 times as preferable with respect to speed as one that can fly 2.2 Mach? How would you combine performance on speed, range, payload and cost? Many people erroneously think that just by normalizing each attribute on a similar scale, such as 0 to 100, they can just multiply attribute values by weights and add. What would you do with the qualitative attributes, such as maneuverability?

Consider the following study by a Washington D.C. based think tank:

> The study attempted to evaluate 100 cities in the United States in terms of their attractiveness to minorities. The evaluation was performed with a matrix of seven columns and 100 rows. Each column represented a factor, such as employment or housing. Each row represented a city. The study (somehow) ranked each of the 100 cities for each of the factors. The city considered best for a particular factor was “given” a 1 while the city considered the worst was “given” a 100. The numbers in each row were then added to obtain a total score for each city. The cities were then ranked by their “attractiveness” scores, with a low score being more attractive than a high score.

[4] We will deviate from this approach and focus on objectives rather than attributes

Several things are wrong with this "evaluation". Take a minute and write down what you think is wrong.

Most people will immediately note that the factors received equal weight. While this is indeed generally inappropriate (we will consider how factors can be "weighted" later) there are even more serious errors involving the misuse of numbers.

Misuse of Numbers

Not weighting the factors was an error of omission and is often fairly obvious. A more serious error, one of commission rather than omission, and one that is more likely to go unnoticed, is the inappropriate addition of ranks. The "scores" in the above example are "ordinal" numbers, representing rank, but nothing more. It is just plain wrong to add these numbers. Any results are meaningless!

How can we identify such mistakes? By thinking about the meaning of the numbers. A city "scoring" a 1 on salary is ranked better than a city "scoring" a 2, but how much better? Is the interval between cities ranked 1 and 2 the same as the interval between cities ranked 21 and 22? Not necessarily. The city ranked first might be ten thousand dollars higher than the city ranked second, while the interval between the 21st and 22nd cities might be only fifty dollars. The numbers in this study are, according to Stevens' classification scheme[5,6,7], ordinal in nature and cannot be added.

In the 1990 edition of *Retirement Places Rated*, author David Savageau ranked 151 retirement places on each of seven criteria and then added the ranks to determine an overall ranking. Using this methodology, Las Vegas Nevada came in 105th. In the next (1994) edition of *Retirement Places Rated* Savageau improved his methodology by 'scoring' (instead of ranking) 183 cities on the seven criteria and then averaging the scores to determine an overall score. This time, Las Vegas came in first! (See page 115 for details).

---

[5] S.S. Stevens, "On the Theory of Scales of Measurement", Science, (103, 1946), pp. 677-680.
[6] F.S. Roberts, Measurement Theory with Applications to decision-making, Utility and the Social Sciences, (London, Addison Wesley, 1979).
[7] Jean Claude Vansnick, "Measurement Theory and Decision Aid", Proceedings of the Third International Summer School on Multiple Criteria Decision Methods, Applications and Software, (Monte Estoril, Portugal: July 1988).

The misuse of numbers is one reason that numerical analyses are sometimes flawed. Unfortunately, decision models based on flawed numerical reasoning[8] may go undetected, leaving the decision-maker left wondering why the results do not make sense. It is any wonder that some of our top-level executives are somewhat suspicious of numerical analyses such as this? But the misuse of numbers is not a reason to forego using them. We just must be careful that a sound theoretical foundation exists for whatever we do.

In order to avoid making such errors, we will take a brief detour and review what is known as levels of measurement. Let us briefly look at Stevens' categorization.

Levels of Measurement

According to Stevens' categorization, there are four levels of measurement. The levels, ranging from lowest to highest are Nominal, Ordinal, Interval, and Ratio. Each level has all of the meaning of the levels below plus additional meaning.

Nominal[9]

Nominal numbers, the lowest level in terms of the meaning conveyed, are just numerical representations for names. Nominal numbers are used for identification purposes only and imply nothing about the ordering. Telephone numbers and social security numbers are nominal. Are you 'older' or 'better' than someone else because your telephone number is higher? Obviously not, and people rarely make mistakes with nominal numbers. However, errors arising from the misuse of ordinal numbers are not so rare as we will see next.

[8] Tom R. Houston, "Why Models Go Wrong", Byte Magazine, (October 1985), pp. 151.

[9] A mathematical definition of nominal is: admits any one - to - one substitution of assigned numbers.

Ordinal[10]

Ordinal numbers[11], as the name suggests, implies an order or ranking among elements. The order may be either increasing or decreasing depending on the application. For example, we could order or rank cities based on population by ranking the city with the highest population as #1, or by ranking the city with the lowest population #1. A ranking implies an ordering among elements but nothing more. It does not imply anything about the differences (or intervals) between items. For example, if we know only that a professional baseball team finished in second place at the end of the season, we do not know if the team was one game behind the first place team or twenty games behind. Care must be taken not to add or multiply ordinal data. Errors arising from the addition of ordinal data are far too common.

Interval[12]

Interval scale data possesses the meaning of Nominal and Ordinal data, as well as having meaning about the intervals between objects. Corresponding intervals on different parts of an interval scale have the same meaning. If we have interval level data then we can infer that the interval between two objects with values of 20 and 5 (an interval of 15) is equivalent to the interval between two objects with values of 80 and 65. Interval level data can be used in arithmetic operations such as addition and multiplication. However, after adding interval level data, one can not infer that a total of 100 is twice as good as a total of 50. If one were to allocate resources based on this inference, then the allocation would be incorrect.

---

[10] A mathematical definition of ordinal is: can be transformed by any increasing monotonic function.

[11] Ordinal numbers are sometimes referred to as rank numbers. Likert scales used in many marketing studies are ordinal numbers.

[12] A mathematical definition of interval is: can be subjected to a linear transformation, or is invariant under the transformation $Y = a X + b$.

Ratio[13]

Ratio level data (sometimes called ratio scale) is the highest level, having Nominal, Ordinal, and Interval properties, as well as the property of ratios. Corresponding ratios on different parts of a ratio scale have the same meaning. If we have ratio scale data, then the ratio between two objects with values of 100 and 50 is equivalent to the ratio of two objects with values of 6 and 3. A ratio scale is often defined as one having a true zero point. However, for our purposes, it is easier to think of a ratio scale as one for which equivalent ratios are considered equal. Temperature measured on the Fahrenheit scale is not a ratio measure, since there we would be wrong to infer that there is twice as much heat when the temperature is 80 degrees as when the temperature is 40 degrees. (If we used the Kelvin scale, which has the ratio property, then such an inference would be correct.)

Numbers at the Racetrack[14]

Suppose the owner of a horse-racing stable is interested in buying a particular horse. He studies the results of the horse's last five races. Let's consider the numbers used to designate the results of a particular race.

> The number worn on the horse and jockey is *nominal*—it identified the horse so that people at the track could look in the racing form to see the horse's name, owner, etc. The number conveys no information about the horse's *order* of finish in the race.
>
> The finishing position for each horse is *ordinal.* The first place horse finished ahead of the second place horse, the second place horse finished ahead of the third place horse, and so on. Knowing that this horse finished first, however conveys no information about how far in front, or the *interval* to the second place horse.
>
> The number of lengths to the next finisher is an *interval* measure. Knowing that the horse finished first by 15 lengths as opposed to 5 lengths is important information not conveyed by the order of finish alone. However, even this measure may not tell us as much as what we want to know.

---

[13] A mathematical definition of ratio is: admits multiplication by a constant, or is invariant under the transformation $Y = a X$. A ratio scale is said to have a true 'zero', however the true zero can be conceptual and need not be observable.

[14] Thanks to Chuck Cox of Compass Associates for this example.

Suppose the horse finished first by 15 lengths in a 2-½ mile race. Is this as strong as finishing first by 15 lengths in a ½ mile race? No. More information would be conveyed if we knew the *ratios* of the times of the first and second place finishes.

Mathematical operations allowed:

Nominal, ordinal, interval, ratio – succeeding scales have additional meaning and can be used in more arithmetic operations as summarized below:

- Addition/subtraction and multiplication/division require at least interval level meaning.
- An interval level number can be multiplied by a constant or a ratio level number but cannot be multiplied by another interval level number.
- There are no restrictions when using ratio level numbers.
- A decision method that produces ratio scale numbers is the most flexible and accurate.

How can you know the level of measure in your numbers?

Although measurements from scientific instruments are often ratio scale, the level of measurement in most social and decision contexts depends on the intent of the subject responding to a question or making a judgment. One need not take a college course to learn how to tell the difference between Nominal, Ordinal, Interval, and Ratio scales. Rather, one needs to ask the question: *What meaning does this data have for the purpose at hand?* For example, if you asked someone to express a preference on a scale of 1 to 5, and they specify a 4 for one item, a 3 for a second item, a 2 for a third item, and a 1 for a fourth item, you can easily infer that the values possess the ordinal property. Whether or not the data possesses the interval property depends on the respondent's intent and understanding of the numbers. The scores would have the interval property only if the subject's intent or understanding was that corresponding intervals are equivalent, e.g., the interval between 1 and 2 is equal to the interval between 4 and 5. Likewise the scores would have the ratio property only if the respondent's intent was, for example, that a score of 5 is five times better than a score of 1. If the respondent's intent was not interval or ratio, the numerical response possesses only the ordinal property. Interval or ratio properties are often assumed for scales with little or no justification. In most graduate schools the letter grade of 'A' represents excellent work, 'B' represents good work, while 'C' is almost failing. The implied interval between an 'A' and a 'B' is much smaller than the interval between a 'B' and a 'C'. But grade point averages are calculated by assigning 4 to an A, 3 to a B, and 2 to a C - which incorrectly measures the interval between an 'A' and 'B' as the same as between a 'B' and 'C'. Furthermore, the grade point average assumes that a 'A' is 1.333 times better than a 'B' and that an 'A' is twice as good as a 'C'. While the former ratio might be reasonable, the latter certainly is not - a 'A' is probably more like 10 or 50 times better than a 'C'. The only conclusion we can draw is that the grade point average calculation is mathematically meaningless! Later we will show how easy it is to obtain ratio level measurement for both quantitative as well as qualitative factors using the Analytic Hierarchy Process.

Let's pause to see where we are and where we are going. We looked at what is perhaps the most common approach to making a choice among alternatives, listing their pros and cons. We then looked at how Benjamin Franklin ingeniously traded off some of the pros and cons, reducing the problem to one of more manageable size. We took a small detour and looked at how we might be lured into misusing numbers if, in an attempt to carry the evaluation process further, we introduce and manipulate numbers without regard to their underlying level of measurement. Let's continue and see how numbers can be used (with care) in a scheme often referred to as weights and scores.

Table 1 – Typical Weights and Scores Matrix

	Crit. 1	Crit. 2	Crit. 3	Crit. 4	Crit. 5	Crit. 6	...	Crit. 7
Weights	80	40	25	90	3	1	...	17
Alt. 1	5	10	2	9	etc.			
Alt. 2								
Alt. 3								
Alt. 4								
....								

Weights and Scores

The typical weights and scores approach consists of a matrix with criteria as column headings and alternatives as rows[15] as shown in Table 1. Criteria are "assigned" weights using a scale such as 0 to 10 or 0 to 100. Then each alternative is scored against each criterion in a similar fashion. The alternative scores for each criterion are then multiplied by the weight for the criterion and summed to give a total score for each alternative

This score represents the overall preference for or performance of the alternative. *If* used carefully, weights and scores can be an effective methodology. There are, however, several practical difficulties:

First of all, when assigning weights, what do the numbers really mean? On a scale of 0 to 100, what is an 80, or what is a 40? When you "give" an 80 to one criterion and "give" a 40 to another, do you really mean that the 80 is twice as important as the 40? If the answer is yes, then the weights possess the ratio scale property. But how can you be consistent enough to insure this if you are dealing with 20, 30 or 100 criteria? If you assigned an 80 to the first criterion and later assign a weight of 10 to the 95th criterion, can you remember what you did earlier and do you really mean that the 95th criterion is only one eighth as important as the first criterion?

[15] Sometimes criteria are shown in rows and alternatives in columns.

Channel capacity and short term memory

Experiments have proven time and time again that the human brain is limited in both its short-term memory capacity and its discrimination ability (channel capacity) to about seven things.

According to James Martin[16], if a person "has to choose from a range of 20 alternatives, he will give inaccurate answers because the range exceeds the bandwidth of his channel for perception. In many cases, seven alternatives are the approximate limit of his channel capacity." Martin's conclusion is based on the results of numerous psychological experiments, including the well known study "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Information Processing," by G. A. Miller.[17] It has been demonstrated time and time again that humans are not capable of dealing accurately with more than about seven to nine things at a time, and we are just fooling ourselves if we try. To demonstrate this to yourself, have a friend make up a seven digit number and read the digits to you. After hearing the last digit, try to write down each of the seven digits in sequence. You will probably be able to do so without a mistake. Now try it with nine digits. You probably will not be able to recall the digits without a mistake. (If you did, you are above average as only about 11% of the population can recall 9 items from their short term memory — now try it with eleven digits!). Although we do not like to admit it, our brains are quite limited in certain respects.

Not only have psychologists demonstrated that humans have difficulty dealing with more than about seven plus or minus two factors at a time, but there is a mathematical basis for this phenomenon. Thomas Saaty, has shown that to maintain reasonable consistency when deriving priorities from

---

[16] James Martin, *Design of Man*-Computer *Dialogues*, (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973).

[17] G.A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Information Processing." *Psychological Review*, (Vol. 63, No. 2, March 1956), pp. 81-97. The distribution is a bell shaped curve with an average of 7. Obviously some people can recall more than seven, some fewer. But only about 11 percent of the population can recall 9 things from their short term memory, still fewer 10 things, and so on.

paired comparisons, n, the number of factors being considered must be less or equal to nine.[18]

Need for Hierarchical Structure

Suppose we try to assign weights to 20, 30 or 100 columns; there are bound to be many mistakes! We can cope with this complexity as we do with other complex situations - by arranging the criteria into groups[19], each group having sub-groups, and so on. This "hierarchical" arrangement has been found to be the best way for human beings to cope with complexity. (See page 13).

Humans have had to learn how to deal with complexity. We discover this time and time again. For example, if you try to recall a sequence of nine or eleven digits as someone reads them, you will probably find yourself grouping (psychologists call this chunking) the digits into groups in an effort to overcome the limitations of your short-term memory. If you prepare a presentation with ten or fifteen bullet items, you will probably find yourself organizing them into categories and sub-categories each with nine or fewer elements so that you do not lose your audience).

Orders of magnitude

Another problem arises when a weights and scores approach involves more than a handful of criteria (columns). Some of the criteria might be "orders of magnitude" more important than others. For example, if one criterion is assigned a .02 on a scale of 0 to 10, and another is assigned a 9, do we really mean that one is 450 times more important than another? Our ability to accurately compare things that differ by orders of magnitude is not nearly as good as our ability to compare things that differ by less than an order of magnitude. A hierarchical grouping avoids this problem as well.

---

[18] Saaty, T.L., The Analytic Hierarchy Process, New York, N.Y., McGraw Hill, 1980, reprinted by RWS Publications, Pittsburgh, 1996, p. 195.

[19] Groups of seven or fewer if we do not want to exceed our capabilities.

Arbitrary assignment

Another difficulty with the weights and scores methodology stems from the “assignment” of weights and scores in what often appears to be an arbitrary fashion. How can we justify that a criterion (such as customer perception) was “given” an 8? What does the 8 really mean? Certainly, we can refer to studies and experience with similar decisions to show how important the criterion is, but why “assign” an 8? Why not a 7, or a 9? Similarly, when scoring a particular alternative with respect to a criterion such as customer perception, we may refer to customer interview studies and past experience and “assign” a 3. But what does the 3 really mean?

Absolute versus relative

This difficulty is due to two causes. The first is the weighing and scoring in an absolute fashion rather than a relative fashion[20]. The second is the implied precision of numbers in situations where we know the precision is not justified. We can lift two objects, one in each hand and estimate that one is 2.7 times heavier than the other, or we can look at two houses and estimate that one house should cost 1.5 times the other. But do we expect that the implied precision of our estimates (e.g. 2.7 rather than 2.5 or 2.8) is appropriate or justifiable? And suppose we had to compare the relative importance of employee moral with customer satisfaction. How can we justify <u>any</u> specific number that we might think of?

Words instead of numbers

Justification would be much easier if we use a less precise way of expressing judgments, such as words instead of numbers. Suppose we were to use words instead of numbers. Words are often easier to justify than numbers. For example, if you say that, with respect to corporate image, alternative A is 3 times more preferable than alternative B, can you justify why it is exactly 3? Why not 2.9, or 3.1? But if you said, instead, that A is

---

[20] Martin observed that there are several ways of increasing our effective channel capacity. One of these is to enable the subject to make relative rather than absolute judgments.

"moderately" more preferable than B, this can be justified with a variety of arguments, including, perhaps, some hard data.

But what can you do with the 'words'?

But what then do you do with verbal judgments such as "moderate"? Words have different meanings to different people. True, anyone can put arbitrary numbers behind words in a computer program. But will the numbers "accurately" reflect the meaning the words had to the individual or group making the judgments?

Even if the decision-maker specifies numerical equivalencies for words, will he or a group of his colleagues consistently remember the assignments accurately enough to insure that the results reflect their actual judgments? Will "errors" due to the use of imprecise words be a problem?