Rev.R.Acad.Cienc.Exact.Fis.Nat. (Esp)
Vol. 93, N.º 4, pp 395-423, 1999
Monográfico: Problemas complejos de decisión. II

BASIC THEORY OF THE ANALYTIC HIERARCHY PROCESS: HOW TO MAKE A DECISION

THOMAS L. SAATY

University of Pittsburgh

1. INTRODUCTION

An assumption arising from the practice of science and engineering since the middle ages is that because nature is physical, we should be able to relate all measurement to physical dimensions. But that is not true. Human thinking and feeling exist in the physical world but they are not matter or gravity or electromagnetism in the strict sense science understands them today. They are intangible. The human experience involves a very large number of intangibles. In general and with few exceptions, intangibles cannot be measured on a physical scale. However, they can be measured in relative terms through comparison with other tangibles or intangibles with respect to attributes they have in common (taken one at a time) and a ratio scale can be derived from them that yields their relative measurement values. The attributes are themselves compared as to their importance with respect to still higher attributes, relative measures derived, and so on up to an overall goal.

Ratio scales are fundamental for capturing proportionality. All order at its most sophisticated level involves proportionality of its parts in making up the whole, and in turn the proportionality of their smaller parts to make up the parts and so on. Without such proportionalities there would be no definable relation among the parts and the resulting structure or function of the system under study would appear to us as arbitrary.

When one speaks of relative measurement, those of us trained in the physical sciences and in mathematics are likely to think of scales used to measure objects. For example, on a scale such as the yard or the meter, each with its units, we divide the corresponding measurements of lengths to get the relative lengths. But that is not what I mean by relative measurement. First, I ask what would I do if I did not have a scale to measure length to define the relative length of two objects? Henri Lebesgue [13] wrote:

«It would seem that the principle of economy would always require that we evaluate ratios directly and not as ratios of measurements. However, in practice, all lengths are measured in meters, all angles in degrees, etc.; that is we employ auxiliary units and, as it seems, with only the disadvantage of having two measurements to make instead of one. Sometimes, this is because of experimental difficulties or impossibilities that prevent the direct comparison of lengths or angles. But there is also another reason.

In geometrical problems, one needs to compare two lengths, for example, and only those two. It is quite different in practice when one encounters a hundred lengths and may expect to have to compare these lengths two at a time in all possible manners. Thus it is desirable and economical procedure to measure each new length. One single measurement for each length, made as precisely as possible, gives the ratio of the length in question to each other length. This explains the fact that in practice comparisons are never, or almost never, made directly but through comparisons with a standard scale.»

But when we have no standard scales to measure things absolutely, we must make comparisons and derive relative measurements from them. The question is how, and what have we learned in this process?

We should note that we are not talking about a proposed theory that we can accept or reject. Comparisons leading to relative measurement is a talent of our brains. It has been neglected in science because we have not learned to formalize it in harmony with the usual way of creating standard scales and comparing or measuring things on them one at a time.

The cognitive psychologist Blumenthal [5] writes:

«Absolute judgment is the identification of the magnitude of some simple stimulus, ..., whereas comparative judgment is the identification of some relation between two stimuli both present to the observer. Absolute judgment involves the relation between a single stimulus and some information held in short-term memory - information about some former comparison stimuli or about some previously experienced measurement scale... To make the judgment, a person must compare an immediate impression with memory impression of similar stimuli....»

Thus relative measurement through comparative judgment is intrinsic to our thinking and should not be carried by us as an appendage whose real function is not understood well or at all and should be kept outside. It is not difficult to see that relative measurement predates and is necessary for creating and understanding absolute measurement. Some of the work reported on here is now well known. But we need it for the subsequent discussion that lays the foundation for relative measurement.

We know from the neurological sciences that sense data are mixed with temperature and other information by the thalamus, before they are recorded in memory. In the end what we sense is what we are, and not fully what is out there. Performance tests that I have conducted on numerous occasions with a diversity of audiences indicate that an individual not experienced in ranking objects according to size, when comparing one object that is very small, with another that is three times larger, would say that they are about the same size. This is particularly true when there are other sizeable objects in the collection. Only by being exposed to many objects and asked to make careful distinctions in size that the individual will begin to show an improved ability to sort and rank the objects according to size. What the person does is to adjust his sensation and impression with what he or she observes. It is not the real objects that one compares, but the impressions one forms about them. One needs such real experiences to institute early in one's mind the possibility of comparing things in pairs. This applies equally to more abstract ideas and their relative importance to a higher order property or goal. He would then be able to say that one idea is more important than another in terms of the satisfaction of the goal and whether, according to his or her understanding and experience, it is much more important or slightly more important. The lesser of the two is always used as the unit in terms of which the more important one is compared as to how much more important it is, and also how many times more, because the feeling of importance is converted to magnitudes on numerous sense experiences and thus there is transfer from the concrete to the abstract so that the two can be combined to make tradeoffs when needed, which happens frequently in daily experience. It is not possible to compare the lesser element with the greater one, because it must first be used as a unit to determine the magnitude of the greater one. Thus there is bias in human thinking in using the smaller of two elements as the unit. It is impossible a priori to ask how much less the smaller element is than the larger without first involving it as the unit of measurement. Thus, priorities of many objects can only be derived on the basis of dominance, and their reciprocal is automatically calculated to determine in a meaningful way, the relative priorities of being «dominated».

We have learned from many applications that wrong decisions may be made in some cases where only one structure is used for the purpose of generating priorities for the alternatives. In general one needs two or more of four separate structures: one for benefits, one for costs, one for opportunities and a fourth for risks. Because one must ask what dominates what in the paired comparisons, and by how much (homogeneous elements with clusters and pivots are used for widely spread alternatives), in the end one multiplies the benefits of each alternative by the opportunities it creates and divides by the costs times the risks.

2. NUMBERS ARE AS GOOD AS THE SCALES TO WHICH THEY BELONG

Numerical scales are our simplest way to express relations between things. There are, in addition to nominal scales which are invariant under one-to-one transformations and used to designate objects by assigning each of them a different name or symbol, four kinds of numerical scales that we use to deal with the world. These scales are, from weakest to strongest: ordinal, interval, ratio, and absolute. We need to say a few words about each. It is worth mentioning at this stage that when there are multiple criteria, it must be possible to combine the rankings with respect to the different criteria, and not every scale admits the use of the arithmetic operations (addition and multiplication) needed to do the combining. Furthermore, there are situations of interdependence among the alternatives that narrow the choice of scale further.

Ordinal Scales: Invariant under strictly monotone increasing transformations $x \geqslant y$ if and only if $f(x) \geqslant f(y)$.

When x is preferred to y, it is assigned any number greater than the number assigned to x (the same number only if x and y are the same). Assigning numbers that indicate order of preference among alternatives is a mapping into an ordinal scale. The only property that one wants to see preserved is monotonicity or simply, greater than. The larger (smaller) the number the higher (lower) the rank. For example, if apples are preferred to oranges and oranges are preferred to bananas, we have an infinite number of ways to assign numbers indicating this ranking. Below are four ways we might do it:

Apples	Oranges	Bananas
3	2	1
77	25	13
1,000,000	1,000	2
.6	.3	.1

Suppose we have one ordinal ranking on taste and have another on juiciness:

	Apples	Oranges	Bananas
Taste	24	17	2
Juiciness	3	2	1

Can we add the numbers for both taste and juiciness to determine which fruit is the most preferred? No. The magnitudes of the numbers are ambiguous, and we can get many different answers that do not lead to a unique outcome. Similarly, although it may not be so obvious, we cannot aggregate students' judgments on how much they enjoy a particular lecture by assigning a number from 1 to 4 and expect to get a meaningful outcome for judging the competence of the lecturer. Conclusion: We cannot use ordinals in ranking when many criteria are taken together to obtain a single overall ranking.

The next three scales are known as *cardinal scales* because the assigned values have meaning beyond a simple order.

Interval Scales: Invariant under positive linear transformations $ax + b$, $a > 0$.

Interval scales have an arbitrary origin and an arbitrary unit. The advantage of keeping the multiplier a positive is that if we then take the ratio of the difference of two readings on an interval scale to the difference of another two readings, we obtain a ratio scale which is defined next. It is important to know that one plus one is not always equal to two if the numbers belong to an interval scale. The temperature scale is an example of an interval scale with a choice as to what zero signifies. In the Celsius scale 0 indicates the freezing point and 100 degrees indicates the boiling point of water, which on the Fahrenheit scale have the respective values, 32 and 212. To establish the unit in each scale, one makes a mark at the 0 Celsius, 32 Fahrenheit level and another mark at the 100 Celsius, 212 Fahrenheit level, then divides the resulting range into 100 equal parts for the Celsius and 180 equal parts for the Fahrenheit. When we apply ordinary arithmetic to such commonplace scales of measurement, we find that some of the things we do are really illegitimate, because our arithmetic operations result in meaningless information. For example, if we measure temperature on an interval scale such as a Fahrenheit scale and if we add 20 degrees of Fahrenheit temperature to 30 degrees, we get 50, which is not 50 degrees Fahrenheit temperature - a much warmer temperature. It is meaningful to take the average of interval scale readings but not their sum. Thus $(ax_1 + b)+(ax_2 + b) = a(x_1 + x_2) + 2b$, which does not have the form $ax + b$. However, if we average by dividing by 2, we do get an interval scale value. We can also multiply interval scale readings by positive numbers whose sum is equal to 1 and add to get an interval scale result, a weighted average. One also cannot multiply numbers from an interval scale, because the result is not an interval scale. Thus $(ax_1 + b)(ax_2 + b) = a^2 x_1x_2 + ab(x_1 + x_2) + b^2$ which again does not have the form $ax + b$.

In decision making where relative measurement finds its best applications because of the need for judgments, interval scales can only be used to rank alternatives with respect to criteria but not to rank the criteria themselves. One cannot measure criteria or goals on an interval scale and then use them for weighting alternatives because one then obtains a product of two interval scales, which as we have seen is not meaningful. Note, however, that the use of interval scales demands the use of tangible «objective» scales to evaluate alternatives with respect to intangible criteria. When no such tangible numerical indicator exists, we must somehow find an absolute scale to define the range in which the alternatives for that criterion can spread, and then assess the given alternatives one by one as to where they fall in that range. In the absence of a numerical range indicator, such a ranking of the alternatives cannot be made. Conclusion: Interval scales cannot legitimately be considered for all our purposes – perhaps only for the alternatives as multicriteria utility people try to do by using ratio scales for the criteria. If there is feedback, this approach would not be valid because we would have to add and multiply numbers from interval scales.

We note that the *ratio* of differences between interval scale readings is meaningful if we have these readings – but how does one create them in the first place? One can take the ratio of the difference between apples and oranges to the difference between apples and bananas to decide how much more apples are preferred to oranges than to bananas. The question is, Can we make these comparisons more directly and more simply? We can, with a finer scale. Next we turn to ratio and absolute scales. These two are intimately related, as we shall see below.

Ratio Scales: Invariant under positive similarity transformations ax, $a > 0$.

Length, weight, time, and many other physical attributes can be measured on a ratio scale. Not only can one add and multiply numbers from the same ratio scale but one can also multiply numbers from two different ratio scales and still obtain a new ratio scale – something that physics does all the time. Ratios are important for gauging a response in proportion to a stimulus or an action in response to an idea or a belief. One cannot arbitrarily assign numbers to things and claim that they are from a ratio scale. One needs to be 100 % sure that the numbers used belong to a ratio scale. The question is – how?

The ratio measurements of objects on a ratio scale are absolute numbers. If one object has a ratio measurement of six pounds and another of two, their ratio is 6/2 = 3. The larger object is three times heavier than the smaller one, which is an absolute quantity, indicating the simple ratio 3 to 1, whether the measurement is in pounds, kilograms, or other units. All ratio scales can be reduced to a comparison of objects on an absolute scale in the ratio of x to 1, where the smaller (less dominant) object is 1 and the larger (more dominant) object is x times the smaller object, in which case the smaller object is $1/x$ as large as the larger object – a reciprocal relation in which the larger object is the unit.

For a ratio scale, we have $ax_1 + ax_2 = a(x_1 + x_2) = ax_3$, which belongs to the same ratio scale, and $ax_1 bx_2 = abx_1x_2 = cx_1x_2 = cx_3$ which belongs to a new ratio scale. However, $ax_1 + bx_2$ does not define a ratio scale, and thus we cannot add measurements from different ratio scales.

Absolute Scales: Invariant under the identity transformation $f(x) = x$ (is an identity).

There is no transformation on absolute numbers that leaves them invariant under some operation other than using their given values. In other words, an absolute number says what it says and it cannot be said in any other way. If there are five people in a room, there is no way to transform the number 5 through arithmetic operations and obtain another number that would meaningfully describe the number of people in the room. If one object is 5 times heavier than another object, there is no way it can be made different and still convey the idea that it is 5 times heavier. This is what is meant by invariance under the identity transformation. Examples of absolute scales are all collections of numbers that indicate magnitude (how many as used in counting) and frequency (how often). By abuse of thought, people often think of numbers measured on what we refer to as an «objective» scale, such as the interval scale of temperature or the ratio scale of weight, as absolute numbers.

It is interesting to note that when we count both men and sheep, to know the number of each in the group we must express it in relative terms, a ratio of the number of each kind to the total number of both kinds. We can then say, of the total, such and such percentage is of one kind and such and such percentage is of the other. Thus we use ratios to express the relative number of each kind of absolute number in terms of a larger absolute number that is the sum of the two and represents a higher order of generality. Conclusion: At the heart of dealing with a variety of things and grouping them together is the notion of proportionality formalized through ratio scales. In a sense, then, ratio scales are philosophically even more fundamental than absolute numbers when many things have to be combined for overall understanding, and that is what our mind does. Kuffler and Nichols [12] p. 57, write: «...the surprising conclusion is that the brain receives little information about the absolute level of uniform illumination...Signals arrive only from the cells with receptive fields situated close to the border...we perceive the difference or contrast at the boundary and it is by that standard that brightness in the uniformly illuminated central area is judged.»

What we need is a theory based on absolute measurement that does not require «objective» scales, which would make it possible not only to measure intangibles but also to combine multidimensional measurements into a unidimensional scale. It is obvious that we can do this only if the scales are relative, not absolute, and are amenable to arithmetic operations. It will be seen that, from the very idea of a relative scale, our scales must be ratio scales. It is on such a relative priority scale that the human mind (and other forms of existence) determines its degree of equilibrium on all the properties it subconsciously perceives at once. At each instant it takes in a variety of data from different dimensions and combines them into an overall assessment of the order and meaning that serve our survival needs at that particular moment. These observations are the basis for what follows.

A judgment or comparison is the numerical representation of a relationship between two elements that share a common parent.

3. THE PARADIGM CASE; CONSISTENCY

We will first show that when the judgments use measurements from a scale to form the ratios, the resulting matrix is consistent and deriving the scale is an elementary but fundamental operation. Later we generalize to the inconsistent case where the numerical values of the judgments are not taken from precise measurements but are ratios estimated according to knowledge and perception.

Let us assume that n activities are being considered by a group of interested people and that their tasks are

a) to provide judgments on the relative importance of these activities, and

b) to ensure that the judgments are quantified to an extent that permits a quantitative interpretation of the judgments among all activities.

Our goal is to describe a method for deriving, from these quantified judgments (i.e., from the relative values associated with pairs of activities), a set of weights to be associated with individual activities in order to put the information resulting from a and b into usable form.

Let $A_1, A_2, ..., A_n$ be the activities. The quantified judgments on pairs of activities (A_i, A_j) are represented by an n-by-n matrix

$$A = (a_{ij}),\ (i,j = 1, 2, ..., n).$$

The problem is to assign to the n activities $A_1, A_2, ..., A_n$ a set of numerical weights $w_1, w_2, ..., w_n$ that reflect the recorded quantified judgments.

First we get a simple question out of the way. The matrix A may have several, or only few, non-zero entries a_{ij}. Zeros are used when the judgment is unavailable. The question arises: how many entries are necessary to ensure the existence of a set of weights that is meaningful in the context of the problem? The answer is: it is sufficient that there be a set of entries that interconnects all activities in the sense that for every two indices *i, j,* there should be some *chain* of (positive) entries connecting i with j:

$$a_{ii_1},\ a_{i_1i_2},\ a_{i_2i_3},\ \ldots,\ a_{i_kj}\ ,$$

Note that a_{ij} itself is such a chain of length 1. (Such a matrix $A = (a_{ij})$ corresponds to a strongly connected graph.) This gives precise meaning to the formulation of task b.

One of the most important aspects of the AHP is that it allows us to measure the overall consistency of the judgments a_{ij}. An extreme example of inconsistent judgments is if we judge one activity to be more important than another and the second more important than the first, $a_{ij} > 1$ and $a_{ji} > 1$. More subtle is the case when the judgments of three alternatives are not «transitive». We might judge one stone two times as heavy as the first, a third stone twice as heavy as the second, but the first and last to be of equal weight. In that case $a_{ij} \neq a_{ik}a_{kj}$. This example leads us to the

Definition

$A = (a_{ij})$ is *consistent* if $a_{ij}a_{jk} = a_{ik}$, i, j, k = 1, ..., n (1)

We see that such a matrix can be constructed from a set of n elements which form a chain (or more generally, a spanning tree, a connected graph without cycles that includes all n elements for its vertices) across the rows and columns.

To interpret our first theorem let us consider the following case. An adult and a child are compared according to their height. If the adult is estimated to be two and a half times taller, that may be demonstrated by marking off several heights of the child end to end. However, if we have an absolute scale of measurement with the child measuring w_1 units and the adult w_2 units, then the comparison would assign the adult the relative value w_2/w_1 and the child w_1/w_2, the reciprocal value. These ratios yield the paired comparison values $(w_1/w_2)/1$ and $1/(w_2/w_1)$, respectively, in which the height of the child serves as the unit of comparison. Such a representation is valid only if w_1 and w_2 belong to a ratio scale so that the ratio w_1/w_2 is independent of the unit used, be it in inches or in centimeters, for example. In this way, we can interpret all ratios as absolute numbers or dominance units.

Let us now form the matrix W whose rows consist of the ratios of the measurements w_i of each of n items with respect to all others.

$$W = \begin{pmatrix} w_1/w_1 & w_1/w_2 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & \cdots & w_2/w_n \\ \vdots & \vdots & & \vdots \\ w_n/w_1 & w_n/w_2 & \cdots & w_n/w_n \end{pmatrix}$$

It is easy to prove the following theorem:

Theorem 1. *A positive n by n matrix has the ratio form $A = (w_i/w_j)$, i,j = 1,...,n, if, and only if, it is consistent.*

Corollary. If (1) is true then A is reciprocal $\left(a_{ji} = \dfrac{1}{a_{ij}}\right)$.

We observe that if W is the matrix above and w is the vector $w = (w_1, \ldots w_n)^T$ then $Ww = nw$. This suggests.

Theorem 2. *The matrix of ratios $A = (w_i/w_j)$ is consistent if and only if n is its principal eigenvalue and $Aw = nw$. Further, $w > 0$ is unique to within a multiplicative constant.*

Proof. The «if» part of the proof is clear. Now for the other half. If A is consistent then n and w are one of its eigenvalues and its corresponding eigenvector, respectively. Now A has rank one because every row is a constant multiple of the first row. Thus all its eigenvalues except one are equal to zero. The sum of the eigenvalues of a matrix is equal to its trace, the sum of the diagonal elements, and in this case, the trace of A is equal to n. Therefore, n is a simple eigenvalue of A. It is also the largest, or principal, eigenvalue of A. Alternatively, $A = Dee^TD^{-1}$ where D is a diagonal matrix with $d_{ii} = w_i$, and $e = (1,\ldots,1)^T$. Therefore, A and ee^T are similar and have the same eigenvalues [20]. The characteristic equation of ee^T is obviously $\lambda^n - n\lambda^{n-1} = 0$, and the result follows.

The solution w of $Aw = nw$, the principal right eigenvector of A, consists of positive entries and is obviously unique to within a positive multiplicative constant (a similarity transformation) thus defining a ratio scale. To ensure uniqueness, we normalize w by dividing by the sum of its entries. Given the comparison matrix A, we can directly recover w as the normalized version of any column of A; $A = wv$, $v = (1/w_1, \ldots, 1/w_n)$. It is interesting to note that for $A = (w_i/w_j)$, all the conclusions of the well-known theorem of Perron are valid without recourse to that theorem. Perron's theorem says that a matrix of positive entries has a simple positive real eigenvalue which dominates all other eigenvalues in modulus and a corresponding eigenvector whose entries are positive that is unique to within multiplication by a constant.

Here, we concern ourselves only with right eigenvectors because of the nature of dominance. In paired comparisons, the smaller element of a pair serves as the unit of comparison. There is no way of starting with the larger of a pair and decomposing it to determine what fraction of it the smaller is without first using the smaller one as a standard for the decomposition.

If A is consistent, then a_{ij} may be represented as a ratio from an existing ratio scale, such as the kilogram scale for weight. It may also be represented by using a judgment estimate as to how many times more the dominant member of the pair has a property for which no scale exists, such as smell or customer satisfaction. Of course, if the measurements from an actual scale are used in the pairwise comparisons, the derived scale of relative magnitudes is not a new scale – it is the same one used to do the measuring. We note that any finite set of n readings $w_1, ..., w_n$ from a ratio scale defines the principal eigenvector of a consistent n by n matrix $W = (w_i/w_j)$.

With regard to the order induced in w by W, in general, we would expect for an arbitrary positive matrix $A = (a_{ij})$, that if for some i and j, $a_{ik} \geqslant a_{jk}$ for all k, then $w_i \geqslant w_j$ should hold. But when A is inconsistent, i.e., it does not satisfy (1), what is an appropriate order condition to be satisfied by the a_{ij}, and how general can such a condition be? We now develop conditions for order preservation that are essentially observations on the behavior of a consistent matrix later generalized to the inconsistent case. The ratio $(w_i/w_j)/1$ may be interpreted as assigning the i^{th} activity the unit value of a scale and the j^{th} activity the absolute value w_j/w_i. In the consistent case, order relations on w_i, $i = 1, ..., n$, can be inferred from the a_{ij} as follows: we factor out w_1 from the first row, w_2 from the second and so on, leaving us with a matrix of identical rows and $w_i \geqslant w_j$ is both necessary and sufficient for $A \leftrightarrow w$.

C. Berge [2] reports on a proposal by T. H. Wei [21] on the measurement of dominance or power of a player in a tournament through a pairwise comparison matrix $B = (b_{ij})$. Each row of B defines the standing of one player relative to the other players in the tournament. We have:

$$b_{ij} = \begin{cases} 0 \text{ if } i \text{ loses to } j \\ 1 \text{ if } i \text{ ties } j \text{ (in particular } b_{ii} = 1) \\ 2 \text{ if } i \text{ wins over } j \end{cases}$$

and thus $b_{ij} + b_{ji} = 2$. The overall power of each player i is *defined* as the i^{th} component of $\lim_{k\to\infty} B^k e/e^T B^k e$, where B^k is the k^{th} power of B. It coincides with a constant multiple of the i^{th} component of the solution of $Bw = \lambda_{\max} w$ where $\lambda_{\max}$ is the principal eigenvalue of the matrix B. From a set of arbitrary nonnegative numbers one obtains a ratio scale w. But under what conditions is the solution relevant to the b_{ij}?

There is a canon about order relations in A and correspondingly in w when A is consistent that we need to observe when A is inconsistent. We begin with a consistent matrix A. By successive application of the consistency condition (1) to each factor on the left of the condition itself, we obtain:

$$A = (1/n)\, A^2 = ... = (1/n)^{k-1} A^k = ...$$

and in normalized form

$$\frac{A}{e^T A e} = \frac{A^2}{e^T A^2 e} = \cdots = \frac{A^k}{e^T A^k e} = \cdots \qquad (2)$$

which shows that every power of A must be considered in the preservation of consistency. When A is consistent, the consistency condition (1) can be stated in equivalent terms for an arbitrary power of A. This is a useful observation for developing an order condition to be satisfied when A is inconsistent. Here the power of A gives different measurements of dominance due to intransitivity. The normalized sum of the rows of A give dominance in paths of length one; those of A^2 in paths of length two and so on. If we define a sequence of successive series of these vectors, then its limit is the principal right eigenvector.

Five Conditions on A For Preserving Order

A weaker condition for order preservation than

(i) $$(A)_i \geqslant (A)_j \qquad \text{implies } w_i \geqslant w_j$$

is

(ii) $$(Ae)_i \geqslant (Ae)_j \qquad \text{implies } w_i \geqslant w_j$$

where $(A)_i$ and $(Ae)_i$ denote the ith row and ith row sum of A, and its generalization to powers of A given in the normalized form:

(iii) $$\frac{(A^m e)_i}{e^T A^m e} \geqslant \frac{(A^m e)_j}{e^T A^m e} \qquad \text{implies } w_i \geqslant w_j$$

The condition for order preservation must include all powers of A, and is given here in terms of their sum. For sufficiently large integer $N > 0$, and for $p \geqslant N$,

(iv) $$\sum_{m=1}^{p} \frac{(A^m e)_i}{e^T A^m e} \geqslant \sum_{m=1}^{p} \frac{(A^m e)_j}{e^T A^m e} \qquad \text{implies } w_i \geqslant w_j$$

and by (2):

(v) $$\lim_{p\to\infty} \frac{1}{p} \sum_{m=1}^{p} \frac{(A^m e)_i}{e^T A^m e} \geqslant \lim_{p\to\infty} \frac{1}{p} \sum_{m=1}^{p} \frac{(A^m e)_j}{e^T A^m e} \qquad \text{implies } w_i \geqslant w_j$$

Theorem 3. *If A is consistent, then*

$$\lim_{p\to\infty} \frac{1}{p} \sum_{m=1}^{p} \frac{(A^m e)_i}{e^T A^m e} \to c w_i > 0$$

and (i)-(v) are true.

Proof. Follows from $A^m = n^{m-1} A$ where n is the principal eigenvalue of A, and $A = (w_i/w_j)$.

It appears that the problem of constructing ratio scales from a_{ij} has a natural principal eigenvalue structure. Our task is to extend this formulation to the case where A is no longer consistent.

4. SMALL PERTURBATIONS AND RATIO SCALE APPROXIMATION

Because we are interested in the construction of an appropriate matrix W of ratios that serves as a «good» approximation to a given reciprocal matrix A, we begin by assuming that A itself is a perturbation of W. We need the following kind of background information.

For an unrepeated eigenvalue of a positive matrix A it is known [11,19,22] that a small perturbation $A(\varepsilon)$ of A gives rise to a perturbation $\lambda(\varepsilon)$ that is analytic in the neighborhood of $\varepsilon = 0$ and small because $A(\varepsilon)$ is reciprocal. The following known theorems give us a part of what we need.

Theorem 4. *(Existence): If λ is a simple eigenvalue of A, then for small $\varepsilon > 0$, there is an eigenvalue $\lambda(\varepsilon)$ of $A(\varepsilon)$ with power series expansion in ε:*

$$\lambda(\varepsilon) = \lambda + \varepsilon\lambda^{(1)} + \varepsilon^2\lambda^{(2)} + \dots$$

and corresponding right and left eigenvectors $w(\varepsilon)$ and $v(\varepsilon)$ such that

$$w(\varepsilon) = w + \varepsilon w^{(1)} + \varepsilon^2 w^{(2)} + \dots$$
$$v(\varepsilon) = v + \varepsilon v^{(1)} + \varepsilon^2 v^{(2)} + \dots$$

Let Θ_{ij} be a perturbation of a reciprocal matrix A such that $B = (a_{ij} + \Theta_{ij})$ is also positive [7].

Theorem 5. *If a positive reciprocal matrix A has the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ where the multiplicity of λ_j is m_j with $\sum_{j=1}^{n} m_j = n$, then given $\varepsilon > 0$ there is a $\delta(\varepsilon) > 0$ such that if $|a_{ij} + \Theta_{ij} - a_{ij}| \leqslant \delta$ for all i and j the matrix B has exactly m_j eigenvalues in the circle $|\mu_j - \lambda_j| < \varepsilon$ for each j $= 1, \dots, s$ where $\mu_1, \dots, \mu_s$ are the eigenvalues of B.*

If A is a consistent matrix, then it has one positive eigenvalue $\lambda_1 = n$ and all other eigenvalues are zero. For a suitable $\varepsilon > 0$ there is a $\delta(\varepsilon) > 0$ such that for $|\Theta_{ij}| < \varepsilon$ the perturbed matrix B has one eigenvalue in the circle $|\mu_1 - n| < \varepsilon$ and the remaining eigenvalues fall in a circle $|\mu_j - 0| < \varepsilon$, $j = 2, \dots, n$.

Theorem 6. *If n is a simple eigenvalue of A which dominates the remaining eigenvalues in modulus, for sufficiently small ε, $n(\varepsilon) \equiv \lambda_{\max}$ dominates the remaining eigenvalues of $A(a_{ij}(\varepsilon))$ in modulus.*

When A is inconsistent, several conditions on a_{ij} and on w_i, along with uniqueness, must be met to enable us to approximate A by ratios. Our conditions are divided into two categories. One category deals with the order induced by a_{ij} as absolute numbers $(w_i/w_j)/1$ or $1/(w_j/w_i)$ from a standard scale, on the components of the scale w. The other category deals with the equality or near equality of the a_{ij} to the ratios w_i/w_j formed from the derived scale w.

When A is inconsistent, how do we construct W so that the order preservation condition (v) still holds? Later we address the other question; what conditions must A satisfy to ensure that w_i/w_j is a «good» approximation to a_{ij}?

Let us consider estimates of ratios given by an expert who may make small perturbations ε_{ij} in $W = (w_i/w_j)$. Comparisons by ratios allow us to write $a_{ij} = (w_i/w_j)\,\varepsilon_{ij}$, $\varepsilon_{ij} > 0$, $i, j = 1, \dots, n$. In that case, A takes the form $A = W \circ E = DED^{-1}$ where $W = (w_i/w_j)$, $E = (\varepsilon_{ij})$, D a diagonal matrix with w as diagonal vector, and o refers to the Hadamard or elementwise product of the two matrices. The principal eigenvalue of A coincides with that of E. The principal eigenvector of A is the elementwise product of the principal eigenvectors $w = (w_1, \dots, w_n)^T$, and $e = (1, \dots, 1)^T$ of W and of E respectively [20].

The distinction we make between an arbitrary positive matrix and a reciprocal matrix is that we can control a step by step modification of a reciprocal matrix so that in the representation $A = WoE = DED^{-1}$, the ε_{ij}, $i, j = 1, \dots, n$ are small. The purpose is to ensure that perturbing the principal eigenvalue and eigenvector of W yields the principal eigenvalue and eigenvector of A.

Why do we need such a perturbation? Because we assume that there is an underlying ratio scale that we attempt to approximate. By improving the consistency of the matrix, we obtain an approximation of the underlying scale by the principal eigenvector of the resulting matrix.

Theorem 7. *w is the principal eigenvector of a positive matrix A if, and only if, $Ee = \lambda_{\max} e$.*

Note that e is the principal eigenvector of E and E *is* a perturbation of the matrix $e^T e$. When $Ee \neq \lambda_{\max} e$ the principal eigenvector of A is another vector $w' \neq w$ and $A = W'oE'$ where $E'e = \lambda_{\max} e$.

Corollary. w is the principal eigenvector of a positive reciprocal matrix $A = WoE$, if and only if, $Ee = \lambda_{\max} e$ and $\varepsilon_{ji} = (e_{ij})^{-1}$.

Assume that A is an arbitrary positive matrix that is a small perturbation E of $W = (w_i/w_j)$. Then we have

Theorem 8. *(Order preservation): A positive matrix A satisfies condition (v), if and only if, the derived scale w is the principal eigenvector of A, i.e.,* $Aw = \lambda_{\max} w$.

Proof. We give two proofs of this theorem, the first is based on the well known theorem of Perron and the second, which is more appropriate for our purpose is based on perturbation.

Let

$$s_k = \frac{A^k e}{e^T A^k e} \tag{3}$$

and

$$t_m = \frac{1}{m} \sum_{k=1}^{m} s_k \tag{4}$$

The convergence of the components of t_m to the same limit as the components of s_m is the standard Cesaro summability. Since,

$$s_k = \frac{A^k e}{e^T A^k e} \rightarrow w \text{ as } k \rightarrow \infty \tag{5}$$

where w is the normalized principal right eigenvector of A, we have

$$t_m = \frac{1}{m} \sum_{k=1}^{m} \frac{A^k e}{e^T A^k e} \rightarrow w \text{ as } m \rightarrow \infty \tag{6}$$

For the second proof, first assume that A has only simple eigenvalues. Using Sylvester's formula:

$$f(A) = \sum_{i=1}^{n} f(\lambda_i) \frac{\prod_{j \neq i} (A - \lambda_i I)}{\prod_{j \neq i} (\lambda_j - \lambda_i)}, \quad \lambda_{\max} = \lambda_1$$

we have on writing $f(A) = A^k$, dividing through by $\lambda^k_{\max}$, multiplying on the left by $(A - \lambda_{\max} I)$ to obtain the characteristic polynomial of A then multiplying on the right by e we obtain:

$$\lim_{k \rightarrow \infty} \frac{A^k e}{\lambda^k_{\max}} = cw, \textit{ for some constant } c > 0$$

Sylvester's formula for multiple eigenvalues of multiplicity m_i shows that one must consider derivatives of $f(\lambda)$ of order no more than m_i. However, it is easy to verify by interchanging derivative and limit, that when each term is divided by $\lambda^k{}_{\max}$ its value tends to zero as $k \rightarrow \infty$, and the result again follows.

Therefore, it is necessary to obtain the principal eigenvector w to capture order properties from A, but not sufficient to ensure that $W = (w_i/w_j)$ is a good approximation to A. The method we use to derive the scale w from a positive inconsistent matrix must also satisfy the following conditions on what constitutes a good numerical approximation to the a_{ij} by ratios. The first two are local conditions on each a_{ij}, the second two are global conditions on all a_{ij} through the principal eigenvalue and eigenvector as functions of the a_{ij}.

Four Conditions for Good Approximations

1. *Reciprocity*

The reciprocal condition is a local relation between pairs of elements: $a_{ji} = 1/a_{ij}$, needed to ensure that, as perturbations of ratios, a_{ij} and a_{ji} can be approximated by ratios from a ratio scale that are themselves reciprocal. It is a necessary condition for consistency.

2. *Homogeneity* – Uniformly Bounded Above and Below

Homogeneity is also a local condition on each a_{ij}. To ensure consistency in the paired comparisons, the elements must be of the same order of magnitude which means that our perceptions in comparing them, should be of nearly the same order of magnitude. Thus we require that the a_{ij} be uniformly bounded above by a positive constant K and, because of the reciprocal condition, they are automatically uniformly bounded below away from zero:

$$1/K \leqslant a_{ij} \leqslant K, \quad K > 0, i, j = 1, ..., n$$

It is a fact that people are unable to directly compare widely disparate objects such as an apple and a watermelon according to weight. If they are not comparable, it is possible to aggregate them in such homogeneous clusters to make the comparisons by introducing hypothetical elements of gradually increasing or decreasing sizes with which they can be compared.

For example (Figure 1), to compare an unripe cherry tomato with a watermelon, we compare it with a small green tomato and a lime in one cluster, then compare the lime with a grapefruit and a honeydew melon in a second cluster, and finally compare the honeydew melon with a sugar baby watermelon and an oblong watermelon in a thied cluster. The relative measurements in the clusters can be combined because we included the largest element (the cantaloupe) in the small cluster as the smallest element of the adjacent larger cluster. Then the relative weights of the elements in the second cluster are all divided by the relative weight of the common element and multiplied by its relative weight in the smaller cluster. In this manner, relative measurement of the elements in the two clusters can be related and the two clusters combined after obtaining relative measurement by paired

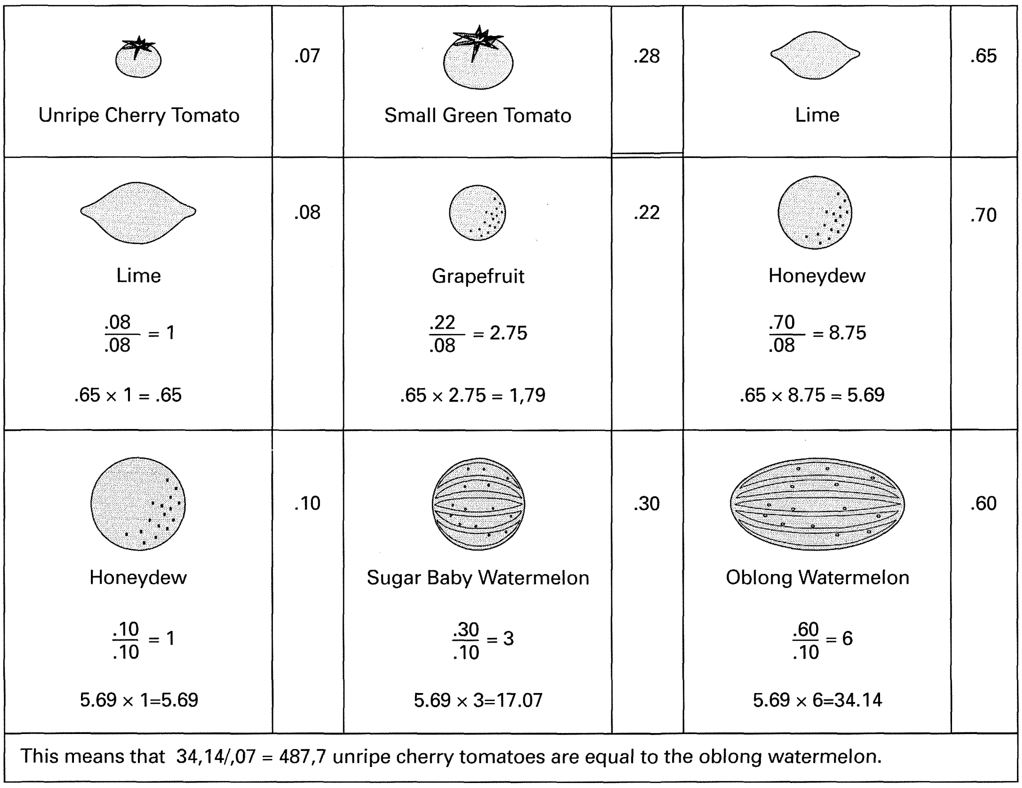

Figura 1. Clustering to Compare the very Small with the very Large.

comparisons in each cluster. The process is continued from cluster to adjacent cluster. Here we see that in the end more than 487 cherry tomatoes make up a watermelon. This kind of clustering has to be done with respect to each criterion. For example, instead of size we could have used relative greenness for clustering and comparisons.

3. *Near Consistency*

The near consistency condition which is global, is formed in terms of the (structural parameters) $\lambda_{\max}$ and n of A and W. It is a less familiar and more intricate condition that we need to discuss at some length. The requirement that comparisons be carried out on homogeneous elements ensures that the coefficients in the comparison matrix are not too large and generally of the same order of magnitude, i.e., from *1* to *9*. Knowing this constrains the size of the perturbations ε_{ij}, whose sum as we shall see below, is measured in terms of the near consistency condition $\lambda_{\max}-n$.

The object then is to apply this condition to develop algorithms to explore changing the judgments and their approximation by successively decreasing the inconsistency of the judgments and then approximating them with ratios from the derived scale. The simplest such algorithm is one which identifies that a_{ij} for which $a_{ij}w_j/w_i$ is maximum and indicates decreasing it in the direction of w_i/w_j. Another algorithm due to Harker [9] utilizes the gradient of the a_{ij}. In the end, we obtain either a consistent matrix or a closer approximation to a consistent one depending on whether the information available allows for making the proposed revisions in a_{ij}.

Because consistency is necessary and sufficient for A to have the form $A = (w_i/w_j)$, we use w to explore possible changes in a_{ij} to modify A «closer» to that form. We form a consistent matrix $W' = (w'_i/w'_j)$, whose elements are approximations to the corresponding elements of A. We have $a_{ij} = (w'_i/w'_j)\ \varepsilon_{ij}$, $\varepsilon_{ij} > 0$.What we have to deal with is the converse of: given a problem, find a good approximation to its solution. It is, given a problem with its exact solution, use the properties of this solution to revise the problem, i.e. the judgments which give rise to a_{ij}. Repeat the process to a level of admissible consistency. (see below)

4. *Uniform Continuity*

Uniform continuity implies that w_i, $i = 1, ..., n$ as a function of a_{ij} should be relatively insensitive to small changes in the a_{ij} in order that the ratios w_i/w_j remain good approximations to the a_{ij}. For example, it holds in

w_i as the ith component of the principal eigenvector because it is an algebraic function of $\lambda_{\max}$ (whose value is shown to lie near n because of 3), and of the a_{ij} and $1/a_{ij}$, which are bounded.

Let us now turn to more elaboration of the near consistency condition in 3). We first show the interesting result, that inconsistency or violation of (1) by various a_{ij} can be captured by a single number $\lambda_{max}-n$, which measures the deviation of all a_{ij} from w_i/w_j.

Assume that the reciprocal condition $a_{ji} = 1/a_{ij}$ and boundedness $1/K \leqslant a_{ji} \leqslant K$, where $K > 0$ is a constant, hold. Let $a_{ij} = (1 + \delta_{ij})\, w_i/w_j$, $\delta_{ij} > -1$, be a perturbation of $W = (w_i/w_j)$, where w is the principal eigenvector of A.

Theorem 9. $\lambda_{\max} \geqslant n$.

Proof. Using $a_{ji} = 1/a_{ij}$, and $Aw = \lambda_{\max}w$, we have

$$\lambda_{\max} - n = \frac{1}{n} \sum_{1 \leqslant i \leqslant j \leqslant n} \frac{\delta_{ij}^2}{1 + \delta_{ij}} \geqslant 0 \tag{7}$$

Theorem 10. *A is consistent if, and only if,* $\lambda_{\max} = n$.

Proof. If A is a consistent, then because of (1), each row of A is a constant multiple of a given row. This implies that the rank of A is one, and all but one of its eigenvalues λ_i, $i = 1, ..., n$, are zero. However, it follows from our earlier argument that, $\sum_{i=1}^{n} \lambda_i = Trace(A) = n$. Therefore $\lambda_{\max} = n$. Conversely, $\lambda_{\max} = n$, implies $\varepsilon_{ij} = 0$, and $a_{ij} = w_i/w_j$.

From (2) we can determine the magnitude of the «greatest» perturbation by setting one of the terms equal to $\lambda_{\max}-n$ and solving for δ_{ij} in the resulting quadratic. An average perturbation value is obtained by replacing $\lambda_{\max}-n$ in the previous result by $(\lambda_{\max}-n)/(n - 1)$.

A measure of inconsistency is obtained by taking the ratio of $\lambda_{\max}-n$ to its average value over a large number of reciprocal matrices of the same order n, whose entries are randomly chosen in the interval $[1/K, K]$. If this ratio is small (e.g., 10 % or less –for example 5 % for 3 by 3 matrices) [8, 15, 16], we accept the estimate of w. Otherwise, we attempt to improve consistency and derive a new w. After each iteration, we assume that the new matrix is a perturbation of W and its eigenvalue and eigenvector are perturbations of n and w, respectively.

In his experimental work in the 1950's, the psychologist George Miller [14], found that in general, people (such as chess experts looking ahead a few moves to decide on a good next move) could deal with information involving simultaneously only a few facts: seven plus or minus two. With more, they become confused and cannot handle the information. Since the individual needs to maintain consistency in his decision matrix, he cannot consider more than a few options at a time. This is in harmony with the established fact that for a reciprocal matrix (though not in general) the principal eigenvalue is stable for small perturbations when n is small.

We have seen that only order preserving derived scales w are of interest. There are many ways to obtain w from A. Most of them are error minimizing procedures such as the method of least squares:

$$\sum_{i,j=1}^{n} \left(a_{ij} - \frac{w_i}{w_j}\right)^2 \tag{8}$$

which also produces nonunique answers. Only the principal eigenvector satisfies order preserving requirements when there is inconsistency. We summarize with:

Theorem 11. *If a positive n by n matrix A is: reciprocal, homogeneous, and near consistent, then the scale w derived from* $Aw = \lambda_{\max}w$ *is order preserving, unique to within a similarity transformation and uniformly continuous in the* a_{ij}, $i, j, = 1, ..., n$.

Similar results can be obtained when A is nonnegative. Also we have extended this discrete approximation of A by W to the continuous case of A reciprocal kernel and its eigenfunction [17,18].

5. SOME STRUCTURAL PROPERTIES OF POSITIVE RECIPROCAL MATRICES

We make the following observations on the structure of reciprocal matrices. The elementwise product of two n by n reciprocal matrices is a reciprocal matrix. It follows that the set of reciprocal matrices is closed under the operation Hadamard product. The matrix e^Te is the identity: $e^Te = e^Teoe^Te = ee^T$ and A^T is the inverse of A, $AoA^T = A^ToA = e^Te$. Thus the set G of n by n reciprocal matrices is an abelian group. Because every subgroup of an abelian group is normal, in particular, the set of n by n consistent matrices is a normal subgroup ($EoWoE^T = W$) of the group of positive reciprocal matrices.

Two matrices A and B are *R-equivalent* ($A\ R\ B$) if, and only if, there are a vector w and positive constants a and b such that $(1/a)\, Aw = (1/b)\, Bw$. The set of all consistent matrices can be partitioned into disjoint equivalence classes. Given a consistent matrix W and a perturbation matrix E such that $Ee = ae$, $a > 0$ a constant, we use the Hadamard product to define $A' = WoE$ such that $(1/a)\, A'w = (1/n)$ Ww. A' and W are R-equivalent. There is a 1-1 correspondence between the set of all consistent matrices and the set of all matrices A' defined by such Hadamard products. An R-equivalence class $Q(W)$ is the set of all A' such that $A'R\ W$. The set of equivalence classes $Q(W)$ forms a partition of the set of reciprocal matrices. It is known that all the elements in $Q(W)$ are connected by perturbations $E, E', E'', \ldots$, corresponding to a fixed value of $a > 0$ such that $(EoE'oE''..)e = ae$. Thus given an arbitrary reciprocal matrix A, there exists an equivalence class to which A belongs.

DeTurck [6] has proved that: The structure group G of the set of positive reciprocal $n \times n$ matrices has $2n!$ connected components. It consists of nonnegative matrices which have exactly one nonzero entry in each row and column. These matrices can be written as $D \cdot S$, where D is a diagonal matrix with positive diagonal entries and S is a permutation matrix, and the negatives of such matrices. The connected component G_0 of the identity consists of diagonal matrices with positive entries on the diagonal. If A is a positive reciprocal matrix with principal right eigenvector $w = (w_1, w_2, \ldots, w_n)^T$ and $D\varepsilon G_0$ is a diagonal matrix with positive diagonal entries $d_1, d_2, \ldots d_n$ then $I_D(A) = DAD^{-1}$ is a positive reciprocal matrix with principal eigenvector $w' = (d_1 w_1, \ldots, d_n w_n)^T$. The principal eigenvalue is the same for both matrices. If $v = (v_1, \ldots, v_n)^T$ and $w = (w_1, \ldots, w_n)^T$ are two positive column vectors, then conjugation by the diagonal matrix D_{vw} with entries $v_1/w_1, \ldots, v_n/w_n$ on the diagonal maps A_w onto A_v. The corresponding diagonal matrix D_{wv} provides the inverse map. Moreover, D_{wv} maps the consistent matrix of A_w to the consistent matrix of A_v.

Roger A. Horn (personal communication) proved the following:

Theorem 12. Let $v_1, \ldots, v_n$ be given positive numbers, and set $v = [v_1, \ldots, v_n]^T$, $w = [1/v_1, \ldots, 1/v_n]^T$. Suppose $X = [x_1, \ldots, x_n]^T$ and $Y = [y_1^T, \ldots, y_n^T]^T$ are n-by-n nonnegative matrices such that $XY = vw^T - I_n$. Then $Z = YX + I_n = [z_{ij}]$ is a positive reciprocal matrix, that is, $z_{ij} = 1/z_{ji}$ for all $i, j = 1, \ldots, n$.

Proof. $0 = trace\ XY = trace\ YX = \sum_{i=1}^{n} y_i^T x_i \Rightarrow y_i^T x_i$ *for all i.*

For $i \neq j$ $I_n + x_i y_j^T + x_j y_i^T = \sum_{k \neq i,j} (-x_k) y_k^T$ is a sum of at most $n-1$ rank-1 matrices, so has rank at most $n-1$ and is singular. Thus, $0 = \det(I_n + x_i y_i^T + x_j y_j^T) = \det(I_n + [x_i x_j][y_i^T y_j^T]) = \det(I_2 + [y_i^T y_j^T]^T [x_i x_j]) = 1 - (y_i^T x_j)(y_j^T x_i) = 1 - z_{ij} z_{ji} = 0$.

6. DECISION MAKING-HOW TO APPLY RELATIVE MEASUREMENT

An often crucial disadvantage of many traditional decision-making methods is that they require specialized expertise to design the appropriate structure and then to embed the decision-making process in it. A decision-making approach should have the following characteristics:

- be simple in construct,
- be capable of dealing with risk and opportunity under uncertainty (in the AHP we deal with separate hierarchies for benefits, costs, risks and opportunities, and combine the outcomes for the alternatives from each thus obtaining an overall synthesis for the most preferred alternative),
- be adaptable to both groups and individuals,
- be natural to our intuition and general thinking,
- encourage compromise and consensus building, and
- not require inordinate specialization to master and communicate.

In addition, the details of the processes leading up to the decision-making process should be easy to review.

At the core of the problems that our method addresses is the need to assess the benefits, the costs, the risks and the opportunities of the proposed solutions. We must answer such questions as the following: Which consequences weigh more heavily than others? Which aims are more important than others? What is likely to take place? What should we plan for and how do we bring it about? These and other questions demand a multicriteria logic. It has been demonstrated over and over by practitioners who use the theory discussed in this paper, that multicriteria logic gives different and often better answers to these questions than ordinary logic and does it efficiently. The main reason is that in this logic we are able to include numerical intensities between the elements considered and can work to combine the micro and the macro in stages obtaining an overall synthesis.

To make a decision one needs various kinds of knowledge, information, and technical data. These concern

- details about the problem for which a decision is needed,
- the people or actors involved,
- their objectives and policies,
- the influences affecting the outcomes, and
- the time horizons, scenarios, and constraints.

The set of potential outcomes or alternatives from which to choose are the essence of decision making. In laying out the framework for making a decision, one needs to sort the elements into groupings or clusters that have similar influences or effects. One must also arrange them in some rational order to trace the outcome of these influences. Briefly, we see decision making as a process that involves the following steps:

1. Structure a problem with a model that shows the problem's key elements and their relationships.

2. Elicit judgments that reflect knowledge, feelings, or emotions.

3. Represent those judgments with meaningful numbers.

4. Use these numbers to calculate the priorities of the elements of the hierarchy.

5. Synthesize these results to determine an overall outcome.

6. Analyze sensitivity to changes in judgment.

The analytic hierarchy process (AHP), the decision-making process described in this paper, meets these criteria. It is about breaking a problem down and then aggregating the solutions of all the subproblems into a conclusion. It facilitates decision making by organizing perceptions, feelings, judgments, and memories into a framework that exhibits the forces that influence a decision. In the simple and most common case, the forces are arranged from the more general and less controllable to the more specific and controllable. The AHP is based on the innate human ability to make sound judgments about small problems. It has been applied in a variety of decisions and planning projects in nearly 20 countries.

Here *rationality* is defined to be:

- Focusing on the goal of solving the problem;

- Knowing enough about a problem to develop a complete structure of relations and influences;

- Having enough knowledge and experience and access to the knowledge and experience of others to assess the priority of influence and dominance (importance, preference, or likelihood to the goal as appropriate) among the relations in the structure;

- Allowing for differences in opinion with an ability to develop a best compromise.

How to Structure a Hierarchy

Perhaps the most creative part of decision making that has a significant effect on the outcome is modeling the problem. In the AHP, a problem is structured as a hierarchy. This is then followed by a process of prioritization, which we describe in detail later. Prioritization involves eliciting judgments in response to questions about the dominance of one element over another when compared with respect to a property. The basic principle to follow in creating this structure is always to see if one can answer the following question: Can I compare the elements on a lower level using some or all of the elements on the next higher level as criteria or attributes of the lower level elements?

A useful way to proceed in structuring a decision is to come down from the goal as far as one can by decomposing it into the most general and most easily controlled factors. One can then go up from the alternatives beginning with the simplest subcriteria that they must satisfy and aggregating the subcriteria into generic higher level criteria until the levels of the two processes are linked in such a way as to make comparison possible.

Here are some suggestions for an elaborate design of a hierarchy: (1) Identify the overall goal. What are you trying to accomplish? What is the main question? (2) Identify the subgoals of the overall goal. If relevant, identify time horizons that affect the decision. (3) Identify criteria that must be satisfied to fulfill the subgoals of the overall goal. (4) Identify subcriteria under each criterion. Note that criteria or subcriteria may be specified in terms of ranges of values of parameters or in terms of verbal intensities such as high, medium, low. (5) Identify the actors involved. (6) Identify the actors' goals. (7) Identify the actors' policies. (8) Identify options or outcomes. (9) For yes-no decisions, take the most preferred outcome and compare the benefits and costs of making the decision with those of not making it. (10) Do a benefit/cost analysis using marginal values. Because we are dealing with dominance hierarchies, ask which alternative yields the greatest benefit; for costs, which alternative costs the most, and for risks, which alternative is more risky. We now illustrate the process with an example in which opportunity is combined with other benefits and risk is combined with other costs. We have numerous examples in which they are treated separately particularly when a new system is being designed. In other words only when the complexity requires separate considerations that one uses different structures for each.

An Example-The Hospice Problem

Westmoreland County Hospital in Western Pennsylvania, like hospitals in many other counties around the nation, has been concerned with the costs of the facilities and manpower involved in taking care of terminally ill patients. Normally these patients do not need as much medical attention as do other patients. Those who best utilize the limited resources in a hospital are patients who require the medical attention of its specialists and advanced technology equipment – whose utilization depends on the demand of patients admitted into the hospital. The terminally ill need medical attention only episodically. Most of the time such patients need psychological support. Such support is best given by the patient's family, whose members are able to supply the love and care the patients most need. For the mental health of the patient, home therapy is a benefit. From the medical standpoint, especially during a crisis, the hospital provides a greater benefit. Most patients need the help of medical professionals only during a crisis. Some will also need equipment and surgery. The planning association of the hospital wanted to develop alternatives and to choose the best one considering various criteria from the standpoint of the patient, the hospital, the community, and society at large. In this problem, we need to consider the costs and benefits of the decision. Cost includes economic costs and all sorts of intangibles, such as inconvenience and pain. Such disbenefits are not directly related to benefits as their mathematical inverses, because patients infinitely prefer the benefits of good health to these intangible disbenefits. To study the problem, one needs to deal with benefits and with costs separately.

Approaching the Problem

I met with representatives of the planning association for several hours to decide on the best alternative. To make a decision by considering benefits and costs, one must first answer the question: In this problem, do the benefits justify the costs? If they do, then either the benefits are so much more important than the costs that the decision is based simply on benefits, or the two are so close in value that both the benefits and the costs should be considered. Then we use two hierarchies for the purpose and make the choice by forming ratios of the priorities of the alternatives (benefits b_i/costs c_i) from them. One asks which is most beneficial in the benefits hierarchy (Figure 2) and which is most costly in the costs hierarchy (Figure 3). If the benefits do not justify the costs, the costs alone determine the best alternative – that which is the least costly. In this example, we decided that both benefits and costs had to be considered in separate hierarchies. In a risk problem, a third hierarchy is used to determine the most desired alternative with respect to all three: benefits, costs, and risks. In this problem, we assumed risk to be the same for all contingencies. Whereas for most decisions one uses only a single hierarchy, we constructed two hierarchies for the hospice problem, one for benefits or gains (which model of hospice care yields the greater benefit) and one for costs or pains (which model costs more).

The planning association thought the concepts of benefits and costs were too general to enable it to make a decision. Thus, the planners and I further subdivided each (benefits and costs) into detailed subcriteria to enable the group to develop alternatives and to evaluate the finer distinctions the members perceived between the three alternatives. The alternatives were to care for terminally ill patients at the hospital, at home, or partly at the hospital and partly at home.

For each of the two hierarchies, benefits and costs, the goal clearly had to be choosing the best hospice. We placed this goal at the top of each hierarchy. Then the group discussed and identified overall criteria for each hierarchy; these criteria need not be the same for the benefits as for the costs.

The two hierarchies are fairly clear and straightforward in their description. They descend from the more general criteria in the second level to secondary subcriteria in the third level and then to tertiary subcriteria in the fourth level on to the alternatives at the bottom or fifth level.

At the general criteria level, each of the hierarchies, benefits or costs, involved three major interests. The decision should benefit the recipient, the institution, and society as a whole, and their relative importance is the prime determinant as to which outcome is more likely to be preferred. We located these three elements on the second level of the benefits hierarchy. As the decision would benefit each party differently and the importance of the benefits to each recipient affects the outcome, the group thought that it was important to specify the types of benefit for the recipient and the institution. Recipients want physical, psycho-social and economic benefits, while the institution wants only psychosocial and economic benefits. We located these benefits in the third level of the hierarchy. Each of these in turn needed further decomposition into specific items in terms of which the decision alternatives could be evaluated. For example, while the recipient measures economic benefits in terms of reduced costs and improved productivity, the institution needed the more specific measurements of reduced length of stay, better utilization of resources, and increased financial support from the community. There was no reason to decompose the societal benefits into a third level subcriteria, hence societal benefits connects directly to the fourth level. The group considered three models for the decision alternatives, and located them on the bottom or fifth level of the hierarchy: In Model 1, the hospital provided full care to the patients; In Model 2, the family cares for the patient at home, and the hospital provides only emergency treatment (no nurses go to the house); and in Model 3, the hospital and the home share patient care (with visiting nurses going to the home).

In the costs hierarchy there were also three major interests in the second level that would incur costs or pains: community, institution, and society. In this decision the costs incurred by the patient were not included as a separate factor. Patient and family could be thought of as part of the community. We thought decomposition was necessary only for institutional costs. We included five such costs in the third level: capital costs, operating

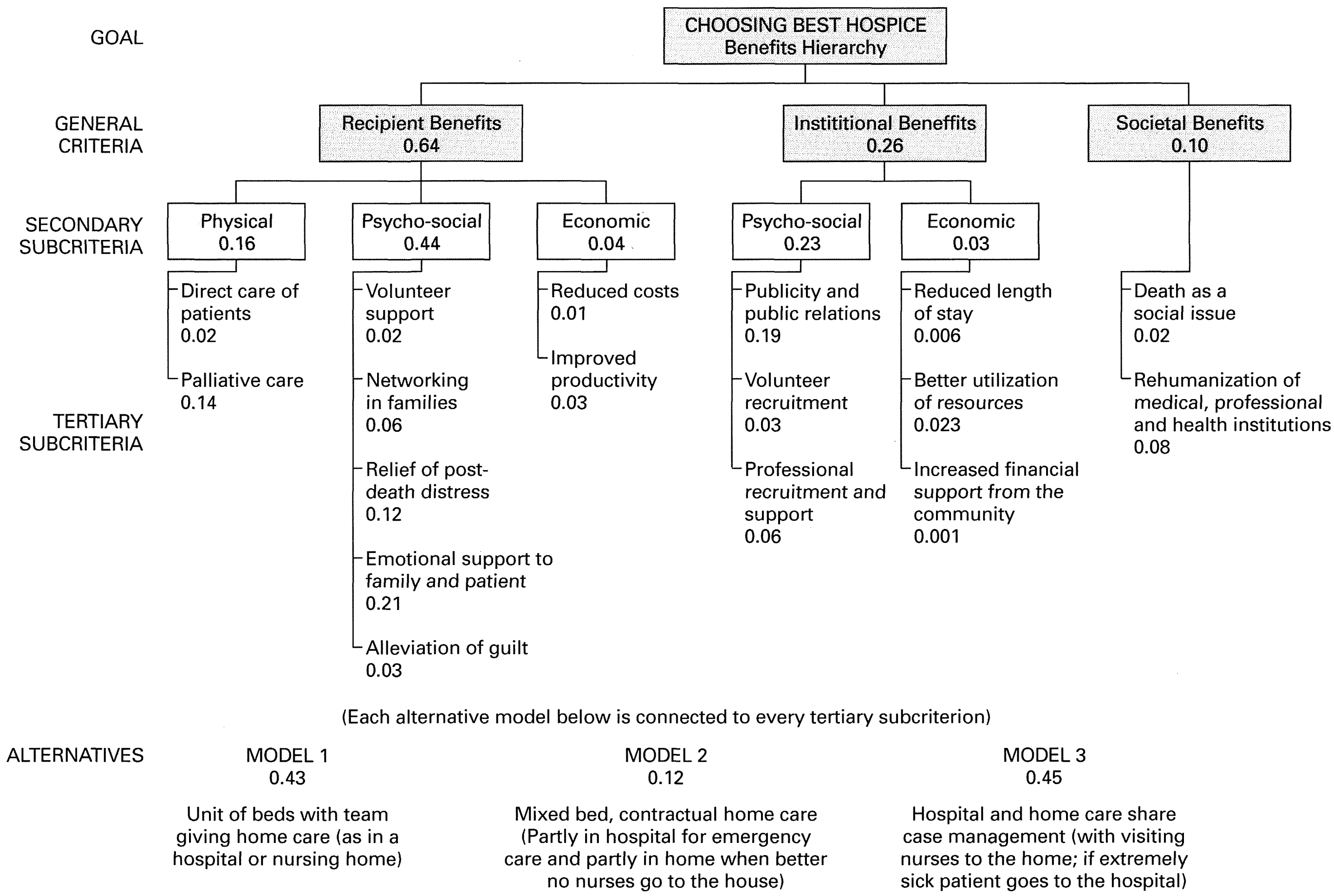

Figura 2. Hospice Benefits Hierarchy.

costs, education costs, bad debt costs, and recruitment costs. Educational costs apply to educating the community and training the staff. Recruitment costs apply to staff and volunteers. Since both the costs hierarchy and the benefits hierarchy concern the same decision, they both have the same alternatives in their bottom levels, even though the costs hierarchy has fewer levels.

Judgments and Comparisons

A judgment or comparison is the numerical representation of a relationship between two elements that share a common parent. The set of all such judgments can be represented in a square matrix in which the set of elements is compared with itself. Each judgment represents the dominance of an element in the column on the left over an element in the row on top. It reflects the answers to two questions: which of the two elements is more important with respect to a higher level criterion, and how strongly, using the 1- 9 scale shown in Table 1 for the element on the left over the element at the top of the matrix. If the element on the left is less important than that on the top of the matrix, we enter the reciprocal value in the corresponding position in the matrix. It is important to note that the lesser element is always used as the unit and the greater one is estimated as a multiple of that unit. From all the paired comparisons we calculate the priorities and exhibit them on the right of the matrix. For a set of n elements in a matrix one needs $n(n-1)/2$ comparisons because there are n 1's on the diagonal for comparing elements with themselves and of the remaining judgments, half are reciprocals. Thus we have $(n^2-n)/2$ judgments. When a rough estimate is needed or when pressed for time or when a very reliable expert is providing the judgments, one occasionally elicits only the minimum of n-1 judgments.

As usual with the AHP, in both the cost and the benefits models, we compared the criteria and subcriteria according to their relative importance with respect to the parent element in the adjacent upper level. For example, in the first matrix of comparisons of the three benefits criteria with respect to the goal of choosing the best hospice alternative, recipient benefits are moderately more important than institutional benefits and are assigned the

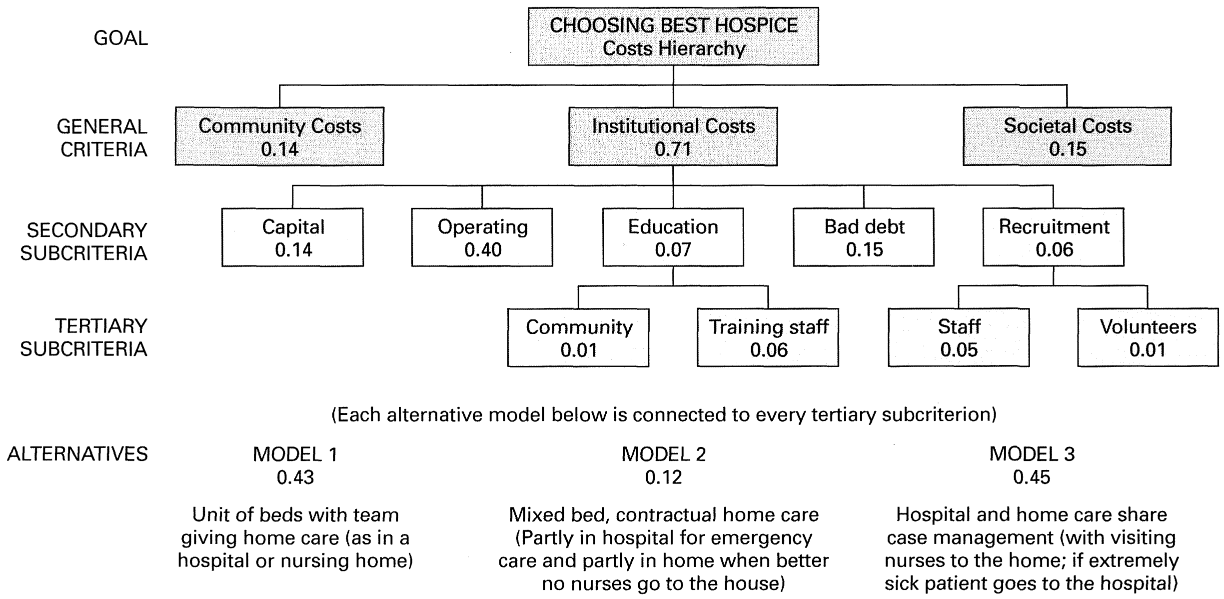

Figura 3. Hospice Costs Hierarchy.

absolute number 3 in the (1,2) or first-row second-column position. Three signifies three times more. The reciprocal value is automatically entered in the (2,1) position, where institutional benefits on the left are compared with recipient benefits at the top. Similarly a 5, corresponding to strong dominance or importance, is assigned to recipient benefits over social benefits in the (1,3) position, and a 3, corresponding to moderate dominance, is

Tabla 1. The Fundamental Scale

Intensity of Importance	**Definition**	**Explanation**
1	Equal Importance.	Two activities contribute equally to the objective.
3	Moderate importance.	Experience and judgment slightly favor one activity over another.
5	Strong importance.	Experience and judgment strongly favor one activity over another.
7	Very strong or demonstrated importance.	An activity is favored very strongly over another; its dominance demonstrated in practice.
9	Extreme importance.	The evidence favoring one activity over another is of the highest possible order of affirmation.
2, 4, 6, 8	For compromise between the above values.	Sometimes one needs to interpolate a compromise judgment numerically because there is no good word to describe it.
Reciprocals of above	If activity *i* has one of the above nonzero numbers assigned to it when compared with activity *j*, then *j* has the reciprocal value when compared with *i*.	A comparison mandated by choosing the smaller element as the unit to estimate the larger one as a multiple of that unit.
Rationals	Ratios arising from the scale.	If consistency were to be forced by obtaining *n* numerical values to span the matrix.
1.1-1.9	For tied activities.	When elements are close and nearly indistinguishable; moderate is 1.3 and extreme is 1.9.

assigned to institutional benefits over social benefits in the (2,3) position with corresponding reciprocals in the transpose positions of the matrix.

A scale of absolute numbers used to assign numerical values to judgments made by comparing two elements with the smaller element used as the unit and the larger one assigned a value from this scale as a multiple of that unit. *Importance* here is generic and can be replaced by *preference* or *likelihood*, the latter indicating that one can use it for risk analysis as the China example given later in the paper shows.

If there is a real ratio scale of measurement and it is desired to use it instead of the fundamental scale, one can use these values if one accepts the linearity inherent in the values of the scale and does not wish or is not allowed to (because for example, the problem belongs to someone else whose priorities may not be known at the time) interpret these values as priorities by putting them in ranges and applying the fundamental scale to compare the relative importance of these ranges. Otherwise, the values should be interpreted according to the fundamental scale.

There are numerous examples in the literature that serve to give validation to this scale, the protein example below is one.

Figure 4 below shows five areas to which the reader can apply to the paired comparison process in a matrix and use the 1-9 scale to test the validity of the procedure. We can approximate the priorities in the matrix by assuming that it is consistent. We normalize each column and then take the average of the corresponding entries in the columns.

The actual relative values of these areas are A = 0.47, B = 0.05, C = 0.24, D = 0.14, and E = 0.09 with which the answer may be compared. By comparing pairwise more than two alternatives in a decision problem, one is able to obtain better values for the derived scale because of redundancy in the comparisons, which helps improve the overall accuracy of the judgments.

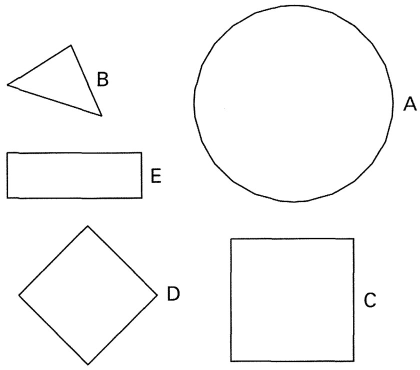

Figura 4. Five Figures to Compare in Pairs to Reproduce Their Relative Weights.

Note that the 1-9 scale can be extended to $1-\infty$ by a process of clustering as illustrated with the comparison of the unripe cherry tomato and the water melon. We now return to our hospice example.

Judgments in a matrix may not be consistent. In eliciting judgments, one makes redundant comparisons to improve the validity of the answer, given that respondents may be uncertain or may make poor judgments in comparing some of the elements. Redundancy gives rise to multiple comparisons of an element with other elements and hence to numerical inconsistencies. For example, where we compare recipient benefits with institutional benefits and with societal benefits, we have the respective judgments 3 and 5. Now if $x = 3y$ and $x = 5z$ then $3y = 5z$ or $y = 5/3\, z$. If the judges were consistent, institutional benefits would be assigned the value 5/3 instead of the 3 given in the matrix. Thus the judgments are inconsistent. In fact, we are not sure which judgments are more accurate and which are the cause of the inconsistency. Inconsistency is inherent in the judgment process. Inconsistency may be considered a tolerable error in measurement only when it is of a lower order of magnitude (10 percent) than the actual measurement itself; otherwise the inconsistency would bias the result by a sizable error comparable to or exceeding the actual measurement itself.

RELATIVE AMOUNT OF PROTEIN IN SEVEN FOODS

Protein in Food	What food has more protein? A	B	C	D	E	F	G	Estimated Values	Actual Values
A: Steak	1	9	9	6	4	5	1	0.345	0.370
B: Potatoes	1/9	1	1	1/2	1/4	1/3	1/4	0.031	0.040
C: Apples	1/9	1	1	1/3	1/3	1/5	1/9	0.030	0.000
D: Soybean	1/6	2	3	1	1/2	1	1/6	0.065	0.070
E: Whole Wheat Bread	1/4	4	3	2	1	3	1/3	0.124	0.110
F: Tasty Cake	1/5	3	5	1	1/3	1	1/5	0.078	0.090
G: Fish	1	4	9	6	3	5	1	0.328	0.320

When the judgments are inconsistent, the decision maker may not know where the greatest inconsistency is. The AHP can show one by one in sequential order which judgments are the most inconsistent, and that suggests the value that best improves consistency. However, this recommendation may not necessarily lead to a more accurate set of priorities that correspond to some underlying preference of the decision makers. Greater consistency does not imply greater accuracy and one should go about improving consistency (if one can given the available knowledge) by making slight changes compatible with one's understanding. If one cannot reach an acceptable level of consistency, one should gather more information or reexamine the framework of the hierarchy[16].

Under each matrix I have indicated a consistency ratio (CR) comparing the inconsistency of the set of judgments in that matrix with what it would be if the judgments and the corresponding reciprocals were taken at random from the scale. For a 3-by-3 matrix this ratio should be about five percent, for a 4-by-4 about eight percent, and for larger matrices, about 10 percent [15, 16].

Priorities are numerical ranks measured on a ratio scale. A ratio scale is a set of positive numbers whose ratios remain the same if all the numbers are multiplied by an arbitrary positive number. An example is the scale used to measure weight. The ratio of these weights is the same in pounds and in kilograms. Here one scale is just a constant multiple of the other. The object of evaluation is to elicit judgments concerning relative importance of the elements of the hierarchy to create scales of priority of influence.

Because the benefits priorities of the alternatives at the bottom level belong to a ratio scale and their costs priorities also belong to a ratio scale, and since the product or quotient (but not the sum or the difference) of two ratio scales is also a ratio scale, to derive the answer we divide the benefits priority of each alternative by its costs priority. We then choose the alternative with the largest of these ratios. It is also possible to allocate a resource proportionately among the alternatives.

I will explain how priorities are developed from judgments and how they are synthesized down the hierarchy by a process of weighting and adding to go from *local* priorities derived from judgments with respect to a single criterion to *global* priorities derived from multiplication by the priority of the criterion and *overall* priorities derived by adding the global priorities of the same element. The local priorities are listed on the right of each matrix. If the judgments are perfectly consistent, and hence CR = 0, we obtain the local priorities by adding the values in each row and dividing by the sum of all the judgments, or simply by normalizing the judgments in any column, by dividing each entry by the sum of the entries in that column. If the judgments are inconsistent but have a tolerable level of inconsistency, we obtain the priorities by raising the matrix to large powers, which is known to take into consideration all intransitivities between the elements, such as those I showed above between x, y, and z [16]. Again, we obtain the priorities from this matrix by adding the judgment values in each row and dividing by the sum of all the judgments. To summarize, the global priorities at the level immediately under the goal are equal to the local priorities because the priority of the goal is equal to one. The global priorities at the next level are obtained by weighting the local priorities of this level by the global priority at the level immediately above and so on. The overall priorities of the alternatives are obtained by weighting the local priorities by the global priorities of all the parent criteria or subcriteria in terms of which they are compared and then adding. (If an element in a set is not comparable with the others on some property and should be left out, the local priorities can be augmented by adding a zero in the appropriate position.) In Table 2 we compare the criteria under benefits.

The process is repeated in all the matrices by asking the appropriate dominance or importance question. For example, for the matrix comparing the subcriteria of the parent criterion institutional benefits (Table 3), psychosocial benefits are regarded as very strongly more important than economic benefits, and 7 is entered in the (1,2) position and 1/7 in the (2,1) position.

In comparing the three models for patient care, we asked members of the planning association which model they preferred with respect to each of the covering or parent secondary criterion in level 3 or with respect to the tertiary criteria in level 4. For example, for the subcriterion direct care (located on the left-most branch in the benefits hierarchy), we obtained a matrix of paired

Table 2. The judgments in this matrix are the responses to the question: Which criterion is more important with respect to choosing the best hospice alternative and how strongly?

Choosing Best Hospice	Recipient Benefits	Institutional Benefits	Social Benefits	Priorities
Recipient Benefits	1	3	5	.64
Institutional Benefits	1/3	1	3	.26
Societal Benefits	1/5	1/3	1	.11

C.R. = .33

Table 3. The judgments in this matrix are the responses to the question: Which subcriterion yields the greater benefit with respect to institutional benefits and how strongly?

Institutional benefits	Psycho-social	Economic	Priorities
Psycho-social	1	7	.875
Economic	1/7	1	.125

C.R. = .000

Table 4. The judgments in this matrix are the responses to the question: Which model yields the greater benefit with respect to direct care of patient and how strongly?

Direct of patient	Model I	Model II	Model III	Priorities
Model I: Unit/Team	1	5	3	.64
Model II: Mixed/Home Care	1/5	1	1/3	.10
Model III: Case Management	1/3	3	1	.26

C.R. = .33

comparisons (Table 4) in which Model 1 is preferred over Models 2 and 3 by 5 and 3 respectively and Model 3 is preferred by 3 over Model 2. The group first made all the comparisons using semantic terms for the fundamental scale and then translated them to the corresponding numbers.

For the costs hierarchy, I again illustrate with three matrices. First the group compared the three major cost criteria and provided judgments in response to the question: which criterion is a more important determinant of the cost of a hospice model? Table 5 shows the judgments obtained.

The group then compared the subcriteria under institutional costs and obtained the importance matrix shown in Table 6.

Finally we compared the three models to find out which incurs the highest cost for each criterion or subcriterion. Table 7 shows the results of comparing them with respect to the costs of recruiting staff.

Our procedure for synthesis involves multiplying the priorities of the alternatives by those of the criteria and adding as shown below. This additive procedure is mandatory because we needed to give back priorities derived from using actual measurements. In the following example, we have two criteria «price» and «repair cost». We also have three items, A, B, and C, whose values are as follows:

Normalized sum

	Price	Repair Cost	Total	Total Normalized
A	200	150	350	.269
B	300	50	350	.269
C	500	100	600	.462

Weighting, adding and then normalizing

Price (1000/1300)	Repair Cost (300/1300)	Weighting and adding
200/1000	150/300	.269
300/1000	50/300	.269
500/1000	100/300	.462

Note that the priority of each criterion is the quotient of the sum of the values of the items under it to the sum of the values of the items under both criteria.

Table 5. The judgments in this matrix are the responses to the question: Which criterion is a greater determinant of cost with respect to the care method and how strongly?

Choosing Best hospice (costs)	Community	Institutional	Societal	Priorities
Community Costs	1	1/5	1	.14
Institutional Costs	5	1	5	.71
Societal Costs	1	1/5	1	.14

C.R. = .000

Table 6. The judgments in this matrix are the responses to the question: Which criterion incurs greater institutional costs and how strongly?

Institutional costs	Capital	Operating	Education	Bad Debt	Recruitment	Priorities
Capital	1	1/7	1/4	1/7	1	.05
Operating	7	1	9	4	5	.57
Education	4	1/9	1	1/2	1	.10
Bad Debt	7	1/4	2	1	3	.21
Recruitment	1	1/5	1	1/3	1	.07

C.R. = .08

As shown in Table 8, we divided the benefits priorities by the costs priorities for each alternative to obtain the best alternative, model 3, the one with the largest value for the ratio. Table 8 shows two ways or modes of synthesizing the local priorities of the alternatives using the global priorities of their parent criteria: The distributive mode and the ideal mode. In the distributive mode, the weights of the alternatives sum to one. It is used when there is dependence among the alternatives and a unit priority is distributed among them. The ideal mode is used to obtain the single best alternative regardless of what other alternatives there are. In the ideal mode, the local priorities of the alternatives under each criterion are divided by the largest value among them. For each criterion one alternative becomes an ideal with value one. Synthesis is obtained by multiplying these values by the priorities of their corresponding criteria and then adding. An alternative that is best for every criterion receives a composite priority of one. All other alternatives receive a smaller value. The composite values also belong to a ratio scale. Adding an irrelevant alternative does not change the ranking of the highly ranked alternatives. In addition, if new alternatives are introduced that are assigned greater values than the best alternative without keeping the values assigned before to the existing alternatives, there can be no reversal in the ranks of the old alternatives.

In our example, for both modes the local priorities are weighted by the global priorities of the parent criteria and synthesized and the benefit-to-cost ratios formed. In this case, both modes lead to the same outcome for hospice, which is model 3. As we shall see below, we need both modes to deal with the effect of adding (or deleting) alternatives on an already ranked set.

Model 3 has the largest ratio scale values of benefits to costs in both the distributive and ideal modes, and the hospital selected it for treating terminal patients. This need not always be the case. In this case, there is dependence of the personnel resources allocated to the three models because some of these resources would be shifted based on the decision. Therefore the distributive mode is the appropriate method of synthesis. If the alternatives were sufficiently distinct with no dependence in their definition, the ideal mode would be the way to synthesize.

I also performed marginal analysis to determine where the hospital should allocate additional resources for the greatest marginal return. To perform marginal analysis, I first ordered the alternatives by increasing cost priorities and then formed the benefit-to-cost ratios corresponding to the smallest cost, followed by the ratios of the differences of successive benefits to costs. If this difference in benefits is negative, the new alternative is dropped from consideration and the process continued. The alternative with the largest marginal ratio is then chosen. For the costs and corresponding benefits from the synthesis rows in Table 8, I obtained:

Costs:	.20	.21	.59
Benefits:	.12	.45	.43

Marginal Ratios:

$$\frac{.12}{.20} = .60 \qquad \frac{.45 - .12}{.21 - .20} = 33 \qquad \frac{.43 - .45}{.59 - .21} = -0.05$$

The third alternative is not a contender for resources because its marginal return is negative. The second alter-

Table 7. The entries in this matrix respond to the question: Which model incurs greater cost with respect to institutional costs for rrecruiting staff and how strongly?

Institutional costs for recruiting staff	Model I	Model II	Model III	Priorities
Model I: Unit/Team	1	4	4	.66
Model II: Mixed/Home Care	1/4	1	1	.17
Model III: Case Management	1/4	1	1	.17

C.R. = .000

Tabla 8. The benefit/cost ratios of the three models given in the bottom row of the table are obtained for both the distributive and ideal modes. Here one multiplies each of the six columns of priorities of a model by the column of criteria weights on the left and adds to obtain the synthesis of overall priorities, once for the benefits (top half of table) and once for the costs (bottom half of table) and forms the ratios of corresponding synthesis numbers to arrive at the benefit/cost ratio (bottom row of table)

		Distributive Mode			Ideal Mode		
Benefits	**Priorities**	**Model 1**	**Model 2**	**Model 3**	**Model 1**	**Model 2**	**Model 3**
Direct Care of Patient	.02	0.64	0.10	0.26	1.000	0.156	0.406
Palliative Care	.14	0.64	0.10	0.26	1.000	0.156	0.406
Volunteer Support	.02	0.09	0.17	0.74	0.122	0.230	1.000
Networking in Families	.06	0.46	0.22	0.32	1.000	0.478	0.696
Relief of Post Death Stress	.12	0.30	0.08	0.62	0.484	0.129	1.000
Emotional Support of Family and Patient	.21	0.30	0.08	0.62	0.484	0.129	1.000
Alleviation of Guilt	.03	0.30	0.08	0.62	0.484	0.129	1.000
Reduced Economic Costs for Patient	.01	0.12	0.65	0.23	0.185	1.000	0.354
Improved Productivity	.03	0.12	0.27	0.61	0.197	0.443	1.000
Publicity and Public Relations	.19	0.63	0.08	0.29	1.000	0.127	0.460
Volunteer Recruitment	.03	0.64	0.10	0.26	1.000	0.156	0.406
Professional Recruitment and Support	.06	0.65	0.23	0.12	1.000	0.354	0.185
Reduced Length of Stay	.006	0.26	0.10	0.64	0.406	0.406	1.000
Better Utiliation of Resources	.023	0.09	0.22	0.69	0.130	0.130	1.000
Increased Monetary Support	.001	0.73	0.08	0.19	1.000	1.000	0.260
Death as a Social Issue	.02	0.20	0.20	0.60	0.333	0.333	1.000
Rehumanization of Institutions	.08	0.24	0.14	0.62	0.387	0.226	1.000
Synthesis		0.428	0.121	0.451	0.424	0.123	0.453
Costs							
Community Costs	.14	0.33	0.33	0.33	1.000	1.000	1.000
Institutional Capital Costs	.03	0.76	0.09	0.15	1.000	0.118	0.197
Institutional Operating Costs	.40	0.73	0.08	0.19	1.000	0.110	0.260
Institutional Costs for Educating the Community	.01	0.65	0.24	0.11	1.000	0.369	0.169
Institutional Costs for Training Staff	.06	0.56	0.32	0.12	1.000	0.571	0.214
Institutional Bad Debt	.15	0.60	0.20	0.20	1.000	0.333	0.333
Institutional Costs of Recruiting Staff	.05	0.66	0.17	0.17	1.000	0.258	0.258
Institutional Costs of Recruiting Volunteers	.01	0.60	0.20	0.20	1.000	0.333	0.333
Societal Costs	.15	0.33	0.33	0.33	1.000	1.000	1.000
Synthesis		0.583	0.192	0.224	0.523	0.229	0.249
Benefict/Cost Ratio		0.734	0.630	2.013	0.811	0.537	1.819

native is best. In fact, in addition to adopting the third model, the hospital management chose the second model of hospice care for further development.

A Second Example Combining Benefits, Costs and Risks -The Wisdom of a Trade War with China over Intellectual Property Rights

This example was developed jointly with the author's colleague Professor Jen S. Shang in mid February 1995 to understand the issues when the media were voicing strong conflicting concerns prior to the action to be taken in Beijing later in February. Many copies of the analysis were sent to congressmen and senators and to the chief U.S. negotiator in Washington, and to several newspapers in the U.S. and in China. A telephone call was received from Mr. Mickey Kantor's office, the chief US negotiator after the meeting in Beijing congratulating us on the analysis not to sanction China. The person calling said, «Aren't you glad we did not sanction China?» We believe that this short and concise analysis may have had some effect on that decision. The full write up of about 7 pages is not included here because of space limitation. The reader will have no difficulty following the analysis carried out in the three hierarchies shown below. I have kept the tense of the the next few paragraphs as it was when the paper was written to better convey the sense of urgency in which it was written.

There are many and strong conflicting opinions about what to do with Chinese piracy of U.S. technology and management know-how. Should the U.S. sanction China on February 26? The basic arguments in favor of imposing tariffs derive from the U.S. perceived need not to allow China to become a runaway nation with an inward oriented closed economy. Some also argue convincingly that a nation whose economy will equal that of the U.S. in three decades must be taught to play by the rules. We have made a brief study of the decision to impose tariffs on Chinese products in the U.S. It is not the immediate small injury to U.S. corporations from such an action that is of major concern, but what might happen in the future. The effect of the tariffs will be decisively more intangible with long-term results that can aggravate trade in the Pacific.

Our findings based on benefits, costs, and risks and on all the factors we could bring to bear on the outcome is a definite and very decisive ***No***, which means that it is not in the best overall interest of the U.S. to take strong action against China. Since usually we are not told much about what China says, we also summarize some arguments gleaned from Chinese newspapers. We explain our analysis and offer the reader the opportunity to perform a similar evaluation based on the factors given here plus others we may have overlooked. In our opinion, the costs are too high to treat China in the same style as an outlaw nation even though China can and should do better as a member of the world community.

To arrive at a rational decision, we considered the factors that influence the outcome of the decision, and arranged them in three hierarchies: one for the benefits of implementing such a sanction, one for the costs and a third for the risks and uncertainties that can occur (see Figure 5). Each hierarchy has a goal followed by the criteria that affect the performance of the goal. The alternatives are listed at the last level of the hierarchy. They are: Yes - to sanction China or No – not to sanction China.

In each hierarchy, we synthesize the values for *Yes* and for *No* by multiplying each alternative's priority with the importance of its parent criterion, and adding to obtain the overall result for *Yes* and for *No*. A user-friendly computer software program, Expert Choice, was used to do all the calculations. To combine the results from the three hierarchies, we divide the benefit results for *Yes* by the costs and by the risks for *Yes* to obtain the final outcome. We do the same for *No* and select *Yes* or *No* depending on which has the larger value. While *Yes*'s benefits are high, the corresponding costs and risks are also high. Its ratio is less than that of the *No* decision. *No* dominates *Yes* both when no risk is considered and also when projected risk is taken into account. Including risk by using possible scenarios of the future can be a powerful tool in assessing the decision on the effect of the future.

To ensure that the outcome not be construed as a result of whimsical judgments, we performed a comprehensive sensitivity analysis. Sensitivity analysis assists the decision maker to discover how changes in the priorities affect the recommended decision. The *Yes* and *No* weights are fixed because they are our best judgments based on the facts. So we fixed the *Yes* and *No* judgments as shown in Figure 6 and varied the importance of each factor. There is a wide range of admissible priority value that a policy maker may choose for each factor. Our sensitivity analysis covers all the reasonable priorities a politician might choose. We changed each factor's importance from its value indicated in the hierarchy to the near extreme values 0.2 and 0.8. This gave us six variations in each hierarchy because there are three factors in each. With three hierarchies, we generated 216 (6*6*6) data points. In this simulation, we found that it is only when long term negative competition is thought to be unimportant that sanctions would be justified. From Figure 6 depicting the 216 possibilities, we see that *No* dominates *Yes* appreciably. Regardless of the weights one assigns to the factors, over 90% of the cases lead to *No*, not to sanction China.

Deng Rong, the daughter of Deng Xiaoping, the most senior elder statesman of China, said recently «sanctions are never the best way to resolve a dispute. One should talk things over and consider the interests of the people.» Our analysis supports this attitude.

7. GENERALIZATION TO THE CONTINUOUS CASE

The expression encountered for deriving a ratio scale for pairwise comparisons in the finite case

$$\sum_{j=1}^{n} a_{ij} w_j = \lambda_{\max} w_i \tag{9}$$

with $a_{ij} = 1/a_{ij}$ or $a_{ij}a_{ij} = 1$ (the reciprocal property), $a_{ij}) > 0$ and

$$\sum_{j=1}^{n} w_i = 1 \tag{10}$$

generalizes to the continuous case through Fredholm's integral equation of the second kind:

$$\int_a^b K(s, t)\, w(t)\, dt = \lambda_{\max} w(s) \tag{11}$$

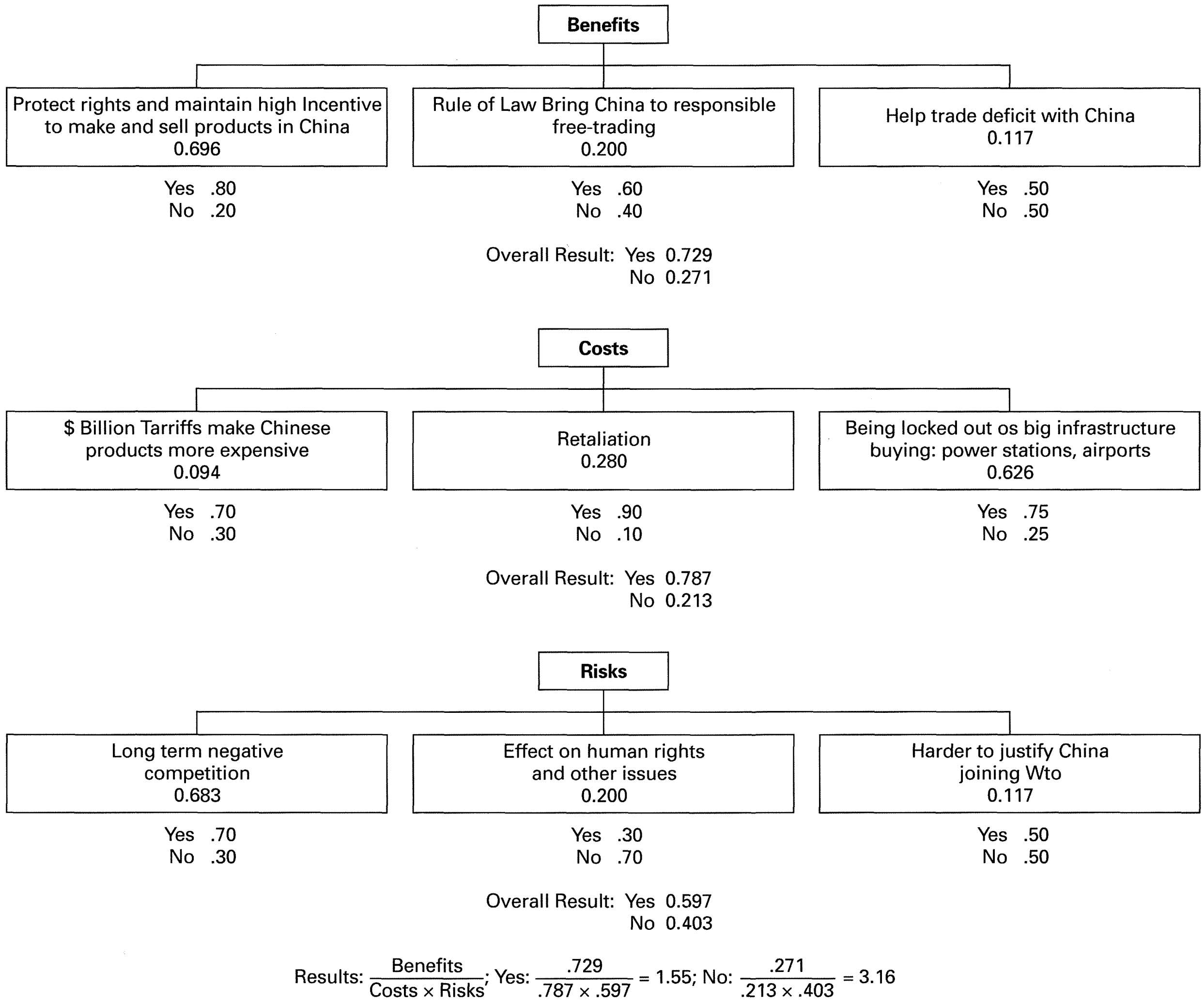

Figura 5. Benefits, Costs and Risks of Sanctioning China.

or more conventionally

$$\lambda \int_a^b K(s, t)\, w(t)\, dt = w(s) \tag{12}$$

with the normalization condition:

$$\int_a^b w(s)\, ds = 1 \tag{13}$$

where instead of the matrix A we have a positive kernel, $K(s,t) > 0$. The problem is to determine the principal right eigenfunction $w(s)$ of K.

We can easily see by substituting in the equation that $Cw(t)$ is also an eigenfunction corresponding to the same λ. The value $\lambda = 0$ is not a characteristic value because we have the corresponding solution $w(t) = 0$ for every value of t, which is the trivial case. Here also, we have the reciprocal property

$$K(s, t)\, K(t, s) = 1 \tag{14}$$

so that $K(s, t)$ is not only positive, but also reciprocal. An example of this type of kernel is

$$K(s, t) = e^{s-t} = e^s/e^t$$

As in the finite case, the kernel $K(s, t)$ is consistent if it satisfies the relation

$$K(s, t)\, K(t, u) = K(s, u), \text{ for all } s, t, \text{ and } u \tag{15}$$

It follows by putting $s = t = u$, that $K(s, s) = 1$ for all s which is analogous to having ones down the diagonal of the matrix in the discrete case.

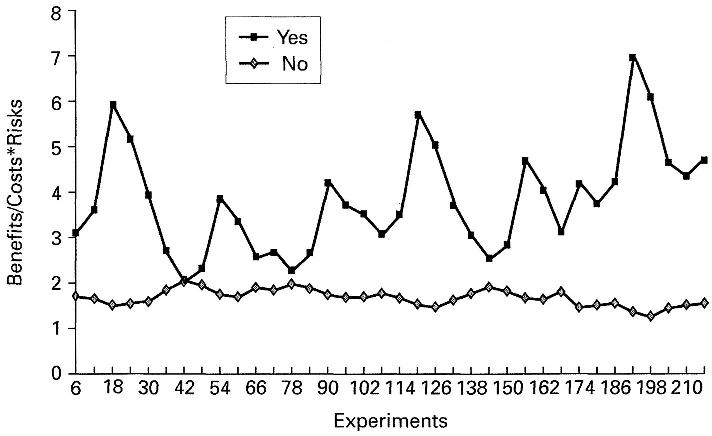

Figura 6. The dominance of *No* over *Yes*.

We now generalize the result that a matrix is consistent if and only if it has the form $A = (w_i/w_j)$ which is equivalent to multiplying a column vector that is the transpose of $(w_1, ..., w_n)$ by the row vector $(1/w_1, ..., 1/w_n)$. As we see below, the kernel $K(s, t)$ is separable and can be written as

$$K(s, t) = k_1(s)\, k_2(t)$$

Theorem 13. $K(s, t)$ is consistent if and only if it is separable of the form:

$$K(s, t) = k(s)/k(t) \tag{16}$$

Proof. (Necessity) $K(t, u_0) \neq 0$ for some $u_0 \in S$, otherwise $K(t, u_0) = 0$ for all u_0 would contradict

$K(u_0, u_0) = 1$ for $t = u_0$. Using (15) we obtain

$$K(s, t)\, K(t, u_0) = K(s, u_0)$$

$$K(s, t) = \frac{K(s, u_0)}{K(t, u_0)} = \frac{k(s)}{k(t)}$$

for all $u_0 \in S$ and the result follows.

(Sufficiency) If (16) holds, then it is clear that $K(s, t)$ is consistent.

We now prove that as in the discrete case of a consistent matrix, whereby the eigenvector is given by any normalized column of the matrix, that an analogous result obtains in the continuous case.

Theorem 14. If $K(s, t)$ is consistent, the solution of (12) is given by

$$w(t) = \frac{k(s)}{\int_s k(s)\, ds} \tag{17}$$

Proof. We replace $w(t)$ in (12) by $\lambda \int_s K(t, u)\, w(u)\, du$ inside the integral and repeat the process n times. Passing to the limit we obtain:

$$w(s) = \lim_{n\to\infty} \lambda^n \int_s \int_s \cdots \int_s K(s, s_1)\, K(s_1, s_2) \cdots K(s_{n-1}, s_n)\, ds_1 \cdots ds_n.$$

Since $K(s, t)$ is consistent, we have:

$$w(s) = \lim_{n\to\infty} \lambda^n \int_s \int_s \cdots \int_s K(s, s_n)\, ds_1\, ds_2 \cdots ds_n =$$

$$= \lim_{n\to\infty} \lambda \int_s K(s, s_n)\, ds_n$$

With $\int_s w(s)\, ds = 1$ we have

$$w(s) = \lim_{n\to\infty} \int_s K(s, s_n)\, ds_n / [\int_s \lim_{n\to\infty} \int_s K(s, s_n)\, ds_n\, ds]$$

$$= \lim_{n\to\infty} \int_s K(s, s_n)\, ds_n / \int_s \int_s K(s, s_n)\, ds_n\, ds]$$

Also, because $K(s, s_n)$ is consistent we have $K(s, s_n) = k(s)/k(s_n)$ and $w(s) = k(s)/\int_s k(s) ds$.

We now determine the form of $k(s)$ and also of $w(s)$. We have for a ratio scale $w(s) > 0$, for $s > 0$, and $w(0)$ arbitrary.

We know that the solution of our homogeneous Fredholm equation has the general form:

$$w(s) = k(s)/\int_s k(s) ds = a\, k(s)$$

which is a constant times $k(s)$. But we know more about the consistent kernel $K(s, t)$. Recall that in the discrete case, the normalized eigenvector was independent of whether all the elements of the pairwise comparison matrix A are multiplied by the same constant a or not, and thus we can replace A by aA and obtain the same eigenvector. Generalizing this result we have:

$$K(as, at) = aK(s, t) = k(as)/k(at) = a\, k(s)/k(t)$$

Because K is a degenerate kernel, we can replace $k(s)$ above by $k(as)$ and obtain $w(as)$. We have now derived from considerations of ratio scales the following condition to be satisfied by a ratio scale:

Theorem 15. $w(s)$ is an eigenfunction solution of (12) with a consistent kernel K that is homogeneous of order one, if and only if the following functional equation holds

$$w(as) = bw(s)$$

where $b = aa$.

We have transformed the condition of solvability to a functional equation, the fundamental equation of ratio scales. Luckily, this equation was first studied as a functional equation without knowledge of its connection to ratio scales by J. Aczel and M. Kuczma [1] in 1991 who called it a Folk_Theorem. The solution has been recently developed in detail in the complex domain by my friend J. Aczel, for a book on neural firing I just completed. Here is a brief part of that solution. If we substitute $s = a^u$ in the equation we have:

$$w(a^{u+1}) - bw(a^u) = 0.$$

Again if we write $w(a^u) = b^u p(u)$, we get:

$$p(u + 1) - p(u) = 0$$

which is a periodic function of period one in the variable u (such as $\cos u/2\pi$). Note that if a and s are real, then so is u which may be negative even if a and s are both assumed to be positive.

By dividing its variable by its period, any periodic function can be reduced to a periodic function of period one. Thus, whatever is known about periodic functions applies to periodic functions of period one and conversely. If P is periodic of period T, i.e. $P(x + T) = P(x)$, then $p(x) = P(Tx)$ will be periodic of period 1: $p(x + 1) = P\big(T(x + 1)\big) = P(Tx + T) = P(Tx) = p(x)$; the converse operation is obvious.

If in the last equation $p(0)$ is not equal to 0, we can introduce $C = p(0)$ and $P(u) = p(u)/C$, we have for the general response function $w(s)$,

$$w(s) = Ce^{\log b \frac{\log s}{\log a}}\, P\left(\frac{\log s}{\log a}\right)$$

where P is also periodic of period 1 and $P(0)=1$. Note that $C > 0$ only if $p(0)$ is positive. Otherwise, if $p(0) < 0$, $C < 0$ (see Saaty [17].)

8. ABSOLUTE MEASUREMENT: EVALUATING EMPLOYEES FOR RAISES

Employees are evaluated for raises. The criteria are Dependability, Education, Experience, and Quality. Each criterion is subdivided into intensities, standards, or subcriteria as shown in Fig. 7. Priorities are set for the criteria by comparing them in pairs, and these priorities are then given in a matrix. The intensities are then pairwise compared according to priority with respect to their parent criterion (as in Table 9) and their priorities are divided by the largest intensity for each criterion (second column of priorities in Figure 7). Finally, each individual is rated in Table 10 by assigning the intensity rating that applies to him or her under each criterion. The scores of these subcriteria are weighted by the priority of that criterion and summed to derive a total ratio scale score for the individual. This approach can be used whenever it is possible to set priorities for intensities of criteria, which is usually possible when sufficient experience with a given operation has been accumulated. Salary raises can be made proportionately to the final priorities.

The priorities for the intensities themselves are also established through a pairwise comparison process as shown in Table 9. Note that the priorities of the intensities shown in Figure 7 are *weighted* priorities, that is, the priorities obtained from the comparisons shown in Table 9 have been weighted by the priority of their parent element. The intensities for each criterion may be weighted differently, even though the words used such as Outstanding, Above Average may be the same.

Here again one has both the distributive and the ideal modes. In the ideal mode all the intensities under each criterion are divided by the priority of the highest intensity. In this case introducing additional alternatives have no effect on the ranking of the other alternatives. Note that with absolute measurement one can take out the few ranking alternatives and pairwise compare just that set to see if a finer ranking can be obtained with paired comparisons.

9. RANK PRESERVATION AND REVERSAL

Given the assumption that the alternatives of a decision are completely independent of one another, can and should the introduction (deletion) of new (old) alternatives change the rank of some alternatives without introducing new (deleting old) criteria, so that a less preferred alternative becomes most preferred? Incidentally, how one prioritizes the criteria and subcriteria is even more important than how one does the alternatives which are

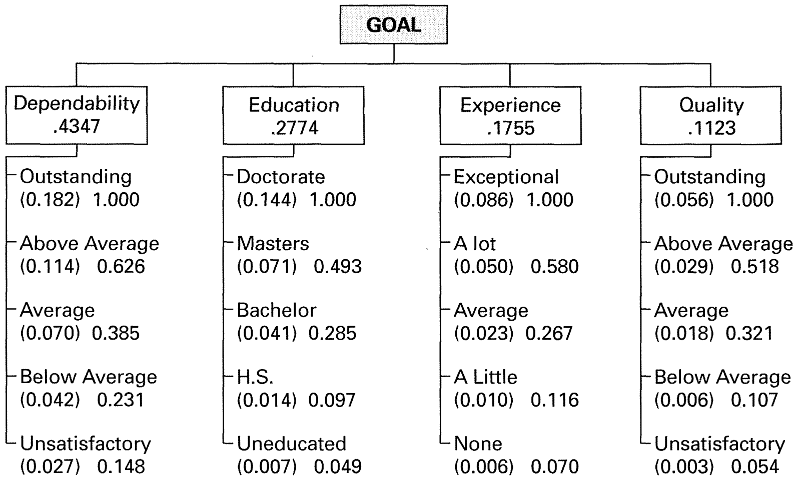

Figura 7. Employee Evaluation Hierarchy.

themselves composites of criteria. Can rank reverse among the criteria themselves if new criteria are introduced? Why should that not be as critical a concern? The answer is simple. In its original form utility theory assumed that criteria could not be weighted and the only important elements in a decision were the alternatives and their utilities under the various criteria. Today utility theorists imitate the AHP by rating, and some even by comparing the criteria, somehow. There was no concern then about what would happen to the ranks of the alternatives should the criteria weights themselves change as there were none. The tendency, even today, is to be unconcerned about the theory of rank preservation and reversal among the criteria.

The house example of a previous section teaches us an important lesson. If we add a fourth house to the collection, the priority weights of the criteria Price and Remodeling Cost would change accordingly. Thus the measurements of the alternatives and their number which we call structural factors, always affect the importance of the criteria. When the criteria are incommensurate and their functional priorities are determined in terms of yet higher level criteria or goals, one must still weight such functional importance of the criteria by the structural effect of the alternatives. What is significant in all this is that the importance of the criteria always depends on the measurements of the alternatives. If we assume that the alternatives are measured on a different scale for each criterion, it becomes obvious that normalization is the instrument that provides the structural effect to update the importance of the criteria in terms of what alternatives there are. Finally, the priorities of the alternatives are weighted by the priorities of the criteria that depend on the measurements of the alternatives. This implies that the overall ranking of any alternative depends on the measurement and number of all the alternatives. To always preserve rank means that the priorities of the criteria should not depend on the measurements of the alternatives but should only derive from their own functional importance with respect to higher goals. This implies that the alternatives should not depend on the measurements of other alternatives. Thus one way to always preserve rank is to rate the alternatives one at a time. In the AHP this is done through absolute measurement with respect to a complete set of intensity ranges with the largest value intensity value equal to one. It is also possible to preserve rank in relative measurement by using an ideal alternative with full value of one for each criterion.

The logic about what can or should happen to rank when the alternatives *depend* on each other has always

Table 9. Ranking Intensities

	Outstanding	Above Average	Average	Below Average	Unsatis-factory	Priorities
Outstanding	1.0	2.0	3.0	4.0	5.0	0.419
Above Average	1/2	1.0	2.0	3.0	4.0	0.263
Average	1/3	1/2	1.0	2.0	3.0	0.630
Below Average	1/4	1/3	1/2	1.0	2.0	0.097
Unsatisfactory	1/5	1/4	1/3	1/2	1.0	0.062
Inconsistency Ratio = 0.015						

Table 10. Ranking Alternatives

	Dependability .4347	**Education .2774**	**Experience .1775**	**Quality .1123**	**Total**
1. Adams, V.	Outstanding	Bachelor	A Little	Outstanding	0.646
2. Becker, L.	Average	Bachelor	A Little	Outstanding	0.379
3. Hayat, F.	Average	Masters	A Lot	Below Average	0.418
4. Kesselman, S.	Above Average	H.S.	None	Above Average	0.369
5. O'Shea, K.	Average	Doctorate	A Lot	Above Average	0.605
6. Peters, T.	Average	Doctorate	A Lot	Average	0.583
7. Tobias, K.	Above Average	Bachelor	Average	Above Average	0.456

been that *anything* can happen. Thus, when the criteria functionally depend on the alternatives, which implies that the alternatives, which of course depend on the criteria, would then depend on the alternatives themselves, rank may be allowed to reverse. The Analytic Network Process (ANP) is the generalization of the AHP to deal with ranking alternatives when there is functional dependence and feedback of any kind. Even here, one can have a decision problem with dependence among the criteria, but with no dependence of criteria on alternatives and rank may still need to be preserved. The ANP takes care of functional dependence, but if the criteria do not depend on the alternatives, the latter are kept out of the supermatrix and ranked precisely as in a hierarchy.

Examples of rank reversal abound in practice, and they do not occur because new criteria are introduced (see chapter 5 [16] for examples of both rank and preference reversals by utility theorists). The requirement that rank always be preserved or that it should be preserved with respect to irrelevant alternatives has been shown to be false with many counterexamples. To every rule or generalization that one may wish to set down about rank, it is possible to find a counterexample that violates that rule. Here is the last and most extreme form of four variants of an attempt to qualify what should happen to rank given by Luce and Raiffa, each of which is followed by a counterexample. They state it but and then reject it. *The addition of new acts to a decision problem under uncertainty never changes old, originally non-optimal acts into optimal ones. The all-or-none feature of the last form may seem a bit too stringent ... a severe criticism is that it yields unreasonable results.* There are numerous examples given in the literature where it is shown that this is an unreasonable assumption, the most elementary of which is what happens to the rank of an alternative when a million (or an entire universe) of copies of it are introduced. Most of the time its rank is decreased. The effects of copies, phantoms,decoys and other types of alternatives have been examined in the literature [16].The AHP has a theory and implementation procedures and guidelines for when to preserve rank and when to allow it to reverse. One mode of the AHP allows an irrelevant alternative to cause reversal among the ranks of the original alternatives.

Guidelines for Selecting the Distributive or Ideal Mode

The distributive mode of the AHP produces preference scores by normalizing the performance scores; it takes the performance score received by each alternative and divides it by the sum of performance scores of all alternatives under that criterion. This means that with the Distributive mode the preference for any given alternative would go up if we reduce the performance score of another alternative or remove some alternatives. The Ideal mode compares each performance score to a fixed benchmark such as the performance of the best alternative under that criterion. This means that with the Ideal mode the preference for any given alternative is independent of the performance of other alternatives, except for the alternative selected as a benchmark. Saaty and Vargas (1993) have shown by using simulation, that there are only minor differences produced by the two synthesis modes. This means that the decision should select one or the other if the results diverge beyond a given set of acceptable data.

The following guidelines were developed by Millet and Saaty (2000), in a forthcoming paper, to reflect the core differences in translating performance measures to preference measures of alternatives. *The Distributive (dominance) synthesis mode should be used when the decision maker is concerned with the extent to which each alternative dominates* all other alternatives under the criterion. *The Ideal (performance) synthesis mode should be used when the decision maker is concerned with how well each alternative performs* relative to a fixed benchmark. In order for dominance to be an issue the decision-maker should regard inferior alternatives as relevant even after the ranking process is completed. This suggests a simple test for the use of the Distributive mode: *if the decision maker indicates that the preference for a top ranked alternative under a given criterion would improve if the performance of any lower ranked alternative was adjusted downward, then one should use the Distributive synthesis mode.* To make this test more actionable we can ask the decision maker to imagine the amount of money he or she would be willing to pay for the top ranked alternative. If the decision maker would be willing to pay more for a top ranked alternative after learning that the performance of one of the lower-ranked alternatives was adjusted downward, then the Distributive mode should be used.

Consider selecting a car: Two different decision makers may approach the same problem from two different points of views even if the criteria and standards are the same. The one who is interested in «getting a well performing car» should use the Ideal mode. The one who is interested in «getting a car that stands out» among the alternatives purchased by co-workers or neighbors, should use the Distributive mode. The first requires knowledge of the functions which the particular alternative performs and how well it compares with a standard or benchmark. The second requires comparison with the other alternatives to determine its importance.

10. GROUP DECISION MAKING

Here we consider two issues in group decision making. The first is how to aggregate individual judgments, and the second is how to construct a group choice from individual choices.

How to Aggregate Individual Judgments

Let the function $f(x_1, x_2, ..., x_n)$ for synthesizing the judgments given by n judges, satisfy the

(i) Separability condition (S): $f(x_1, x_2, ..., x_n) = g(x_1)g(x_2) ... g(x_n)$ for all $x_1, x_2, ..., x_n$ in an interval P of positive numbers, where g is a function mapping P onto a proper interval J and is a continuous, associative and cancellative operation. [(S) means that the influences of the individual judgments can be separated as above.]

(ii) *Unanimity condition* (U): $f(x, x, ..., x) = x$ for all x in P. [(U) means that if all individuals give the same judgment x, that judgment should also be the synthesized judgment.]

(iii) *Homogeneity condition* (H): $f(ux_1, ux_2, ..., ux_n) = uf(x_1, x_2, ..., x_n)$ where $u > 0$ and x_k, ux_k $(k = 1, 2, ..., n)$ are all in P. [For ratio judgments (H) means that if all individuals judge a ratio u times as large as another ratio, then the synthesized judgment should also be u times as large.]

(iv) *Power conditions* (P_p): $f(x_1^p, x_2^p, ..., x_n^p) = f^p(x_1, x_2, ..., x_n)$. [$(P_2)$ for example means that if the *kth* individual judges the length of a side of a square to be x_k, the synthesized judgment on the area of that square will be given by the square of the synthesized judgment on the length of its side.]

Special case $(R = P_{-1})$: $f(1/x_1, 1/x_2, ..., 1/x_n) = 1/f(x_1, x_2, ..., x_n)$. [($R$) is of particular importance in ratio judgments. It means that the synthesized value of the reciprocal of the individual judgments should be the reciprocal of the synthesized value of the original judgments.]

In this regard, we have the following theorems:

Theorem 16. *The general separable* (S) *synthesizing functions satisfying the unanimity* (U) *and homogeneity* (H) *conditions are the geometric mean and the root-mean-power.* If moreover the reciprocal property (R) is assumed even for a single n-tuple $(x_1, x_2, ..., x_n)$ of the judgments of n individuals, where not all x_k are equal, then only the geometric mean satisfies all the above conditions.

In any rational consensus, those who know more should, accordingly, influence the consensus more strongly than those who are less knowledgeable. Some people are clearly wiser and more sensible in such matters than others, others may be more powerful and their opinions should be given appropriately greater weight. For such unequal importance of voters not all g's in (S) are the same function. In place of (S), the weighted separability property (WS) is now: $f(x_1, x_2, ..., x_n) = g_1(x_1)g_2(x_2) ... g_n(x_n)$. [(WS) implies that not all judging individuals have the same weight when the judgments are synthesized and the different influences are reflected in the different functions $(g_1, g_2, ..., g_n)$.]

Theorem 17. *The general weighted-separable* (WS) *synthesizing functions with the unanimity* (U) *and homogeneity* (H) *properties are the weighted geometric mean* $f(x_1, x_2, ..., x_n) = x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}$ *and the weighted root-mean-powers*

$$f(x_1, x_2, ..., x_n) = \sqrt[\gamma]{q_1 x_1^\gamma + q_2 x_2^\gamma \cdots q_n x_n^\gamma},$$

where $q_1 + q_2 + \cdots q_n = 1$, $q_k > 0$ $(k = 1, 2, ..., n)$, $\gamma \neq 0$, *but otherwise* $q_1, q_2, ..., q_n, \gamma$ *are arbitrary constants.*

If f also has the reciprocal property (R) and for a single set of entries $(x_1, x_2, ..., x_n)$ of judgments of n individuals, where not all x_k are equal, then *only the weighted geometric mean* applies. We give the following theorem which is an explicit statement of the synthesis problem that follows from the previous results, and applies to the second and third cases of the deterministic approach:

Theorem 18. *If* $x_1^{(i)}, ..., x_n^{(i)}$ $i = 1, ..., m$ *are rankings of n alternatives by m independent judges and if* a_i *is the importance of judge i developed from a hierarchy for evaluating the judges, then* $(\prod_{i=1}^{m} x_1^{a_i})^{1/m}, ..., (\prod_{i=1}^{m} x_n^{a_i})^{1/m}$ *are the combined ranks of the alternatives for the m judges.*

The power or priority of judge i is simply a replication of the judgment of that judge (as if there are as many other judges as indicated by his/her power a_i), which implies multiplying his/her ratio buy itself a_i times, and the result follows.

On the Construction of Group Choice from Individual Choices

Given a group of individuals, a set of alternatives (with cardinality greater than 2), and individual ordinal preferences for the alternatives, Arrow proved with his Impossibility Theorem that it is impossible to derive a rational group choice (construct a social choice function that aggregates individual preferences) from ordinal preferences of the individuals that satisfy the following four conditions, i.e., at least one of them is violated:

Decisiveness: the aggregation procedure must generally produce a group order.

Unanimity: if all individuals prefer alternative A to alternative B, then the aggregation procedure must produce a group order indicating that the group prefers A to B.

Independence of irrelevant alternatives: given two sets of alternatives which both include A and B, if all individuals prefer A to B in both sets, then the aggregation procedure must produce a group order indicating that the group, given any of the two sets of alternatives, prefers A to B.

No dictator: no single individual preferences determine the group order.

The main conclusion about group decision making, using the ratio scale approach of the AHP, is that it can be shown that because now individual preferences are cardinal rather than ordinal, it is possible to derive a rational group choice satisfying the above four conditions. It is possible because: a) Individual priority scales can always be derived from a set of pairwise cardinal preference judgments as long as they form at least a minimal spanning tree in the completely connected graph of the elements being compared; and b) The cardinal preference judgments associated with group choice belong to a ratio scale that represents the relative intensity of the group preferences.

11. CONCLUDING REMARKS

All people make decisions and have been making them since the beginning of life on this earth. A good decision theory must uncover this natural part in people and formalize it for general use and for making decisions still better. That is what the AHP is about. It must not require a high degree of technical education to understand how to use the process, because even a lay person should find it familiar and natural. We need to remember that even today there are people in the world who may not know about numbers at all, and are still able to make decisions with their feelings without resorting to the use of numbers. That is why the fundamental scale of the AHP allows the use of words and feelings that correspond to numbers as an abstraction.

Measurement is quantitative information useful for discriminating among magnitudes and among orders of magnitudes. Numerical discrimination is different from cognitive discrimination. Creativity and understanding are linked to our cognitive ability and not to our ability to make precise measurements. It is rare that extreme precision is needed for any sort of understanding and discrimination. It is the way we are made to create understanding. The more precise we are, we still need to be more precise, but the complexity of the world brings in new information that nullifies that precision and requires new precision and so on. Even in science measurement and precision are subject to interpretation. It is the goals we pursue that need to be served and we are in control of the importance and meaningfulness of these goals as they serve our well being and survival. Precision in the preparation of drugs is necessary, but there is such flexibility that the same size pill is prescribed for all adults regardless of the size of their bodies. Precision in designing the gears of a clock is mandatory, but precision in time is one and only one aspect of experience that may have to be traded off with other factors. In fact, time is subjective and what is considered good in punctuality by some may be regarded as some kind of militancy by others. Strict punctuality is a human normative invention not respected in the biology of cells and in birth and death. The question is whether we can access the world directly and satisfactorily with the very judgments we use to evaluate measurement. Note that if we have several criteria measured on the same absolute scale, we must deal with them in a particular way through grouping and normalization, in order to obtain the correct outcome one obtains by multiplying and adding numbers. One needs to keep this in mind in going back and forth from absolute to relative scales on many criteria.

A generalization of the theory to dependence and feedback appears in a book by the author called *The Analytic Network process* (ANP).

REFERENCES

1. Aczel, Janos D. and Marek Kuczma. (1991). *Generalizations of a «Folk-Theorem»: On Simple Functional Equations in a Single Variable.* Results in Mathematics **19,** 5-21. Basel.
2. Berge, C. (1962). *Theory of Graphs and Its Applications.* Wiley, New York.

3. Birkhoff G. (1957). Extensions of Jentzsch's Theorem. *Trans. Amer. Math. Soc.* **85,** 219-227.
4. Blocksma, M. (1989). *Reading the Numbers.* Penguin Group, Viking-Penguin Inc., New York.
5. Blumenthal, A. L. (1977). *The Process of Cognition.* Prentice Hall.
6. DeTurck, D. M. (1987). The Approach to Consistency in The Analytic Hierarchy Process. *Mathematical Modelling* **9**/3-5, 345-352.
7. Franklin, J. N. (1968). *Matrix Theory.* Prentice-Hall, Englewood Cliffs, NJ.
8. Golden, B. L. and Wang, Q. (1989). *An Alternate Measure of Consistency.* In The Analytic Hierarchy Process-Applications and Studies, eds, Golden, Wasil and Harker, Springer-Verlag, New York.
9. Harker, P. T. (1987). Alternative modes of questioning in the analytic hierarchy process. *Mathematical Modeling* **9**, 353-360.
10. Kohlberg, E. and John W. Pratt. (1982). The Contraction Mapping Approach to the Perron Frobenius Theory: Why Hilbert's Metric? *Mathematics of Operations Research* **7**/**2**.
11. Kuffler, Stephen, W. and Nicholls, J. G. (1976). *From Neuron to Brain.* Sunderland, MA: Sinauer Associates.
12. Lancaster, P. and Miron Tismenetsky. (1985). *Linear Transformations and Matrices.* The Theory of Matrices Second Edition: With Applications, Academic Press, Inc., Orlando.
13. Lebesgue, H. (1928). *Leçons sur l'integration.* 2nd ed., Gauthier-Villars, Paris.
14. Miller, G. A. (1956). The Magical Number Seven Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review* **63**, 81-97.
15. Saaty, T. L. (1990). *Multicriteria Decision Making: The Analytic Hierarchy Process.* Revised and reprinted, RWS Publications, 4922 Ellsworth Ave, Pittsburgh, PA. Original version published by McGraw-Hill, 1980.
16. Saaty, T. L. (1994). Fundamentals of Decision Making and Priority Theory, RWS Publications, 4922 Ellsworth Ave., Pittsburgh, PA.
17. Saaty, T. L. (2000). *The Brain: Unraveling the Mystery of How it Works.* RWS Publications, 4922 Ellsworth Ave., Pittsburgh, PA.
18. Saaty, T. L. and Vargas, L. G. (1993). A Model of Neural Impulse Firing and Synthesis. *Journal of Mathematical Psychology* **37**, 200-219.
19. Stewart, G. W. (1973). *Eigenvalues and Eigenvectors. Introduction to Matrix Computations.* Academic Press, Inc., Orlando.
20. Vargas, L. G. (1983). *Analysis of Sensitivity of Reciprocal Matrices.* Applied Mathematics and Computation **12**, 301-320, Elsevier Science Publishing Co., Inc., New York.
21. Wei, T. H. (1952). *The Algebraic Foundations of Ranking Theory.* Thesis, Cambridge, MA.
22. Wilkinson, J. H. (1965). *The Algebraic Eigenvalue Problem.* Clarendon Press, Oxford.