On the Measurement of Intangibles. A Principal Eigenvector Approach to Relative Measurement Derived from Paired Comparisons

Thomas L. Saaty

Introduction

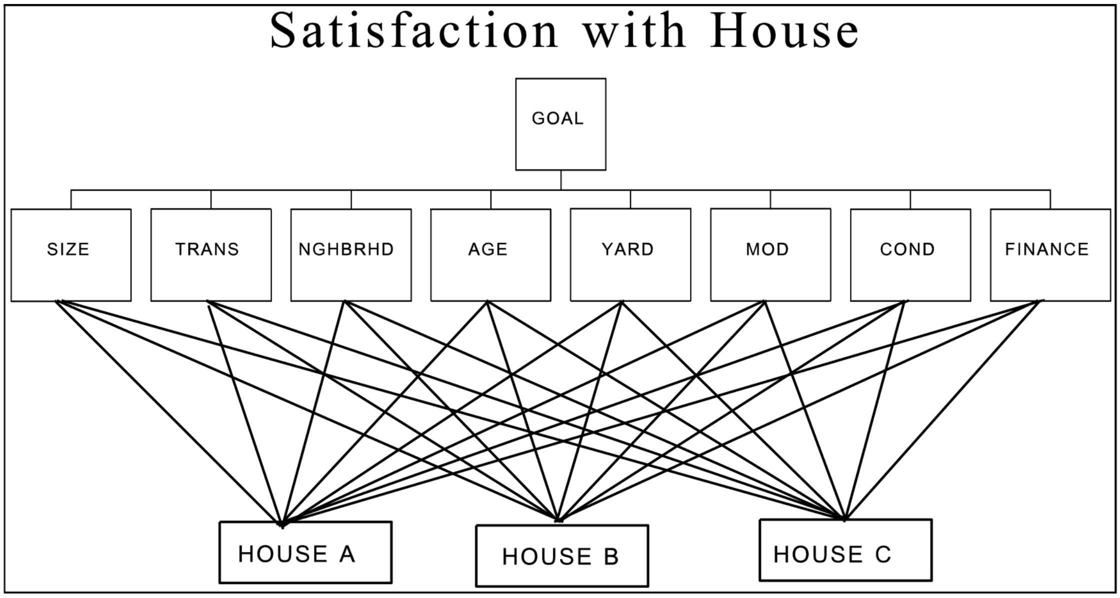
Nearly all of us have been brought up to believe that clear-headed logical thinking is our only sure way to face and solve problems. But experience suggests that logical thinking is not natural to us. Indeed, we have to practice, and for a long time, before we can do it well. Since complex problems usually have many related factors, traditional logical thinking leads to sequences of ideas so tangled that the best solution cannot be easily discerned.

For a very long time people believed and argued strongly that it is impossible to express the intensity of human feelings with numbers. The epitome of such a belief was expressed by A. F. MacKay who writes [12] that pursuing the cardinal approaches is like chasing what cannot be caught. It was also expressed by Davis and Hersh [5]: "If you are more of a human being, you will be aware there are such things as emotions, beliefs, attitudes, dreams, intentions, jealousy, envy, yearning, regret, longing, anger, compassion and many others. These things—the inner world of human life—can never be mathematized." In their book [11] Lawrence LeShan and Henry Margenau write: "We cannot as we have indicated before, quantify the observables in the domain of consciousness. There are no rules of correspondence possible that would enable us to quantify our feelings. We can make statements of the relative intensity of feelings, but we cannot go beyond this. I can say: I feel angrier at him today than I did yesterday. We cannot, however, make meaningful statements such as, I feel three and one half times angrier than I did yesterday.... The physicists' schema, so faithfully emulated by generations of psychologists, epistemologists and aestheticians, is probably blocking their progress, defeating possible insights by its prejudicial force. The schema is not false—it is perfectly reasonable—but it is bootless for the study of mental phenomena."

The Nobel Laureate Henri Bergson [4] writes: "But even the opponents of psychophysics do not see any harm in speaking of one sensation as being more intense than another, of one effort as being greater than another, and in thus setting up differences of quantity between purely internal states. Common sense, moreover, has not the slightest hesitation in giving its verdict on this point; people say they are more or less warm, or more or less sad, and this distinction of more and less, even when it is carried over to the region of subjective facts and unextended objects, surprises nobody."

Thomas L. Saaty holds the Chair of Distinguished University Professor at the University of Pittsburgh, with appointments in several departments. His email address is saaty@katz.pitt.edu.

DOI: http://dx.doi.org/10.1090/noti944



The criteria important to the family are:

1. Size of House: Storage space; size of rooms; number of rooms; total area of house.
2. Transportation: Convenience and proximity of bus service.
3. Neighborhood: Degree of traffic, security, taxes, physical condition of surrounding buildings.
4. Age of House: How long ago house was built.
5. Yard Space: Front, back, and side space, and space shared with neighbors.
6. Modern Facilities: Dishwasher, garbage disposal, air conditioning, alarm system, and other such items.
7. General Condition: Extent to which repairs are needed; condition of walls, carpet, drapes, wiring; cleanliness.
8. Financing: Availability of assumable mortgage, seller financing, or bank financing.

The next step is to make comparative judgments. The family assesses the relative importance of all possible pairs of criteria with respect to the overall goal, Satisfaction with House, coming to a consensus judgment on each pair; and their judgments are arranged into a matrix. The question to ask when comparing two criteria is: which is more important and how much more important is it with respect to satisfaction with a house?

The matrix of pairwise comparison judgments on the criteria given by the home-buyers in this case is shown in Table 1. The judgments are entered using the fundamental scale of the AHP Table 0: a criterion compared with itself is always assigned the value 1 so the main diagonal entries of the pairwise comparison matrix are all 1. We are permitted to interpolate values between the integers, if desired. Reciprocal values are automatically entered in the transpose position, so the family must make a total of 28 pairwise judgments.

Figure 1. Decomposition of the problem into a hierarchy.

If we were to ask what in practical terms measurement means, one would most likely propose a scale that is applied to measure objects: a set of numbers, a set of objects and a mapping from the objects to the numbers. Then we can agree that appropriate judgment must be used to interpret the scale readings and use them in practice. So judgment is also essential. But there is another way to think of a scale.

Henri Lebesgue, who was concerned with questions of measure theory and measurement, wrote [10]:

> "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,

Intensity of Importance	*Definition*	*Explanation*
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgment slightly favor one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgment strongly favor one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favored very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favoring one activity over another is of the highest possible order of affirmation
1.1-1.9	When activities are very close a decimal is added to 1 to show their difference as appropriate	A better alternative way to assigning the small decimals is to compare two close activities with other widely contrasting ones, favoring the larger one as needed a little value over the smaller one using the 1-9 values.
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 logical assumption
Measurements from ratio scales		When it is desired to use such numbers in physical applications. Alternatively, often one estimates the ratios of such magnitudes by using judgment

Table 0. Fundamental scale of absolute numbers.

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."

Lebesgue did not go far enough in examining why we have to compare. When we deal with intangible factors, which by definition have no scales of measurement, we can compare them in pairs. Making comparisons is a talent we all have. Not only can we indicate the preferred object, but we can also discriminate among intensities of preference. The philosopher, Arthur Schopenhauer [25], said, "Every **truth is the reference** of a judgment to something outside it, and **intrinsic truth** is a contradiction."

A common kind of decision problem we face is something like this: which house to buy? The different houses being considered have some attributes or criteria in common that are important to the decision maker. If one house were best on every criterion, the choice would be easy, but usually the house that is the best on one criterion (e.g., cost) is worst on another (e.g., size). How should one make the tradeoff? We describe and discuss a mathematical model, the Analytic Hierarchy Process (AHP), not a prescriptive (normative) but a *descriptive* psychophysical process that can be used to make such decisions by dealing with the measurement of intangibles using human judgment. Intangibles can be nonphysical influences that are passing and very transient. No conceivable instrument can be devised to measure them other than the mind itself, which must also interpret their meaning. Intangibles leave an impact on our minds, which are biologically endowed to respond to influences,

	Size	Trans.	Nbrhd.	Age	Yard	Modern	Cond.	Finance	Normalized Priority Vector w
Size	1	5	3	7	6	6	1/3	1/4	.175
Trans.	1/5	1	1/3	5	3	3	1/5	1/7	.062
Nbrhd.	1/3	3	1	6	3	4	1/2	1/5	.103
Age	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8	.019
Yard	1/6	1/3	1/3	3	1	1/2	1/5	1/6	.034
Modern	1/6	1/3	1/4	4	2	1	1/5	1/6	.041
Cond.	3	5	2	7	5	5	1	1/2	.221
Finance	4	7	5	8	6	6	2	1	.345

$\lambda_{max} = 8.811$, Consistency Ratio (C.R.) =.083

Table 1. The family's pairwise comparison matrix for the criteria.

Size of House	A	B	C	Distributive Priorities	Idealized Priorities
A	1	5	9	.743	1.000
B	1/5	1	4	.194	0.261
C	1/9	1/4	1	.063	0.085

C.R. = .07

Yard Space	A	B	C	Distributive Priorities	Idealized Priorities
A	1	6	4	.691	1.000
B	1/6	1	1/3	.091	0.132
C	1/4	3	1	.218	0.315

C.R. = .05

Transportatio n	A	B	C	Distributive Priorities	Idealized Priorities
A	1	4	1/5	.194	0.261
B	1/4	1	1/9	.063	0.085
C	5	9	1	.743	1.000

C.R. = .07

Modern Facilities	A	B	C	Distributive Priorities	Idealized Priorities
A	1	9	6	.770	1.000
B	1/9	1	1/3	.068	0.088
C	1/6	3	1	.162	0.210

C.R. = .05

Neighborhood	A	B	C	Distributive Priorities	Idealized Priorities
A	1	9	4	.717	1.000
B	1/9	1	1/4	.066	0.092
C	1/4	4	1	.217	0.303

C.R. = .04

General Condition	A	B	C	Distributive Priorities	Idealized Priorities
A	1	1/2	1/2	.200	0.500
B	2	1	1	.400	1.000
C	2	1	1	.400	1.000

C.R. = .00

Age of House	A	B	C	Distributive Priorities	Idealized Priorities
A	1	1	1	.333	1.000
B	1	1	1	.333	1.000
C	1	1	1	.333	1.000

C.R. = .00

Financing	A	B	C	Distributive Priorities	Idealized Priorities
A	1	1/7	1/5	.072	0.111
B	7	1	3	.650	1.000
C	5	1/3	1	.278	0.430

C.R. = .06

Table 2. Pairwise comparison matrices for the alternative houses.

	Size (.175)	Trans (.062)	Nghbd (.103)	Age (.019)	Yard (.034)	Modrn (.041)	Cond (.221)	Finance (.345)	Composite priority vector
									Distributive Mode
A	.743	.194	.717	.333	.691	.770	.200	.072	.346
B	.194	.063	.066	.333	.091	.068	.400	.649	.369
C	.063	.743	.217	.333	.218	.162	.400	.279	.285

									Ideal Mode
A	1.00	.261	1.00	1.00	1.00	1.00	.500	.111	.315
B	.261	.085	.092	1.00	.132	.088	1.00	1.00	.383
C	.085	1.00	.303	1.00	.315	.210	1.00	.430	.302

Table 3. Distributive and ideal synthesis.

by making comparisons both consciously and subconsciously. It is a way of measurement that took place long before Nicole Oresme, Pierre de Fermat, and finally René Descartes more rigorously introduced general coordinate systems for physical measurements and assumed that they were extensible from zero to infinity by using an *arbitrarily chosen* unit applied *uniformly* over the entire range of measurement. Taking ratios removes the arbitrariness of the unit and creates relative absolute scales, invariant under the identity transformation. The example below is more staid than dealing with evanescent phenomena like political decisions, but it serves to illustrate the ideas whose mathematical foundations have been developed in a separate paper and in books.

Choosing the Best House

Consider the following (hypothetical) example: a family wishing to purchase a house identifies eight criteria that are important to them. The problem is to select one of three candidate houses. The first step is to structure the problem into a hierarchy (see Figure 1). On the first (top) level is the overall goal of Satisfaction with House. On the second level are the eight criteria that contribute to the goal, and on the third (bottom) level are the three candidate houses that are to be evaluated by considering the criteria on the second level.

The criteria important to the family are:

1. Size of House: Storage space; size of rooms; number of rooms; total area of house.
2. Transportation: Convenience and proximity of bus service.
3. Neighborhood: Degree of traffic, security, taxes, physical condition of surrounding buildings.
4. Age of House: How long ago the house was built.
5. Yard Space: Front, back, and side space and space shared with neighbors.
6. Modern Facilities: Dishwasher, garbage disposal, air conditioning, alarm system, and other such items.
7. General Condition: Extent to which repairs are needed; condition of walls, carpet, drapes, wiring; cleanliness.
8. Financing: Availability of assumable mortgage, seller financing or bank financing.

The next step is to make comparative judgments. The family assesses the relative importance of all possible pairs of criteria with respect to the overall goal, Satisfaction with House, coming to a consensus judgment on each pair; and their judgments are arranged into a matrix. The questions to ask when comparing two criteria are, which is more important and how much more important is it with respect to satisfaction with a house?

The matrix of pairwise comparison judgments on the criteria given by the home-buyers in this case is shown in Table 1. The judgments are entered using the fundamental scale of the Analytic Hierarchy Process (AHP) Table 0: a criterion compared with itself is always assigned the value 1 so the main diagonal entries of the pairwise comparison matrix are all 1. We are permitted to interpolate values between the integers, if desired. Reciprocal values are automatically entered in the transpose position, so the family must make a total of twenty-eight pairwise judgments.

We have assumed that an element with weight zero is eliminated from comparison because zero can be applied to the whole universe of factors not included in the discussion.

The foregoing integer-valued fundamental scale of response used in making paired comparison judgments can be derived from the logarithmic response function of Weber Fechner in psychophysics as follows. For a given value of the stimulus, the magnitude of response remains the same until the value of the stimulus is increased sufficiently large in proportion to the value of the stimulus, thus preserving the proportionality of relative increase in stimulus for it to be detectable for a new response. This suggests the idea of just noticeable differences (jnd), well known in psychology. Thus, starting with a stimulus s_0, successive magnitudes of the new stimuli take the form [2]

$$\begin{aligned} s_1 &= s_0 + \Delta s_0 = s_0 + \frac{\Delta s_0}{s_0} s_0 = s_0(1+r), \\ s_2 &= s_1 + \Delta s_1 = s_1(1+r) = s_0(1+r)^2 \equiv s_0\alpha^2, \\ &\vdots \\ s_n &= s_{n-1}\alpha = s_0\alpha^n \quad (n = 0, 1, 2, \dots), \\ \frac{\Delta s_0}{s_0} &= \frac{\Delta s_1}{s_1} = \frac{\Delta s_2}{s_2} = \cdots . \end{aligned}$$

We consider the responses to these stimuli to be measured on a ratio scale ($b = 0$). A typical response has the form $M_i = a \log \alpha^i$, $i = 1, \dots, n$, or one after another they have the form

$$M_1 = a \log \alpha,\ M_2 = 2a \log \alpha, \dots, M_n = na \log \alpha.$$

We take the ratios M_i/M_1, $i = 1, \dots, n$, of these responses in which the first is the smallest and serves as the unit of comparison, thus obtaining the *integer* values $1, 2, \dots, n$ of the fundamental scale of the AHP. It appears that numbers are intrinsic to our ability to make comparisons and that they were not an invention by our primitive ancestors. We must be grateful to them for the discovery of the symbolism. In a less mathematical vein, we note that we are able to distinguish ordinally between high, medium, and low at one level and for each of them in a second level below that also to distinguish between high, medium, and low, giving us nine different categories. We assign the value one to (low, low), which is the smallest, and the value nine to (high, high), which is the highest, thus covering the spectrum of possibilities between two levels and giving the value nine for the top of the paired comparisons scale as compared with the lowest value on the scale. The scale consists of absolute numbers, which unlike the familiar numbers that belong to a ratio scale that are invariant under a similarity transformation (multiplication by a positive number) are invariant under the identity transformation.

In his book *The Number Sense: How the Mind Creates Mathematics* [6], the mathematician and cognitive neuropsychologist Stanislas Dehaene writes more than twenty-five years after we derived this scale that "Introspection suggests that we can mentally represent the meaning of numbers 1 through 9 with actual acuity. Indeed, these symbols seem equivalent to us. They all seem equally easy to work with, and we feel that we can add or compare any two digits in a small and fixed amount of time like a computer. In summary, the invention of numerical symbols should have freed us from the fuzziness of the quantitative representation of numbers."

It can be shown that when comparisons involve greater contrast than 9, the elements can be aggregated into clusters, with a pivot element from one cluster to an adjacent one, again applying the same kind of comparisons within the next cluster using the scale 1–9. One then divides by the weight of the pivot element in the second cluster and multiplies by its weight from the first cluster, and now the priorities in the two clusters are comparable and can be put together and so on. This type of clustering is called homogeneous clustering [22]. In addition, as we shall see later, for stability of the priorities with respect to small changes in judgment, each cluster must not contain more than a few elements: about seven [20].

In the AHP model, the *vector of priorities* for the criteria is obtained by computing the principal eigenvector, the classical Perron vector, of the pairwise comparison matrix. The Perron vector is a necessary condition for obtaining the priorities when the judgments are inconsistent. As we shall see below, consistency and near consistency are important concepts in our considerations. Because the pairwise comparison matrix has positive entries, Perron's theorem ensures that there is a unique positive vector (denoted by w) whose entries sum to one that is an eigenvector of the pairwise comparison matrix, associated with an eigenvalue (denoted by $\lambda_{\max}$) of strictly largest modulus. That eigenvalue, the Perron eigenvalue, is positive and algebraically simple (multiplicity one as a root of the characteristic equation) [14]. Consistency of the family's set of judgments is measured by the *consistency ratio* (C.R.), which we explain later.

Table 1 shows that size dominates transportation strongly since a 5 appears in the (size, transportation) position. In the (finance, size) position we have a 4, which means that finance is between moderately and strongly more important than size. The priority vector shows that financing is the most important criterion to the family as the entry of w corresponding to finance has the largest value, 0.345.

Consistency, which was alluded to previously, is an elaboration of the common sense view expressed in this statement: if you prefer spring to summer by 2, summer to winter by 3, and spring to winter by 6, then those three judgments are consistent.

The family's next task is to compare the houses in pairs with respect to how much better (more dominant) one is than the other in satisfying each of the eight criteria. There are eight 3-by-3 matrices of judgments since there are eight criteria and three houses are to be compared for each criterion. The matrices in Table 2 contain the judgments of the family. In order to facilitate understanding of the judgments, we give a brief description of the houses.

In Table 1 an element on the left of the matrix is compared for dominance over another at the top.

House A: This house is the largest. It is located in a good neighborhood with little traffic and low taxes. Its yard space is larger than that of either house B or C. However, its general condition is not very good, and it needs cleaning and painting. It would have to be bank financed at a high interest rate.

House B: This house is a little smaller than house A and is not close to a bus route. The neighborhood feels insecure because of traffic conditions. The yard space is fairly small, and the house lacks basic modern facilities. On the other hand, its general condition is very good, and it has an assumable mortgage with a rather low interest rate.

House C: House C is very small and has few modern facilities. The neighborhood has high taxes but is in good condition and seems secure. Its yard is bigger than that of house B but smaller than house A's spacious surroundings. The general condition of the house is good, and it has a pretty carpet and drapes. The financing is better than for house A but poorer than for house B.

In Table 2 both ordinary (distributive) and idealized priority vectors of the three houses are given for each of the criteria. The idealized priority vector is obtained by dividing each element of the distributive priority vector by its largest element. The composite priority vector for the houses is obtained by multiplying each priority vector by the priority of the corresponding criterion, adding across all the criteria for each house and then normalizing. When we use the (ordinary) distributive priority vectors, this method of synthesis is known as the *distributive mode* and yields A = .345, B = .369, and C = .285. Thus house B is preferred to houses A and C in the ratios: .369/.346 and .369/.285, respectively.

In Table 2 again an element on the left of each matrix is compared for dominance over another at the top. If the top element is dominant, a fraction is entered and its inverse, a whole number, is entered in the reciprocal position.

When we use the idealized priority vector, the synthesis is called the *ideal mode*. This yields A = .315, B = .383, C = .302 and B is again the most preferred house. The two ways of synthesizing are shown in Table 3. They need not yield the same ranking. In general, the ideal mode is used when rating the alternatives one at a time (see later) with respect to the criteria and when the criteria priorities are independent of the alternatives. The ideal mode is used to force rank preservation when new houses are added to the collection by only comparing them with respect to the ideal and not the other alternatives. But always preserving rank is not always desirable [16].

The Pairwise Comparison Matrix

In comparing pairs of criteria with respect to the goal, one estimates which of the two criteria is more important and how much more. The result of these comparisons is arranged in a positive matrix $A = [a_{ij}]$ whose entries satisfy the reciprocal property $a_{ji} = 1/a_{ij}$.

We start with a positive reciprocal matrix such as Table 1. In an n-by-n table, $n(n-1)/2$ judgments must be made, which is why the house-buying family had to make $(8 \times 7)/2 = 28$ judgments. These judgments are made *independently*, but they are not really "independent". If the family feels that financing is twice as important as size and that size is twice as important as age, for consistency of judgments we should *expect* them to feel that financing is four times as important as age. The mathematical expression of our expectation is the set of identities

$$a_{ij} = a_{ik}/a_{jk} \quad \text{for all } i, j, k = 1, \ldots, n$$

among the entries of a *consistent* pairwise comparison matrix $A = [a_{ij}]$. Of course, real-world pairwise comparison matrices are very unlikely to be consistent, and we address the consequences of that reality next [16], [19].

Why the Principal Eigenvector?

Suppose a positive square matrix $A = [a_{ij}]$ is consistent. Then A must have unit diagonal entries, since $a_{ii} = a_{ik}/a_{ik}$, for all $i, k = 1, \ldots, n$. Moreover, A must be *reciprocal* since $a_{ij}a_{ji} = 1$ means that $a_{ij} = 1/a_{ji}$. Such a matrix has a very simple structure since $a_{ik} = a_{i1}a_{1k} = a_{i1}/a_{k1}$ for all i, k. Thus the entries in the first column of A determine all other entries! For convenience, write $\alpha_i \equiv a_{i1}$, so that $A = [a_{ij}] = [\alpha_i/\alpha_j]$. If we define the two positive n-by-1 vectors $x \equiv [\alpha_i]$ and $y \equiv [(\alpha_i)^{-1}]$, then it is clear that $A = xy^T$ has rank one. Thus, the positive matrix A has one nonzero eigenvalue and $n-1$ zero eigenvalues. It is easy to check that $Ax = \sum_{j=1}^{n}(\alpha_i/\alpha_j)\alpha_j = [n\alpha_i] = nx$, so the nonzero eigenvalue of A (its Perron eigenvalue, which we have denoted by $\lambda_{\max}$) is n, and an associated positive right eigenvector is $x \equiv [\alpha_i]$. If we set $c \equiv \alpha_1 + \cdots + \alpha_n$, the Perron vector of A (its unique positive eigenvector whose entries sum to one) may be written as $w \equiv x/c = [\alpha_i/c] \equiv [w_i]$. The Perron vector determines all the entries of $A : A = [a_{ij}] = [\alpha_i/\alpha_j] = [(\alpha_i/c)/(\alpha_j/c)] = [w_i/w_j]$. When a matrix is consistent, right and left eigenvectors have reciprocal corresponding entries.

We know that an n-by-n positive consistent matrix $A = [a_{ij}]$ has a unique positive eigenvector $w \equiv [w_i]$ (its Perron vector) whose entries sum to one and whose corresponding eigenvalue (its Perron eigenvalue) is n. Moreover, the ratios of the entries of w are precisely the entries of $A : a_{ij} = w_i/w_j$. If we think of A as a matrix of (perfectly) consistent pairwise comparisons for n given elements, then the n values w_i are a natural set of priorities that underlie the set of pairwise judgments: $a_{ij} = w_i/w_j$. There are several ways to prove that the principal eigenvector is necessary when the judgments are inconsistent [18], [21].

The foregoing discussion is intended to motivate the central and critical choice of the Perron vector as the means to extract a vector of priorities from a given pairwise comparison matrix in the AHP model. If humans made perfectly consistent judgments all the time, the model would be perfect. But they do not, so we must now face the question of assessing the deviation from consistency of an actual pairwise comparison matrix and the consequences of inconsistency for the quality of decisions made according to the AHP model. In passing, we observe that if humans were always perfectly consistent, they would not be able to learn new things that modify or change the relations among what they knew before and they would be like robots. But there is a level of tolerable inconsistency that we must allow beyond which the judgments would appear to be uninformed, random, or arbitrary.

When we compare things, unlike assigning them numbers independently of one another, their priorities always depend on what other things they are compared with. Were we to assume that the universe is interdependent (the subject of the Analytic Network Process (ANP) [17], [18]) with a complicated field of influences in every aspect, our traditional way of assigning numbers from scales of measurement would not be "the natural way" to determine their importance. More and more we are finding with this relative way of thinking that the world is different (has different rank orders) than we think we understand it to be today. Another useful observation is that comparisons are necessary for comparing criteria to derive their priorities because there are no scales for their measurement and also because their importance varies from decision to decision. In the ANP the criteria are compared with respect to each alternative and the alternatives are compared with respect to each criterion to derive their interdependent priorities.

Regrettably, the three laws of thought (identity, excluded middle, and contradiction), known to Plato and Aristotle and even to Leibniz, that are strict requirements we all adhere to in language, logic, science, and mathematics have precluded comparisons, our biological heritage. Without comparisons nothing can be known in and of itself without also knowing other things with which it is compared, including knowing it at an earlier time, so we can ensure it is the one we have in mind (the law of identity), thus recognizing it. Arthur Schopenhauer, who was not equipped to develop a mathematical theory to use comparisons, listed the laws of thought by adding a fourth one in his *On the Fourfold Root of the Principle of Sufficient Reason.* (1) a subject is equal to the sum of its predicates, or $a = a$; (2) no predicate can be simultaneously attributed and denied to a subject, or $a \neq\sim a$; (3) of every two contradictorily opposite predicates one must belong to every subject; and (4) **truth is the *reference of a judgment to something outside* itself as its sufficient reason or ground.**

We are all familiar with the arbitrarily imposed axiom in logic and mathematics that if A dominates B and B dominates C, then A must dominate C. But in the real world team A beats team B, B beats C but C beats A, contradicting theory. It appears that theory needs to be changed to accommodate reality.

When is a Positive Reciprocal Matrix Consistent?

Let $A = [a_{ij}]$ be an n-by-n positive reciprocal matrix, so all $a_{ii} = 1$ and $a_{ij} = 1/a_{ji}$ for all $i, j = 1, \ldots, n$. Let $w = [w_i]$ be the Perron vector of A, let $D = \operatorname{diag}(w_1, \ldots, w_n)$ be the n-by-n diagonal matrix whose main diagonal entries are the entries of w, and set $E = D^{-1}AD = [a_{ij}w_j/w_i] = [\varepsilon_{ij}]$. Then E is similar to A and is a positive reciprocal matrix since $\varepsilon_{ij} = a_{ji}w_i/w_j = (a_{ij}w_j/w_i)^{-1} = 1/\varepsilon_{ij}$. Moreover, all the row sums of E are equal to the Perron eigenvalue of A:

$$\sum_{j=1}^{n} \varepsilon_{ij} = \sum_j a_{ij}w_j/w_i = [Aw]_i/w_i = \lambda_{\max} w_i/w_i = \lambda_{\max}.$$

The computation

(1)
$$n\lambda_{\max} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\varepsilon_{ij}\right) = \sum_{i=1}^{n}\varepsilon_{ii} + \sum_{\substack{i,j=1\\ i\neq j}}^{n}(\varepsilon_{ij}+\varepsilon_{ji}) = n + \sum_{\substack{i,j=1\\ i\neq j}}^{n}(\varepsilon_{ij}+\varepsilon_{ij}^{-1}) \geq n + 2(n^2-n)/2 = n^2$$

reveals that $\lambda_{\max} \geq n$. Moreover, since $x + 1/x \geq 2$ for all $x > 0$, with equality if and only if $x = 1$, we see that $\lambda_{\max} = n$ if and only if all $\varepsilon_{ij} = 1$, which is equivalent to having all $a_{ij} = w_i/w_j$.

The foregoing arguments show that a positive reciprocal matrix A has $\lambda_{\max} \geq n$, with equality if and only if A is consistent. As our measure of deviation of A from consistency, we choose the *consistency index*

$$\mu \equiv \frac{\lambda_{\max} - n}{n-1}.$$

We have seen that $\mu \geq 0$ and $\mu = 0$ if and only if A is consistent. These two desirable properties explain the term "n" in the numerator of μ; what about the term "$n - 1$" in the denominator? Since trace $A = n$ is the sum of all the eigenvalues of A, if we denote the eigenvalues of A that are different from $\lambda_{\max}$ by $\lambda_2, \ldots, \lambda_{n+1}$, we see that $n = \lambda_{\max} + \sum_{i=2}^{n}\lambda_i$, so $n - \lambda_{\max} = \sum_{i=2}^{n}\lambda_i$ and $\mu = -\frac{1}{n-1}\sum_{i=2}^{n}\lambda_i$ is the negative average of the non-Perron eigenvalues of A.

It is an easy, but instructive, computation to show that $\lambda_{\max} = 2$ for every 2-by-2 positive reciprocal matrix:

$$\begin{bmatrix} 1 & \alpha \\ \alpha^{-1} & 1 \end{bmatrix}\begin{bmatrix} 1+\alpha \\ (1+\alpha)\alpha^{-1} \end{bmatrix} = 2\begin{bmatrix} 1+\alpha \\ (1+\alpha)\alpha^{-1} \end{bmatrix}.$$

Thus, every 2-by-2 positive reciprocal matrix is consistent.

Not every 3-by-3 positive reciprocal matrix is consistent, but in this case we are fortunate to have again explicit formulas for the Perron eigenvalue and eigenvector. For

$$A = \begin{bmatrix} 1 & a & b \\ 1/a & 1 & c \\ 1/b & 1/c & 1 \end{bmatrix},$$

we have $\lambda_{\max} = 1 + d + d^{-1}$, $d = (ac/b)^{1/3}$, and

(2)
$$w_1 = bd\Big/\left(1 + bd + \frac{c}{d}\right), \quad w_2 = c/d\left(1 + bd + \frac{c}{d}\right), \quad w_3 = 1\Big/\left(1 + bd + \frac{c}{d}\right).$$

Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
R.I.	0	0	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49	1.52	1.54	1.56	1.58	1.59
First Order Differences		0	0.52	0.37	0.22	0.14	0.10	0.05	0.05	0.04	0.03	0.02	0.02	0.02	0.01

Table 4. Random index.

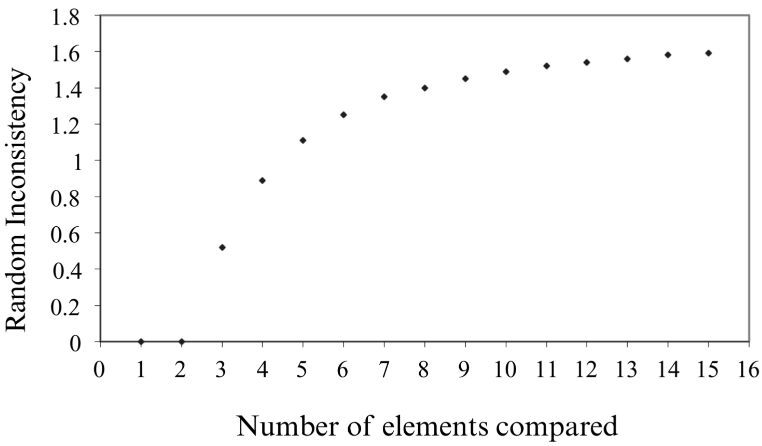


Figure 2. Plot of random inconsistency.

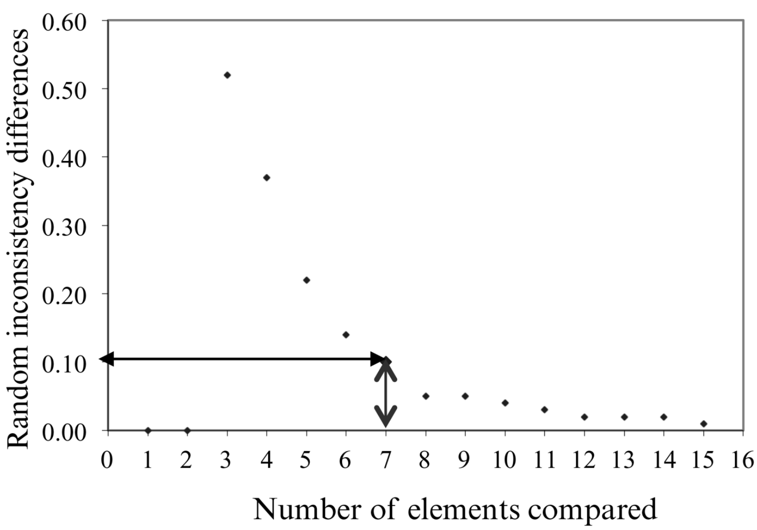


Figure 3. Plot of first differences in random inconsistency—7 is critical.

Note that $\lambda_{\max} = 3$ when $d = 1$ or $c = b/a$, which is true if and only if A is consistent.

In order to get some feel for what the consistency index might be telling us about a positive n-by-n reciprocal matrix A, consider the following simulation: choose the entries of A above the main diagonal at random from the seventeen values $\{1/9, 1/8, \dots, 1/2, 1, 2, \dots, 8, 9\}$. Then fill in the entries of A below the diagonal by taking reciprocals. Put ones down the main diagonal and compute the consistency index. Do this 50,000 times and take the average, which we call the random index. Table 4 shows the values obtained from one set of such simulations for matrices of size $1, 2, \dots, 10$.

Figure 2 is a plot of the first two rows of Table 4. It shows the asymptotic nature of random inconsistency.

The third row of Table 2 gives the differences between successive numbers in the second row. Figure 3 is a plot of these differences and shows the importance of the number seven as a cutoff point beyond which the differences are less than 0.10, where we are not sufficiently sensitive to make accurate changes in judgment on several elements simultaneously.

Since it would be pointless to try to discern any priority ranking from a set of random comparison judgments, we should probably be uncomfortable about proceeding unless the consistency index of a pairwise comparison matrix is very much smaller than the corresponding random index value in Table 4. The *consistency ratio* (C.R.) of a pairwise comparison matrix is the ratio of its consistency index (C. I.) to the corresponding random index value in Table 4.

As a rule of thumb, we do not recommend proceeding if the consistency ratio is more than about 0.10 for $n > 4$. For $n = 3$ and 4 we recommend that the C.R. be less than 0.05 and 0.09, respectively. Thus in general when asked, we require that C.R. not exceed 0.10 by much. How do we explain this outcome in general?

The notion of order of magnitude is essential in any mathematical consideration of changes in measurement. When one has a numerical value say between 1 and 10 for some measurement and one wants to determine whether a change in this value is significant or not, one reasons as follows: a change of a whole integer value is critical because it changes the magnitude and identity of the original number significantly. If the change or perturbation in value is of the order of a percent or less, it would be so small (by two orders of magnitude) and would be considered negligible. However, if this perturbation is a decimal (one order of magnitude smaller), we are likely to pay attention to modify the original value by this decimal without losing the

	Size	Trans	Nbrhd	Age	Yard	Modern	Cond	Finance	w	v
Size	1	5	3	7	6	6	1/3	1/4	.173	.047
Trans.	1/5	1	1/3	5	3	3	1/5	1/7	.054	.117
Nbrhd.	1/3	3	1	6	3	4	6	1/5	.188	.052
Age	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8	.018	.349
Yard	1/6	1/3	1/3	3	1	1/2	1/5	1/6	.031	.190
Modern	1/6	1/3	1/4	4	2	1	1/5	1/6	.036	.166
Cond.	3	5	1/6	7	5	5	1	1/2	.167	.059
Finance	4	7	5	8	6	6	2	1	.333	.020

$\lambda_{\max} = 9.669$, C.R. $= 0.17$

Table 5. A family's housebuying pairwise comparison matrix for the criteria.

	Size	Trans.	Nbrhd.	Age	Yard	Modern	Cond.	Finance
Size	-	0.001721	0.007814	-0.00041	0.00054	0.000906	-0.08415	-0.03911
Trans.	-	-	-0.00331	0.001291	0.002485	0.003249	-0.06021	-0.01336
Nbrhd.	-	-	-	-0.00091	-0.00236	-5.7E-05	0.008376	-0.07561
Age	-	-	-	-	-0.01913	-0.03372	0.007638	**0.094293**
Yard	-	-	-	-	-	-0.01366	-0.01409	0.041309
Modern	-	-	-	-	-	-	-0.02599	0.029355
Cond.	-	-	-	-	-	-	-	0.006487
Finance	-	-	-	-	-	-	-	-

Table 6. Partial derivatives for the house example.

significance and identity of the original number as we first understood it to be. Thus in synthesizing near-consistent judgment values, changes that are too large can cause dramatic change in our understanding, and values that are too small cause no change in our understanding. We are left with only values of one order of magnitude smaller that we can deal with incrementally to change our understanding. It follows that our allowable consistency ratio should be not more than about .10. The requirement of 10% cannot be made smaller, such as 1% or 0.1%, without trivializing the impact of inconsistency. But inconsistency itself is important because without it new knowledge that changes preference cannot be admitted. Assuming that all knowledge should be consistent contradicts experience, which requires continued revision of understanding.

If the C.R. is larger than desired, we do three things: (1) find the most inconsistent judgment in the matrix; (2) determine the range of values to which that judgment can be changed corresponding to which the inconsistency would be improved; (3) ask the family to consider, if they can, changing their judgment to a plausible value in that range. If they are unwilling, we try with the second most inconsistent judgment, and so on. If no judgment is changed, the decision is postponed until better understanding of the criteria is obtained. In our house example the family initially made a judgment of 6 for the a_{37} entry in Table 1 and the consistency index of the set of judgments was C.I. $= (9.669 - 8)/7 = 0.238$. But C.R. $= .238/1.40 = 0.17$ is larger than the recommended value of 0.10. If we are going to ask the family to reconsider, and perhaps change, some of their pairwise comparisons, where should we start?

Three methods are plausible for this purpose. All require theoretical investigation of convergence and efficiency. The first uses an explicit formula for the partial derivatives of the Perron eigenvalue with respect to the matrix entries.

For a given positive reciprocal matrix $A = [a_{ij}]$ and a given pair of distinct indices $k > l$, define $A(t) = [a_{ij}(t)]$ by $a_{kl}(t) \equiv a_{kl} + t$, $a_{lk}(t) \equiv (a_{lk} + t)^{-1}$, and $a_{ij}(t) \equiv a_{ij}$ for all $i \neq k$, $j \neq l$, so $A(0) = A$. We use a linear function of t because multiplying by t when t is zero or close to zero can make the reciprocal very large, and thus we want t to be bounded away from zero. Also, we don't want t to be very large because the judgments would be too widespread, violating the requirement of homogeneity. Thus our assumption on the functional relationship is reasonable. Let $\lambda_{\max}(t)$ denote the Perron eigenvalue of $A(t)$ for all t in a neighborhood of $t = 0$ that is small enough to ensure that all entries of the reciprocal matrix $A(t)$ are positive there. Finally, let $v = [v_i]$ be the

1.00000	1.55965	3.26120	0.70829	1.07648	1.25947	0.32138	0.48143
0.64117	1.00000	1.16165	1.62191	1.72551	2.01882	0.61818	0.88194
0.30664	0.86084	1.00000	0.55848	0.49513	0.77239	**5.32156**	0.35430
1.41185	0.61656	1.79056	1.00000	0.59104	0.51863	1.36123	2.37899
0.92895	0.57954	2.01967	1.69193	1.00000	0.58499	1.07478	1.78893
0.79399	0.49534	1.29467	1.92815	1.70942	1.00000	0.91862	1.52901
3.11156	1.61765	2.25498	0.73463	0.93042	1.08858	1.00000	0.99868
2.07712	1.13386	2.82246	0.42035	0.55899	0.65402	1.00133	1.00000

Table 7. $\varepsilon_{ij} = a_{ij}w_j/w_i$.

	Size	Trans.	Nbrhd.	Age	Yard	Modern	Cond.	Finance	w
Size	1	1.7779	1.756208	0.774933	1.163989	1.418734	0.425449	0.494088	0.174
Trans.	0.562461	1	0.548777	1.556678	1.636746	1.994957	0.717895	0.794016	0.062
Nbrhd.	0.569408	1.822233	2	1.134652	0.994177	1.615679	0	0.675211	0.102
Age	1.290434	0.642394	0.881328	1	0.584131	0.533978	1.64704	2.23156	0.019
Yard	0.859115	0.610968	1.005857	1.711945	1	0.609428	1.315833	1.697915	0.034
Modern	0.704854	0.501264	0.618935	1.872735	1.640883	1	1.079564	1.39304	0.041
Cond.	2.35046	1.392962	0	0.60715	0.759975	0.9263	2	0.774223	0.223
Finance	2.02393	1.259421	1.481018	0.448117	0.588958	0.717855	1.291617	1	0.345

Table 8.

	Size	Trans	Nbrhd	Age	Yard	Modern	Cond	Finance	*w*	*v*
Size	1	5	3	7	6	6	1/3	1/4	.175	.042
Trans.	1/5	1	1/3	5	3	3	1/5	1/7	.062	.114
Nbrhd.	1/3	3	1	6	3	4	1/2	1/5	.103	.063
Age	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8	.019	.368
Yard	1/6	1/3	1/3	3	1	1/2	1/5	1/6	.034	.194
Modern	1/6	1/3	1/4	4	2	1	1/5	1/6	.041	.168
Cond.	3	5	2	7	5	5	1	1/2	.221	.030
Finance	4	7	5	8	6	6	2	1	.345	.021

$\lambda_{\max} = 8.811$, C.R. $= 0.083$

Table 9. Modified matrix in the a_{37} and a_{73} positions.

unique positive eigenvector of the positive matrix A^T that is normalized so that $v^T w = 1$. Then a classical perturbation formula [4, Theorem 6.3.12] tells us that

$$\left.\frac{d\lambda_{\max}(t)}{dt}\right|_{t=0} = \frac{v^T A'(0)w}{v^T w} = v^T A'(0)w$$
$$= \sum_{k \neq l} v_k w_l - \frac{1}{a_{kl}^2} v_l w_k.$$

We conclude that

$$\frac{\partial \lambda_{\max}}{\partial a_{ij}} = v_i w_j - a_{ji}^2 v_j w_i \quad \text{for all } i, j = 1, \ldots, n.$$

Because we are operating within the set of positive reciprocal matrices,

$$\frac{\partial \lambda_{\max}}{\partial a_{ji}} = -\frac{1}{a_{ij}^2} \frac{\partial \lambda_{\max}}{\partial a_{ij}} \quad \text{for all } i \text{ and } j.$$

Thus, to identify an entry of A whose adjustment within the class of reciprocal matrices would result in the largest rate of change in $\lambda_{\max}$, we should examine the $n(n-1)/2$ values $\{v_i w_j - a_{ji}^2 v_j w_i\}$, $i > j$, and select (any) one of largest absolute value. This is the method proposed for positive reciprocal matrices by Harker [8]. Ergu et al. [7] propose another method for dealing with consistency in the ANP. Here is how Harker's method is applied to our house example with the initial judgments in Table 1 replaced by $a_{37} = 6$, $a_{73} = 1/6$ to make it more inconsistent.

Table 6 gives the array of partial derivatives for the matrix of criteria in Table 1.

The (4, 8) entry in Table 5 (in bold print and underlined) is largest in absolute value. Thus, the family could be asked to reconsider their judgment of 1/8 for age vs. finance which indicates that finance is very strongly to extremely more important than age. One needs to know how much to change a judgment to improve consistency, and we show that next. One can then repeat this process with the goal of bringing the C.R. within the desired range. If the indicated judgments cannot be changed fully according to one's understanding, they can be changed partially. Failing the attainment of a consistency level with justifiable judgments, one needs to learn more before proceeding with the decision. Actually, the values used in the original example were $a_{37} = 1/2$, $a_{73} = 2$, derived in the simpler approach described next.

Two other methods, presented here in order of increasing observed efficiency in practice, are conceptually different. They are based on the fact that

$$n\lambda_{\max} - n = \sum_{\substack{i,j=1 \\ i\neq j}}^{n} (\varepsilon_{ij} + \varepsilon_{ij}^{-1}).$$

This suggests that we examine the judgment for which ε_{ij} is farthest from the number 1, that is, an entry a_{ij} for which $a_{ij}w_j/w_i$ is the largest, and see if this entry can reasonably be made smaller. We hope that such a change of a_{ij} also results in a new comparison matrix that has a smaller Perron eigenvalue. To demonstrate how improving judgments works, take the house example matrix in Table 1. To identify an entry ripe for consideration, construct the matrix $\varepsilon_{ij} = a_{ij}w_j/w_i$ (Table 7). The largest value in Table 7 is 5.32156, which focuses attention on $a_{37} = 6$.

How does one determine the most consistent entry for the $(3, 7)$ position? Harker [8] has shown that when we compute the new eigenvector w after changing the $(3, 7)$ entry, we want the new $(3, 7)$ entry to be w_3/w_7 and the new $(7, 3)$ entry to be w_7/w_3. On replacing a_{37} by w_3/w_7 and a_{73} by w_7/w_3 and multiplying by the vector w, one obtains the same product as one would by replacing a_{37} and a_{73} with zeros and the two corresponding diagonal entries with two (see Table 8).

We take the Perron vector of the latter matrix to be our w and use the now-known values of w_3/w_7 and w_7/w_3 to replace a_{37} and a_{73} in the original matrix. The family is now invited to change their judgment towards this new value of a_{37} as much as they can. Here the value was $a_{37} = 0.102/0.223 = 1/2.18$, approximated by $1/2$ from the AHP fundamental scale, and we hypothetically changed it to $1/2$ to illustrate the procedure (see Table 9). If the family does not wish to change the original value of a_{37}, one considers the second most inconsistent judgment and repeats the process.

One by one, each reciprocal pair a_{ij} and a_{ji} in the matrix is replaced by zero and the corresponding diagonal entries a_{ii} and a_{jj} are replaced by 2. The principal eigenvalue $\lambda_{\max}$ is then computed. The entry with the largest resulting $\lambda_{\max}$ is identified for change as described above. This method is in use in the ANP software program SuperDecisions [26]. The SuperDecisions software is used in teaching the subject. Here is the link to the webpage from which the SuperDecisions software can be downloaded, and it is free to educators and researchers: www. superdecisions. com. Incidentally, the name of the software is borrowed from its use of a matrix whose entries are matrices, the supermatrix, and is not an attempt to sound like something extraordinary.

Alternatives in a decision may be compared in pairs or if there are many, they can be rated one at a time by assigning them numbers from appropriate priority scales developed for each criterion such as (high, medium, low), (excellent, outstanding, very good, good, poor and very poor) that are then compared in pairs and their priorities derived, with each scale divided by the largest derived eigenvector component, making the largest value equal to one and the rest proportionately smaller. Comparisons yield a more accurate ranking of alternatives than rating them one at a time because rating involves memory of an ideal that is likely to vary among different people. If the alternatives vary widely, then the scales developed must reflect different orders of magnitude that are appropriately linked together [22].

The Normalized Priority Vector is Unique

To choose the best alternative in a decision, the priorities must be unique. There is more to the concept of priority. When $A = [w_i/w_j]$ is consistent, $A^k = n^{k-1}A$. This says how much a criterion represented by a row of A dominates other criteria through chains of k arcs, uniquely determined by the single arc chains represented by the rows of A itself. But this is not true when A is inconsistent.

Criterion i is said to dominate criterion j in one step if the sum of the entries in row i of A is greater than the sum of the entries in row j. It is convenient to use the vector $e = (1, \ldots, 1)^T$ to express this dominance: criterion i dominates criterion j in one step if $(Ae)_i > (Ae)_j$. A criterion can dominate another criterion in more than one step by dominating other criteria that in turn dominate the second criterion. Two-step dominance is identified by squaring the matrix and summing its rows, three-step dominance by cubing it, and so on. Thus, criterion i dominates criterion j in k steps if $(A^k e)_i > (A^k e)_j$. Criterion i is said simply to *dominate* criterion j if entry i of the vector

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} A^k e / e^T A^k e \tag{3}$$

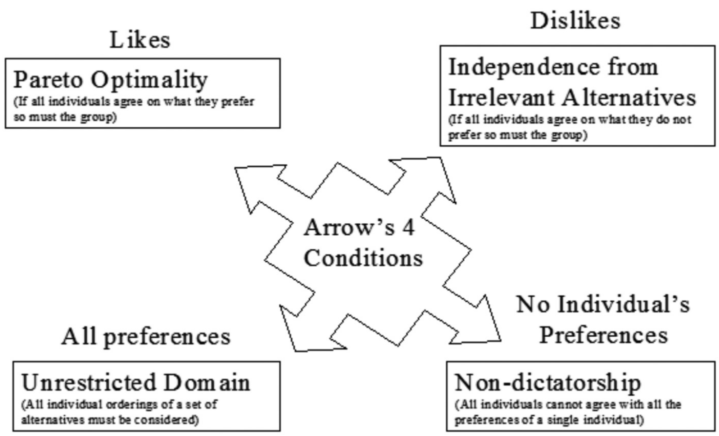


Figure 4. Arrow's four conditions.

is greater than its entry j. But this limit of averages can be evaluated: the Perron-Frobenius Theorem ensures that $A^k/\lambda_{\max}^k \to wv^T$ as $k \to \infty$, so

(4)
$$A^k e/(e^T A^k e) \approx \lambda_{\max}^k w(v^T e)/\lambda_{\max}^k (e^T w)(v^T e) = w \quad \text{as } k \to \infty.$$

Since (3) is a limit of averages of terms of a sequence that converges to the Perron vector w of A, according to Cesaro summability [17], (3) is actually equal to w.

More simply, a priority vector w can be used to weight the columns of its matrix and sum the elements in each row to obtain a new priority vector. Such ambiguity is eliminated if we require that a priority vector satisfy the condition $Aw = cw$, $c > 0$. The constant c is needed because w is derived from absolute scale entries (invariant under the identity transformation) and thus in normalized form also belongs to an absolute scale. It is easy to prove, using the biorthogonality of left and right eigenvectors [7] that c and w must be, respectively, the Perron value and vector of A. With the concept of dominance, we have proved that the Perron vector is necessary for deriving priorities.

Should one be concerned about the often inexact form of the judgments? Wilkinson [27] has studied the stability of an eigenvector of a square matrix with real coefficients. Perturbing the matrix by adding to it the perturbation matrix ΔA yields the following perturbation Δw_1 in the principal eigenvector w_1. The expression below involves all the eigenvalues λ_i of A and all of both its left (v_j) and right (w_j) eigenvectors:

$$\Delta w_1 = \sum_{j=2}^{n} (v_j^T \Delta A w_1/(\lambda_1 - \lambda_j) v_j^T w_j) w_j.$$

The left and right eigenvectors v and w are in normalized form. T indicates transposition.

The eigenvector w_1 is stable when the following hold:

(1) The perturbation ΔA is small as observing the consistency index would ensure.

(2) λ_j is well separated from λ_1; when A is consistent, $\lambda_1 = n$, $\lambda_j = 0$.

(3) The product of left and right eigenvectors is not too large, which is the case for a consistent (and near-consistent) matrix if the elements are homogenous (compared here on the relative dominance scale of the absolute values 1–9) with respect to the criterion of comparison.

(4) The number of their entries is small (hence perhaps why inconsistency becomes problematic as to which element causes it the most for $n > 7$, [13]).

The conclusion is that *n must be small*, and one must compare *homogeneous* elements, which is in harmony with the axioms of the AHP [18].

Synthesis of Individual Judgments into a Representative Group Judgment

Kenneth Arrow's Impossibility Theorem, for which he received the Nobel Prize in 1972, stated that it was not possible to find a representative group judgment from the judgments of individuals using ordinal preferences. However, if one allows cardinal preferences and uses the geometric mean to combine individual judgments as we do in the AHP, it is possible. In 1983 we proved, in a paper coauthored with Janos Aczel, that the unique way to combine reciprocal individual judgments into a corresponding reciprocal group judgment is by using their geometric mean [1].

Arrow proved in his impossibility theorem, using ordinal preferences (either A is preferred to B or it is not) that there does not exist a social welfare function that satisfies all four conditions listed in Figure 4, at once. We showed in April 2011 [24] in a journal of which Arrow is an editor that ***with cardinal intensities of preference and the geometric mean to combine the individual judgments into a representative group judgment, a social welfare function exists that satisfies these four conditions.*** Thus we have a possibility theorem.

Validation and Diverse Uses

How do we test for the validity of the process? One of the things we can do is to get judgments from many people, even those who may not be experts in decision making but who are experts in what they do. Should the answer always match the data available? What if the data themselves are incorrect? What if we don't know enough to create a very complete structure for a decision? These questions have been examined in the literature, but at best one needs to apply the process in a number of decisions to develop confidence in its reliability. We have provided three examples to illustrate its accuracy when used with systematic knowledge and understanding, of both the decisions to which it is applied and of its limitations that revolve around the adequacy of the structure used to represent the decision, and the experience necessary to develop sound and accurate judgments used in making comparisons.

Relative Consumption of Drinks

Table 10 shows how an audience of about thirty people, using consensus to arrive at each judgment, provided judgments to estimate the *dominance* of the consumption of drinks in the U.S. (which drink is consumed more in the U.S. and how much more than another drink?). The derived vector of relative consumption and the actual vector, obtained by normalizing the consumption given in official statistical data sources, are at the bottom of the table.

Optics Example

Four identical chairs were placed on a line from a light source at the distances of nine, fifteen, twenty-one, and twenty-eight yards. The purpose was to see if one could stand by the light and look at the chair and compare their relative brightness in pairs, fill in the judgment matrix, and obtain a relationship between the chairs and their distance from the light source. This experiment was repeated twice with different judges whose judgment matrices are shown in Table 11.

The judges of the first matrix were the author's young children, ages five and seven at that time, who gave their judgments qualitatively. The judge of the second matrix was the author's wife, also a mathematician and not present during the children's judgment process. In Table 12 we give the principal eigenvectors, eigenvalues, consistency indices, and consistency ratios of the two matrices.

First and second trial eigenvectors of Table 12 have been compared with the last column of Table 13 calculated from the inverse square law of optics. How close are the eigenvectors to the actual result for physics? We use a compatibility index that we developed for that purpose. We take the Hadamard product of a matrix of ratios of the entries of one vector with the transpose of a second matrix of the other vector. If the two vectors are identical, each entry of the Hadamard product would be equal to one and the sum of all resulting entries would be equal to n^2. Otherwise, one divides the resulting sum by n^2 and ensures that the ratio is about 1.01. It is interesting and important to observe in this example that the numerical judgments have captured a natural law. It would seem that they might do the same in other areas of perception or thought, like the one on estimating chess championship outcomes that we show in the next example, and, more generally, in continuous versions of these ideas.

Which Drink is Consumed More in the U.S.?
An Example of Estimation Using Judgments

Drink Consumption in the U.S.	Coffee	Wine	Tea	Beer	Sodas	Milk	Water
Coffee	1	9	5	2	1	1	1/2
Wine	1/9	1	1/3	1/9	1/9	1/9	1/9
Tea	1/5	2	1	1/3	1/4	1/3	1/9
Beer	1/2	9	3	1	1/2	1	1/3
Sodas	1	9	4	2	1	2	1/2
Milk	1	9	3	1	1/2	1	1/3
Water	2	9	9	3	2	3	1

The derived scale based on the judgments in the matrix is:

Coffee	Wine	Tea	Beer	Sodas	Milk	Water
.177	.019	.042	.116	.190	.129	.327

with a consistency ratio of .022.

The actual consumption (from statistical sources) is:

.180	.010	.040	.120	.180	.140	.330

Table 10. Relative consumption of drinks.

Relative visual brightness (1st Trial)

	C_1	C_2	C_3	C_4
C_1	1	5	6	7
C_2	1/5	1	4	6
C_3	1/6	1/4	1	4
C_4	1/7	1/6	1/4	1

Relative visual brightness (2nd Trial)

	C_1	C_2	C_3	C_4
C_1	1	4	6	7
C_2	1/4	1	3	4
C_3	1/6	1/3	1	2
C_4	1/7	1/4	1/2	1

Table 11. Pairwise comparisons of the four chairs.

Relative brightness eigenvector (1st Trial)	Relative brightness eigenvector (2nd Trial)
0.61	0.62
0.24	0.22
0.10	0.10
0.05	0.06
$\lambda_{max} = 4.39$, C.I. = 0.13, C.R.= 0.14	$\lambda_{max} = 4.10$, C.I. = 0.03, C.R.= 0.03

Table 12. Principal eigenvectors and corresponding measures.

Note the sensitivity of the results as the closest chair is moved even closer to the light source, for then it absorbs most of the value of the relative index and a small error in its distance from the source yields great error in the values. What is noteworthy from this sensory experiment is the observation or hypothesis that the observed intensity of illumination varies (approximately) inversely with the square of the distance. The more carefully designed the experiment, the better the results obtained from the visual observations.

Distance	Normalized distance	Square of normalized distance	Reciprocal of previous column	Normalized reciprocal	Rounding off
9	0.123	0.015 129	66.098	0.607 9	0.61
15	0.205	0.042 025	23.79	0.218 8	0.22
21	0.288	0.082 944	12.05	0.110 8	0.11
28	0.384	0.147 456	6.78	0.062 3	0.06

Table 13. Inverse square law of optics.

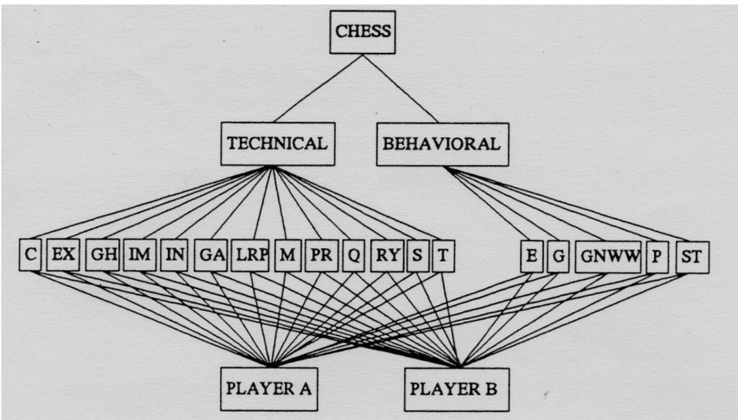


Figure 5. Criteria, factors, and players in chess competition.

World Chess Championship Outcome Validation: The Karpov-Korchnoi Match

The following criteria (Table 14) and hierarchy (Figure 5) were used to predict the outcome of world chess championship matches using judgments of ten grandmasters in the former Soviet Union and the United States who responded to questionnaires mailed to them. The predicted outcomes that included the number of games played, drawn, and won by each player either was exactly as they turned out later or adequately close to predict the winner. The outcome of this exercise was officially notarized before the match took place. The notarized statement was mailed to the editor of the *Journal of Behavioral Sciences* along with the paper later published in May 1980. The prediction was that Karpov would win by six to five games over Korchnoi, which he did.

The AHP, the name of the decision process described above, has been used in various settings to make decisions.

This approach to prioritization provides the opportunity to help focus attention on the important issues in the world and allocate resources to them accordingly.

- Since its early development, the AHP has been used to predict correctly, a few months before the elections, the next candidate to be elected for president. The factors involved varied from election to election depending on the domestic and international circumstances prevailing at the time.
- In 1986 the Institute of Strategic Studies in Pretoria, a government-backed organization, used the AHP to analyze the conflict in South Africa and recommended actions ranging from the release of Nelson Mandela to the removal of apartheid and the granting of full citizenship and equal rights to the black majority. All of these recommended actions were quickly implemented by the white government.
- A company used it in 1987 to choose the best type of platform to build to drill for oil in the North Atlantic. A platform costs around three billion dollars to build, but the demolition cost was an even more significant factor in the decision.
- Xerox Corporation has used the AHP to allocate close to a billion dollars to its research projects.
- IBM used the process in 1991 in designing its successful mid-range AS 400 computer. IBM won the prestigious Malcolm Baldrige award for Excellence for that effort. The book about the AS 400 project has a chapter devoted to how AHP was used in benchmarking.
- The AHP has been used since 1992 in student admissions and prior to that in military personnel promotions and in hiring decisions.
- The process was applied to the U. S. versus China conflict in the intellectual property rights battle of 1995 over Chinese individuals copying music, video, and software tapes and CD's. An AHP analysis involving three hierarchies for benefits, costs, and risks showed that it was much better for the U. S. not to sanction China. Shortly after the study was completed, the U. S. awarded China most-favored nation trading status and did not sanction it.
- In sports it was used in 1995 to predict which football team would go to the Super Bowl and win (correct outcome, Dallas won over my hometown, Pittsburgh). The AHP was applied in baseball to analyze which Padres players should be retained.
- British Airways used it in 1998 to choose the entertainment system vendor for its entire fleet of airplanes
- The Ford Motor Company used the AHP in 1999 to establish priorities for criteria that improve customer satisfaction.

Table 14. Definitions of chess factors.

T (1)	Calculation (Q): The ability of a player to evaluate different alternatives or strategies in light of prevailing situations.
B (2)	Ego (E): The image a player has of himself as to his general abilities and qualification and his desire to win.
T (3)	Experience (EX): A composite of the versatility of opponents faced before, the strength of the tournaments participated in, and the time of exposure to a rich variety of chess players.
B (4)	Gamesmanship (G): The capability of a player to influence his opponent's game by destroying his concentration and self-confidence.
T (5)	Good Health (GH): Physical and mental strength to withstand pressure and provide endurance.
B (6)	Good Nerves and Will to Win (GN): The attitude of steadfastness that ensures a player's health perspective while the going gets tough. He keeps in mind that the situation involves two people and that if he holds out the tide may go in his favor.
T (7)	Imagination (IW: Ability to perceive and improvise good tactics and strategies.
T (8)	Intuition (IN): Ability to guess the opponent's intentions.
T (9)	Game Aggressiveness (GA): The ability to exploit the opponent's weaknesses and mistakes to one's advantage; occasionally referred to as "killer instinct."
T (10)	Long Range Planning (LRP): The ability of a player to foresee the outcome of a certain move, set up desired situations that are more favorable, and work to alter the outcome.
T (11)	Memory M: Ability to remember previous games.
B (12)	Personality (P): Manners and emotional strength, and their effects on the opponent in playing the game and on the player in keeping his wits.
T (13)	Preparation (PR): Study and review of previous games and ideas.
T (14)	Quickness (Q): The ability of a player to see clearly the heart of a complex problem.
T (15)	Relative Youth (RY): The vigor, aggressiveness, and daring to try new ideas and situations, a quality usually attributed to young age.
T (16)	Seconds (S): The ability of other experts to help one to analyze strategies between games.
B (17)	Stamina (ST): Physical and psychological ability of a player to endure fatigue and pressure.
T (18)	Technique M: Ability to use and respond to different openings, improvise middle game tactics, and steer the game to a familiar ground to one's advantage.

- In 2001 it was used to determine the best site to relocate the earthquake-devastated Turkish city of Adapazari.
- A comprehensive analysis as to whether the Unites States should develop an anti-nuclear missile (estimated in the 1990s to cost sixty billion dollars and strongly opposed by scientists as technically infeasible) was presented to the National Defense University (NDU) in February 2002. In December of that year President Bush decided to go for it. The U.S. actually developed prototypes and tested them in stages successfully.
- An application by Professor Wiktor Adamus of Krakow University convinced the prime minister of Poland in 2007 not to adopt the Euro for currency until many years later.
- An AHP application, known to the military at the Pentagon, showed that occupying or bombing Iran in terms of benefits, opportunities, costs, and risks is not the best option for security in the Middle East.
- The AHP was used to assist the Green Bay Packers to hire the best players, perhaps partly the reason why they won the Super Bowl football championship in 2011 by beating the Pittsburgh Steelers. Other teams, including hockey and baseball, are also using it.
- In 1991, 2001, and 2009, AHP was used in three studies by economists to determine the turn-around dates of the U.S. economy and the strength of recovery. These studies were uncannily accurate.
- The latest application made in August 2011 was to the Israeli-Palestinian conflict

when five top people from each side used the AHP to reach an agreement called the Pittsburgh Principles. One of them wrote: "I had been in hundreds of meetings between Israelis and Palestinians where we tried to reach a joint statement but failed because in most of the cases each side was trying to score points and court his own public opinion rather than being objective and trying to be real and responsible."

- The AHP is used by many organizations, including the military, to prioritize their projects and allocate their resources optimally according to these priorities.

Since the AHP helps one organize one's thinking, it can be used to deal with many decisions that are often made intuitively. At a minimum, the process allows one to experiment with different criteria, structures, and judgments and also to test the sensitivity of the outcome to changes in both the structure [19] and the judgments. It appears that if we know how to measure things in relative terms according to the criteria that they share, we can measure anything that way, and that kind of measurement includes, as a special case, the normalized measurement that we make in a scientific field, in which we always have to interpret the significance of the measurements obtained by using expert knowledge and judgment in that field.

This work on the AHP was developed independently of the Theory of Perron although I refer to him abundantly. Consistent matrices automatically satisfy Perron's conditions, lead to his results, and generalize to acceptably inconsistent matrices through perturbation arguments some of which were developed by J. H. Wilkinson [27]. We hope that we can have another opportunity to show the reader how the ANP works and how the discrete mathematics of comparisons has been generalized to the continuous case involving Fredholm's equation whose solution produces results associated with neural firing and synthesis.

References

[1] J. Aczel and T. Saaty, Procedures for synthesizing ratio judgments, *Journal of Mathematical Psychology* **27** (1983), 93–102.

[2] E. Batschelet, *Introduction to Mathematics for Life Scientists*, Springer-Verlag, New York, 1971.

[3] R. A. Bauer, E. Collar, and V. Tang, *The Silverlake Project*, Oxford University Press, New York, 1992.

[4] H. Bergson, The Intensity of Psychic States, Chapter 1 in *Time and Free Will: An Essay on the Immediate Data of Consciousness*, translated by F. L. Pogson, M.A., George Allen and Unwin, London, 1910, pp. 1–74.

[5] P. Davis and P. T. Hersh, *Descartes' Dream*, Harcourt Brace Jovanovich, New York, 1982.

[6] S. Dehaene, *The Number Sense*: *How the Mind Creates Mathematics*, Oxford University Press, USA, 1997.

[7] Daji Ergu, G. Kou, Yi Peng, and Yong Shi, A simple method to improve the consistency ratio of the pairwise comparison matrix in ANP, *European Journal of Operational Research* **213** (2011), no. 1, 246–259.

[8] P. Harker, Derivatives of the Perron root of a positive reciprocal matrix: With applications to the analytic hierarchy process, *Applied Mathematics and Computation* **22** (1987), 217–232.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[10] H. Lebesgue, *Leçons Sur L'integration*, second ed., Gauthier-Villars, Paris, 1928.

[11] L. LeShan and H. Margenau, *Einstein's Space and Van Gogh's Sky*, Macmillan, 1982.

[12] A. F. MacKay, *Arrow's Theorem: The Paradox of Social Choice*, Yale University Press, 1980.

[13] M. Ozdemir and T. L. Saaty, The unknown in decision making: What to do about it, *European Journal of Operational Research* **174** (2006), 349–359.

[14] Y. Peng, G. Kou, G. Wang, W. Wu, and Y. Shi, Ensemble of software defect predictor: An AHP-based evaluation method, *International Journal of Information Technology & Decision Making* **10** (2011), no. 1, 187–206.

[15] T. L. Saaty, *The Analytic Hierarchy Process*, McGraw Hill, New York, 1980. Reprinted by RWS Publications, available electronically free, 2000.

[16] ———, *Fundamentals of Decision Making*, RWS Publications, 2006, 478 pp.

[17] ———, *Theory and Applications of the Analytic Network Process*, RWS Publications, 2006, 352 pp.

[18] ———, *Principia Mathematica Decernendi*, subtitled *Mathematical Principles of Decision Making*, RWS Publications, 2010, 531 pp. http://Rozann@creativedecisions.net.

[19] T. L. Saaty and E. Forman, *The Hierarchon*, a collection of nearly 800 hierarchies in all kinds of life, many actual, applications made by people, and *The Encyclicon*, three volumes of nearly 900 pages of several hundred network decision applications, all published by RWS Publications.

[20] T. L. Saaty and M. Ozdemir, Why the magic number seven plus or minus two, *Mathematical and Computer Modelling* **38** (2003), 233–244.

[21] T. L. Saaty and K. Peniwati, *Group Decision Making: Drawing Out and Reconciling Differences*, RWS Publications, 2008, 385 pp.

[22] T. L. Saaty and J Shang, An innovative orders-of-magnitude approach to AHP-based mutli-criteria decision making: Prioritizing divergent intangible humane acts, *European Journal of Operational Research* **214** (2011), no. 3, 703–715.

[23] T. L. Saaty and L. G. Vargas, Hierarchical analysis of behavior in competition: Prediction in chess, *Behavioral Sciences* **25** (1980), 180–191.

[24] ———, The possibility of group choice: Pairwise comparisons and merging functions, *Social Choice and Welfare*, April 2011.

[25] A. Schopenhauer and K. Hillebrand, *On the Fourfold Root of the Principle of Sufficient Reason*, an essay, Kindle eBook, 2011.

[26] Superdecisions. http://www.superdecisions.com, until 2013.

[27] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.