Rank from Comparisons and from Ratings in the Analytic Hierarchy/Network Processes

Thomas L. Saaty
University of Pittsburgh

Theme (*Weak version*): The rank of a given set of independent alternatives with respect to several criteria must stay the same if new alternatives are added or old ones deleted unless adding or deleting alternatives introduces or deletes criteria and changes judgments.

Theme (*Strong version*): Replace the word *rank* by *priority* with the added condition that the ratios of the priorities of the original alternatives must be the same before and after deleting old alternatives or adding new ones.

Abstract

Rank preservation and reversal are important subjects in multicriteria decision-making particularly if a theory uses only one of two ways of creating priorities: rating alternatives one at a time with respect to an ideal or standard, or comparing them in pairs. It is known that our minds can do both naturally and without being tutored. When rating alternatives, they must be assumed to be independent and rank should be preserved. When comparing alternatives, they must be assumed to be dependent and rank may not always be preserved. However, even in making comparisons rank can be preserved if one uses idealization instead of normalization with the original set of alternatives and preserves that ideal from then on unless that ideal itself is deleted for some reason. So often it is a matter of judgment as to whether it is desirable to force rank preservation or allow rank to adjust as necessary. Examples are given to illustrate the foregoing ideas.

1. Introduction

The Harvard psychologist Arthur Blumenthal tells us in his book *The Process of Cognition*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1977, that there are two types of judgment: "*Comparative judgment* which is the identification of some relation between two stimuli both present to the observer, and *absolute judgment* which involves the relation between a single stimulus and some information held in short term memory about some former comparison stimuli or about some previously experienced measurement scale using which the observer rates the single stimulus." In the Analytic Hierarchy Process (AHP) we call the first *relative measurement* and the second *absolute measurement*. In relative measurement we compare each alternative with many other alternatives and in absolute measurement we compare each alternative with one ideal alternative we know of or can imagine, a process we call *rating alternatives*. The first is descriptive and is conditioned by our observational ability and experience and the second is normative, conditioned by what we know is best, which of course is relative. Comparisons must precede ratings because ideals can only be created through experience using comparisons to arrive at what seems best. It is interesting that in order to rate alternatives with respect to an ideal as if they are independent can only be done after having made comparisons that involve dependence to create the ideal or standard in the first place. Making comparisons is fundamental and intrinsic in us. They are not an intellectual invention nor are they something that can be ignored.

The tradition of measurement makes one think that there is only one way to measure things and that is on a physical scale or assign values to them one at a time. We have been in the habit of creating scales and waiting for things to arrive to be measured on these scales as needed. A unique value is assigned to each thing or element from a scale. The value assigned to an element is unconditional, as it does not depend on the measurements of other elements. That is not the case with scales derived from paired comparisons. Unlike measurement on traditional scales, these scale values exist only after one has the objects or criteria to compare. In addition the values derived for each element are relative to what other values it is compared with, and thus each time an element is compared with other elements it has a different value. The values derived are conditional. Derived relative scales need not have a unit, but they can after they are derived by dividing by the value of one of them if desired. In addition it is possible to create a unity or ideal after a first set is chosen and compare every element that is added thereafter with respect to the standard (unit) and allow it to become larger than that unit or smaller as needed. Relative scales of measurement derived from a fundamental scale of paired comparisons with values that belong to an absolute scale (invariant under the identity transformation) themselves belong to an absolute scale. One can see from the literature of scales that scales derived as in the AHP are a new paradigm in measurement that many people do not understand well even after a degree of exposure. But numerous examples show that it has useful characteristics not available in existing measurement particularly with regard to the measurement of intangibles, and with using judgment and understanding within a sizeable structure to examine possible future happenings. Usually we have been led to believe that intuition is unreliable because single hunches are usually inaccurate. We have found that intuition is very reliable when a knowledgeable person provides judgments that are many and well integrated within an organized structure. Because comparisons are our biological inheritance, and also because experience and judgment are what distinguishes the expert from the non-expert, it appears that we need but to formalize our understanding within a transparent and justifiable scientific framework like the AHP to make it more reliable and usable.

There is no objectivity apart from human values that leads to the ranking of alternatives. Nature has no predetermined rank of alternatives on specially chosen criteria of its own. It is people who establish the criteria and their perceived rankings on these criteria.

When alternatives are thought to be independent of one another they are rated one at a time. IN that case one must be able to say how high or low an alternative rate on that criterion. To do that one msust have sometjing in mind called an ideal so that one gets the feeling about how close or far that alternative is from the ideal and allocate s it to one of various intensity slots of ranking such as very high, high, medium, low, poor and so on. These intensity slots must themselves be compared pairwise to obtain numerical priorities for them that would be associated with the alternatives. One more thought here is that these intensities may fall in different order of magnitude groupings linked by a pivot so that an alternative can be honestly plasec in one or the other of these groupings because not all alternatives are of the same order of importance as with the cherry tomato and the watermelon example shown earlier in this book.

When the criterion is intangible the alternatives of necessity are compared on a scale with respect to an ideal (the best conceivable alternative) one has in mind for that criterion. An ideal is understood or imagined by examining many diverse alternatives. Thus rating with respect to an ideal involves indirect dependence among the alternatives whether present or absent. The excellence of the ideal changes more and more when more is learned about a new and superior alternative; no matter how much imagination is applied, new alternatives that could not be imagined previously will come along and change the excellence of the old ideal to a new ideal. There is never a sure ideal that is permanent. Strictly speaking rating alternatives always compares them indirectly with other alternatives from which an ideal inherits its superior status and hence there is indirect dependence among all the alternatives all the time which implies the possibility of change in rank when alternatives are added or deleted. To preserve rank is to find a way to enforce its staying the same when new alternatives are added or old ones deleted. When alternatives are rated one at a time their scores are created independently and their ranks are always preserved. When they are compared, their ranks depend on each other and these ranks are no longer independent of one another. In that case to preserve rank, one does takes the given set of alternatives and divides their priorities under each criterion by the largest priority among them so that the alternative with the largest value becomes the ideal among them. Any new alternative is only compared pairwise in a 2 by 2 matrix with the ideal alternative for that criterion and if better than that ideal, its priority becomes more than one and stays that way. In this case there can be no rank reversal.

There are situations in life when alternatives are mutually exclusive and exhaustive and no new alternatives can be added or existing ones deleted. The ranking of such alternatives differs from rankings that involve more and more alternatives, decisions in which the number of alternatives is open.

In paired comparisons, the ideal is used for each criterion when the alternatives are mutually exclusive and exhaustive. Otherwise the distributive mode is used and the ideal is formed at the end for further use as for example in BOCR. Thus BOCR answers can be of two kinds: One is for mutually exclusive and exhaustive alternatives with the ideal done for each criterion and one for dependent alternatives that can be increased by adding others or decreased by taking out some. In this case the ideal is formed from the distributive mode at the end. BOCR generally assumes the independence of the four merits. Otherwise they would be included in a single network to indicate their interdependence.

That is not the case when the alternatives are compared with each other: comparisons are necessary when criteria are involved even when these criteria are independent of one another. With comparisons the scores of the alternatives become directly dependent on one another. Is there a way to treat these scores with respect to different criteria to preserve their rank? It is clear that if alternatives are dependent on one another functionally (as distinct from dependence in judgment) their ranks can change. Is there a way to rank alternatives when they are functionally independent of one another so that knowledge of one alternative does not involve knowledge of another? It is clear that dependence in ranking or in priority is an inescapable fact because our idea of what a good high-scoring alternative is depends on what other alternatives we know. Plus the fact that human goals and values are involved in ranking makes the alternatives dependent on one another even when in reality they exist in different places and at different times.

With dependence rank can change. Rank reversal can happen when a limited resource is allocated among a set of independent alternatives and a new alternative is added or when the alternatives bring different resources under different criteria so that the amount of the resource possessed by a criterion changes and this causes the relative weights of the criteria themselves to change. Thus as we will also see below *preserving rank is a tenuous concern even when the alternatives are thought to be functionally independent*. In practice we want to preserve rank to maintain order and reduce the complexity of conflicts. Thus the situation in which rank needs to be preserved must be specified. The best that one can say about them is that they invite attention when there are many alternatives over whose presence we have no control such as admitting students to a university, patients to a hospital, promoting military personnel, and so on, and the idea of fairness (first-come first-served) in ranking them seems to be the motivation to preserve their ranks. In all situations where we have control over what alternatives to consider and what not to consider it seems reasonable to allow rank to change because their quality in our mind depends on what other alternatives we consider.

2. Fundamentals of the AHP using Comparisons

In the AHP paired comparison judgments from a fundamental scale of absolute numbers are entered in a reciprocal matrix. Their numerical values and corresponding intensities are: 1 = equal, 3 = moderately dominant, 5 = strongly dominant, 7 = very strongly dominant and 9 = extremely dominant, along with intermediate values for compromise and reciprocals for inverse judgments and even using decimals to compare homogeneous elements whose comparison falls within one unit. From the matrix an absolute scale of relative values is obtained on normalization (by dividing each value by the sum of all the values) that is used when conditional dependence on the quality and number of other elements in the comparisons is needed, or on idealization (dividing each value by the largest value of any alternative) used when conditional dependence is not needed. Both these modes, the first called *distributive* and the second *ideal*, are required for use in *relative measurement* as derived from paired comparisons. In the consistent case adding the entries in any column and then dividing each entry in that column by its sum, or dividing the sum of each column by their total sum, gives the priorities as the principal right eigenvector of the matrix. Since each judgment is expressed as an absolute number from the fundamental scale, so also is their sum, and their ratio. Every column would give the same vector of priorities because of consistency. In the inconsistent case one solves a system of linear homogeneous equations that have coefficients that are from an absolute scale to obtain the principal right eigenvector for the priorities, and hence the solution also belongs to an absolute scale that becomes relative on normalization or idealization. Thus the AHP uses only absolute scale numbers for judgments and for their resulting priorities. Because it compares alternatives with each other, relative measurement is descriptive not normative.

Because the AHP is a multicriteria process we need to combine the priorities of the alternatives derived under the different criteria. The only possible meaningful way to do this that preserves the influence of the proportionality of priority of the criteria on each corresponding vector of alternative priorities is to multiply and add. This is also validated by the more general Analytic Network Process (ANP) interdependence feedback approach that involves the concept of dominance and raising the supermatrix to powers thus again using products and sums.

3. Fundamentals of the AHP using Absolute Judgment: Rating

When one rates alternatives, they must be independent of one another. The presence or absence of an alternative must have no effect on how one rates any of the others. We call this kind of ranking of alternatives with respect to an ideal (which is an arbitrarily chosen fixed reference point) *absolute measurement* or *rating*. Absolute measurement is analogous to measuring something with a physical device; for example, measuring length with a yardstick.

In order to rate alternatives with respect to an ideal, we need to create intensity levels or degrees of variation of quality on a criterion; for example, excellent, above average, average, below average and poor. We then pairwise compare them to establish priorities and normalize those priorities by dividing by the largest value among them, so that *excellent* would have a value of 1.000 and the others would be proportionately less. Idealizing the priorities by dividing by the largest assures that intensities belonging to large families do not receive small priorities simply because there are many of them. We then rate an alternative by selecting the appropriate intensity level for it on each criterion. Even when we use a numerical scale, say 1 to 100, to rate each alternative we must have an intuitive idea of how high or how low an alternative falls and in the process we subconsciously make comparisons among different levels on the scale. It is not the exact number chosen, but the level of intensity of feeling behind where it should fall, up or down, on the scale that matters. Because it compares the alternatives with respect to a standardized ideal, absolute measurement is normative not descriptive.

The ratings approach is illustrated in the following example of choosing the best city to live in. Figure 1 shows the goal, criteria and their priorities obtained from paired comparisons, and the intensities for each criterion with their idealized values obtained by dividing by the largest value in the vector of priorities derived from their paired comparisons matrix.

The pairwise comparisons for the C*ultural* criterion intensities and the resulting priorities are illustrated in Table 1 below. The values in the *Idealized* column are obtained by dividing each priority in the *Derived* column by the largest, .569. The prioritized intensities become the standards from which one selects the appropriate one to describe a particular city's performance with respect to *Cultural* (interpret this as cultural opportunities). The prioritized intensities in essence become a standardized performance scale, something like a yardstick that can be used to rate a city on culture. Note that for this criterion of culture, judgment in still involved in deciding which intensity to pick. Actual data can also be used in establishing the priorities, usually involving some form of idealization where data is converted to priorities directly.

A score is computed for a city by multiplying the priority of the selected intensity times the priority of the criterion and summing for all the criteria, shown in the Total Score column in Table 2. The Priorities column is obtained by normalizing the Total Score column by dividing by the sum of the values in it. The selected intensities for each alternative, the ratings, are shown in Table 2 below. The priorities corresponding to the ratings are shown in Table 3.

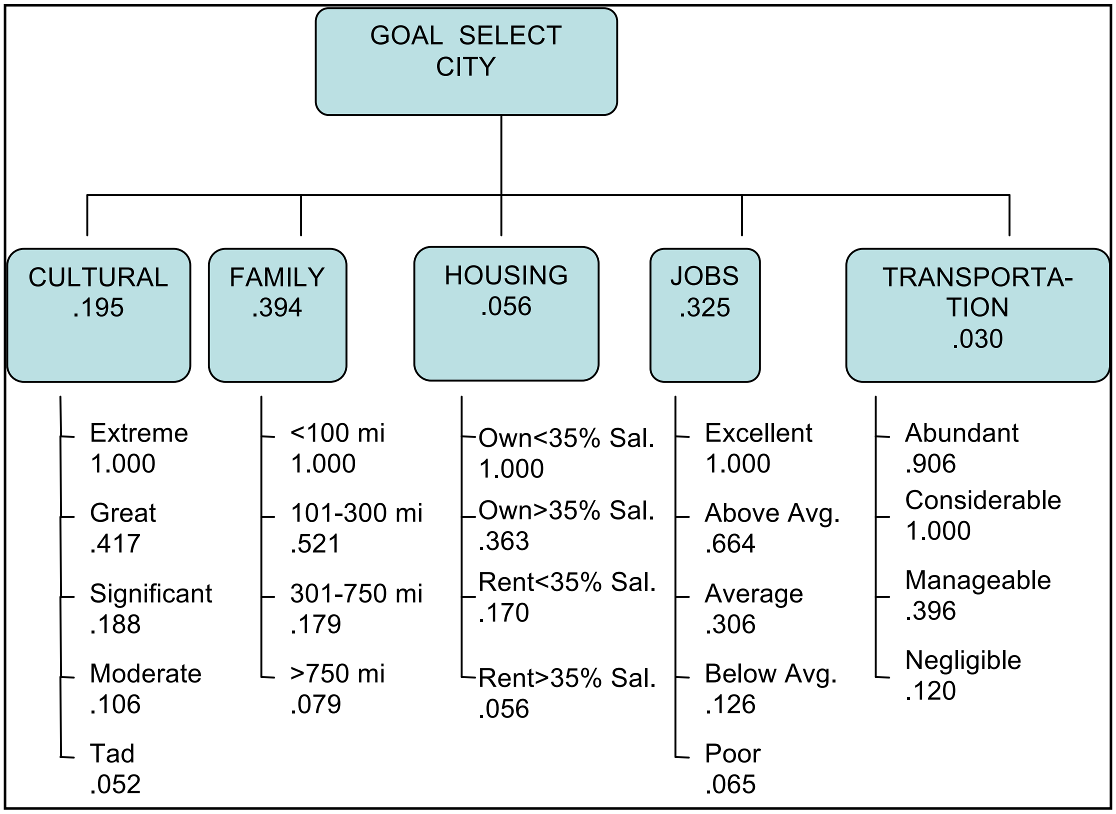

Figure 1 Choosing a City to Live in Using Ratings.

Table 1 Deriving Priorities for the Cultural Criterion Categories

	Extreme	**Great**	**Significant**	**Moderate**	**Tad**	*Derived Priorities*	*Idealized Priorities*
Extreme	1	5	6	8	9	*.569*	*1.000*
Great	1/5	1	4	5	7	*.234*	*.411*
Significant	1/6	1/4	1	3	5	*.107*	*.188*
Moderate	1/8	1/5	1/3	1	4	*.060*	*.106*
Tad	1/9	1/7	1/5	1/4	1	*.030*	*.052*

Inconsistency = .112

Table 2 Ratings of Alternative Cities for each Criterion

Alternatives	**Cultural** *.195*	**Family** *.394*	**Housing** *.056*	**Jobs** *.325*	**Transport** *.030*	**Total Score**	**Priorities (Normal.)**
Pittsburgh	Signific.	<100 mi	Own>35%	Aver-age	Manage-able	.562	.294
Boston	Extreme	301-750 mi	Rent>35%	Above Avg	Abundant	.512	.267
Bethesda	Great	101-300 mi	Rent<35%	Excel-lent	Consider-able	.650	.339
Santa Fe	Signific.	>750 mi	Own>35%	Aver-age	Negligible	.191	.100

Table 3 Priorities of the Ratings in Table 2

Alternatives	**Cultural** *.195*	**Family** *.394*	**Housing** *.056*	**Jobs** *.325*	**Trans-port** *.030*	**Total Score**	**Priorities (Normalized)**
Pittsburgh	0.188	1.000	0.363	0.306	0.396	.562	.294
Boston	1.000	0.179	0.056	0.664	0.906	.512	.267
Bethesda	0.411	0.521	0.170	1.000	1.000	.650	.339
Santa Fe	0.188	0.079	0.363	0.306	0.120	.191	.100

4. Structural and Functional Dependence

Unlike rating alternatives where we compare them to the best possible standard or ideal alternative, in the comparative judgment process we compare each alternative with some or all of the other alternatives. In that case an alternative that is ideally poor on an attribute could have a relatively high priority when compared with still poorer alternatives on that attribute but have low priority on another attribute where it is almost ideally good but is compared with better-valued alternatives. Thus the final rank of any alternative depends on the quality of the alternatives with which it is compared. Hence in making comparisons among alternatives, the priority of any alternative is influenced not only by how many alternatives it is compared with but by their quality.

In general increasing copies of an alternative indefinitely so that the universe is full of them tends to depreciate their value unless there is synergy among them so that the whole is more than the sum of its parts. Synergy happens when the copies support each other's functions, so that they tend to increase each other's value. The first, depreciation of value, is due to structural dependence and the second, appreciating value, is due to functional dependence in which the alternatives directly depend on each other like some industries do.

How do we determine whether alternatives are dependent? We compare them in pairs to see how strongly more a member of a pair influences a third alternative with respect to a common criterion. That is what we do in the ANP to derive dependence priorities. If there is no influence for any such comparison, the alternatives are independent.

A set of independent alternatives should *never* change rank when rated one at a time because they are rated independently of one another. When they are compared in pairs, they become structurally dependent. This is sometimes known as conditional independence in probability theory. If alternatives are dependent, either structurally or functionally, anything can happen to their priorities and to their ranks when new ones added. It is possible to treat alternatives when they are compared as if they are independent by assigning the best of the alternatives the value one with respect to each criterion and the other alternatives proportionately less by using the principal eigenvector and idealizing the results by dividing by the largest one. Any new arrival is compared only with the ideal under each criterion and placed above or below that ideal. That is, its value can be greater than one when necessary. Presumably relative comparison should not be used when it is known in advance that the alternatives are completely independent from one another. Rank reversals in relative measurement occur in practice due to the number and quality of other alternatives. The other alternatives have no influence on an alternative when it is being rated. In real life, the number of alternatives and their quality, both reflected in the normalization process, may affect the rank of any of them. Thus again, the number of alternatives and the quality of other alternatives cannot be included as criteria in a multicriteria setting because then they would make an alternative dependent on other alternatives. They must be part of the structure and mathematical operations of the process of evaluation. We note that the ANP automatically takes into consideration the quality and number of alternatives. There are numerous occasions where for simplicity people try to force rank preservation and get wrong rankings for alternatives that should be ranked as if they are structurally dependent. Thus it is useful in practice to carry out both kinds of rankings. When one obtains different answers one needs to think about whether one wants to be normative and prevent rank reversal for some justifiable reason, or allow it to reverse for some practical and desirable reason related to relative performance rather than ideal performance. The president of a developing country was once told by an interviewer that according to the US Congress his country was not doing well. *He said our progress should not be measured by the ideal standards of the most developed* *country but relative to how we were last year and where we are now.* That is the difference between the two modes.

As for the effect of number on ranking, Corbin and Marley [3] offered the example of the lady who shopped for hats and found two hats she liked almost equally only to discover that there were many copies of the one she liked better, and she bought the other. One would say she did not want to be seen wearing a hat that is worn by many other women, but she only became conscious of that because she learned that there were many hats of the same kind. Now assume that instead of the hats it was computers. In that case she would not change her mind and buy the better computer regardless how many of it there are. The judgments are identical in both cases yet the decision is different. What criterion can one use to account for the difference without violating independence? To say that the hats and computers are independently evaluated one by one prevents one from recognizing that there are many others, yet number has an effect and any criterion that takes it into consideration makes the alternatives dependent because of number. Changing one's preference because of knowledge that there are many of the same alternative, assumes there is dependence. It appears that whether number should or should not influence the outcome is up to the decision maker, and should not be legislated as a law because it can go either way, number can have an influence in one decision and not in another.

Thus rather than being unconditionally independent of each other the alternatives are in fact only conditionally independent. As we shall see below, dependence involves normalization. To require that paired comparisons should yield results as if the alternatives are unconditionally independent often seems artificial and needs number crunching in a way that always produces rankings that are similar to absolute measurement. We see that there are three kinds of relations among alternatives: independence, conditional independence and functional dependence. Most multicriteria decision making (MCDM) literature is concerned with independent alternatives. Together with dependence that requires a network structure, conditional independence that is a special case of networks, but occurs in hierarchic structures, involves paired comparisons and uses normalization in deriving priorities. The alternatives are dependent if when comparing them pairwise, some are perceived to influence a third alternative with respect to a given attribute more than others. Otherwise they are independent but conditionally so if pairwise compared.

5. Normative Versus Descriptive Theories

In MCDM a theory can be normative or it can be descriptive. For example Utility Theory (MAUT, MAVT) is a normative theory whereas the AHP and ANP are descriptive. A descriptive or positive statement is a statement about **what is** that contains no indication of approval or disapproval (e.g., this paper is white; cows eat vegetables). It is clear that a positive statement can be wrong. A normative, or prescriptive "what ought to be" statement tells us how things should be (e.g., people **ought** to be honest). There is no way of disproving this statement. If one disagrees with it, he has no sure way of convincing someone who believes the statement that he is wrong unless one goes out to take samples of what is actually happening and show that the assertions made do not conform to reality. Religion is normative (categorical) about what should be, science is descriptive about what is. In nature that has no judgments to make or criteria to add or delete, the presence of many alternatives, that are otherwise independent of each other, can reduce or increase the survivability and thus also the priority of other living things. How anyone living at a certain time in human progress can believe that they know everything so well that they then set down a criterion of rationality for all time illustrates why utility theory has had profound intellectual problems.

The attention given to rank has been a subject of debate for a long time. In the book by Luce and Raiffa, Games and Decisions, published in 1957, the authors present four variations on the axiom about whether rank should or should not be preserved with *counterexamples* in each case and without concluding that it always should and why.

They write:

"Adding new acts to a decision problem under uncertainty, each of which is weakly dominated by or is equivalent to some old act, has no effect on the optimality or non-optimality of an old act.

and elaborate it with

If an act is non optimal for a decision problem under uncertainty, it cannot be made optimal by adding new acts to the problem.

and press it further to

> *The addition of new acts does not transform an old, originally non-optimal act into an optimal one, and it can change an old, originally optimal act into a non-optimal one only if at least one of the new acts is optimal.*

and even go to the extreme with:

> *The addition of new acts to a decision problem under uncertainty never changes old, originally non-optimal acts into optimal ones.*

and finally conclude with:

> *The all-or-none feature of the last form may seem a bit too stringent ... a severe criticism is that it yields unreasonable results."*

These authors clearly sensed that it is not reasonable to force rank preservation all the time.

Utility theory with its interval scale outcomes, and interval scales that cannot be summed, assumes the strict independence of alternatives and therefore ignores situations that its methodology cannot handle such as the dependence of alternatives on alternatives either in number and kind or in function (as happens in paired comparisons) or criteria on alternatives. In utility theory alternatives are only rated one at a time and even then people noticed with examples that rank should not always be preserved [see references]. But multi-attribute utility advocates thought that with multi-criteria decision-making this is no longer a problem. To explain why ranks were reversing, they thought that there has to be new criteria or change in judgments. But that is not enough as we have seen. Keeney and Raiffa in their book [8] on page 272, in referring to their scaling constants k_Y and k_Z say that "If we assessed $k_Y = .75$ and $k_Z = .25$, we cannot say that Y is three times more important as Z. In fact we cannot conclude that attribute Y is more important than Z**. Going one step further it is not clear how we would precisely define the concept that one attribute is more important than another**." The methodological approach of utility theory has had intrinsic problems and paradoxes like those studied by the Nobel laureate Maurice Allais, and by Daniel Ellsberg. In its original form multi-attribute utility theory (MAUT) banned comparisons of criteria but took up doing that after the AHP showed how and a new theory appeared with the name multi-attribute value theory (MAVT).

If one can compare criteria one can with greater ease also compare alternatives and there is no need for utility functions assumed to exist to use in all decisions. Measurement derived from paired comparison in the AHP is needed in the general framework of the ANP to handle these cases. The new paradigm of relative measurement allows one to include these previously ignored dependencies.

6. Need for Normalization when an Existing Unit of Measurement is Used for all the Criteria

When there is a single unit of measurement for all the criteria, normalization is important for converting the measurements of alternatives to relative values and synthesizing in order to obtain the right answer. Let us see first what happens when we go from scale measurements to relative values with respect to two criteria by using the same kind of measurement such as dollars for two criteria and give the measurements of three alternatives for each. We then add them and then normalize them by dividing by their total with respect to both criteria as in Table 4 to obtain their relative overall outcome.

Table 4 Scale Measurement Converted to Relative Measurement

Alternatives	Criterion C_1	Criterion C_2	Sums	Relative Value of Sums
A_1	1	3	4	4/18 = .222
A_2	2	4	6	6/18 = .333
A_3	3	5	8	8/18 = .444

Normalization is Basic in Relative Measurement

To obtain the relative values in the last column of this table, given that the numbers in the two columns under the criteria are represented in form relative to each other, the AHP requires that the criteria be assigned priorities in the following way. One adds the measurement values under each and divides it by the sum of the measurements with respect to all the other criteria measured on the same scale. This gives the priority of that criterion for that unit of measurement. Multiplying the relative values of the alternatives by the relative values of the criteria, and adding gives the final column of Table 5. Each of the middle three columns in Table 5 gives the value and the value normalized (relative value) in that column.

Table 5 Scale Measurement Converted to Relative Measurement

Alternatives	Criterion C_1 Normalized weight = 6/18	Criterion C_2 Normalized weight = 12/18	Sums and Normalized Sums	AHP Synthesized Weighted Relative Values
A_1	1 1/6	3 3/12	4 4/18	4/18 = .222
A_2	2 2/6	4 4/12	6 6/18	6/18 = .333
A_3	3 3/6	5 5/12	8 8/18	8/18 = .444

The outcome in the last column coincides with the last column of Table 4, as it should. More generally, normalization is always needed when the criteria depend on the alternatives as in the ANP.

One thing we learn from this example is that if we add new alternatives, the ratios of the priorities of the old alternatives remain the same. Let us prove it for example in the case of two criteria C1 and C2. We begin with two alternatives A and B, whose priorities under C1 and C2 are respectively, a_i and b_i $i = 1, 2$ which in relative form are

$$a_i / \sum_{i=1}^{2} a_i \text{ and } \mathrm{b}_i / \sum_{i=1}^{2} b_i .$$

The weights of C1 and C2 are respectively

$$\sum_{i=1}^{2} a_i / (\sum_{i=1}^{2} a_i + \sum_{i=1}^{2} b_i), \sum_{i=1}^{2} b_i / (\sum_{i=1}^{2} a_i + \sum_{i=1}^{2} b_i).$$

Synthesizing by weighting and adding yields for the overall priorities of A and B respectively

$$(a_1+b_1)/(\sum_{i=1}^{2}a_i+\sum_{i=1}^{2}b_i), \text{ and } (a_2+b_2)/(\sum_{i=1}^{2}a_i+\sum_{i=1}^{2}b_i).$$

The ratio of these priorities is $(a_1+b_1)/(a_2+b_2)$ which only depends on their values and not on the priorities of the criteria. We note that the sum of the values of the alternatives is used to normalize the value of each alternative by dividing by it. But this value is also the numerator of the priority of that criterion and cancels out in the weighting process leaving the sum of the values of the alternatives under both criteria in the denominator of the final result. This sum in turn cancels in taking the ratio of the priorities of A and B. Now it is clear that if we add a third alternative C, this ratio of the priorities of A and B remains unaffected by the change in the priorities of the criteria due to C. We conclude that in this case where the priorities of the criteria depend on the alternatives, the ratio of the priorities of the alternatives is invariant to adding a new alternative. This invariance should also hold in the stronger case when the criteria are independent of the alternatives, but the alternatives themselves are structurally independent of one another. When proportionality is not maintained because of structural dependence for each criterion, rank can reverse. Thus when the ideal mode is used the ideal must be preserved so that when new alternatives are added, they are compared with the old ideal allowing values to go above one, and thus the ratios among the existing alternatives can be preserved.

One can say that there is a natural law that binds absolute measurement to relative measurement on several criteria and that law is normalization. However, normalization loses information about the original measurements, the original unit of measurement and its associated zero. For example, normalizing measurement in pennies and corresponding values of measurement in dollars yield the same relative values, losing the information that they come from different orders of magnitude and have different units.

7. Examples

Now we will discuss two examples.

1) Phantom Alternatives as used in Marketing

The following example illustrates an interesting and real occurrence in the world of marketing. A *phantom* alternative A_3 is publicized in the media to deliberately cause rank reversal between A_1 and A_2 with the ideal mode. We begin with A_1 dominating A_2. Introducing A_3 we obtain reversal in the rank of A_2 over A_1 once with A_3 between A_1 and A_2 and once ranking the last of the three. This is the case of a phantom alternative (a car) A_3 that is more expensive and thus less desirable but has the best quality in terms of efficiency. People bought A_1 because it is cheaper but A_2 is a much better car because of its efficiency. Knowing that a (considerably) more expensive car A_3 will be on the market that also has only slightly better efficiency than A_2 makes people shift their preference to A_2 over A_1, without anything happening that causes them to change the relative importance of the criteria: efficiency and cost. Car A_3 is called a phantom because it is never made, it is proposed in advertising in a way to induce people to change their overall choice, although their preferences remain the same as before. Note that we already showed that with consistent judgments that preserve proportionality among the old alternatives, rank reversal could take place with no change in the weights of the criteria. With inconsistency proportionality is no longer preserved and rank reversal is even more natural. We recall that when dealing with intangibles, judgments are rarely consistent no matter how hard one tries unless they are forced to be consistent afterwards through number crunching.

The following example shows that one can preserve the old judgments, but if the new alternatives have slightly different judgments, the rank will change with the ideal mode when the alternative that is ideal is changed. In Part A of the example, on introducing A_3, the ideal changed from A_2 to A_3 under the second criterion and A_3 has the value .2 under the first criterion. In Part B of the example, the only difference is that A_3 has the value .3 under the first criterion. The upshot is A_2 is the best choice in both after introducing the phantom, but in the first case $A_2 > A_3 > A_1$ while in the second case $A_2 > A_1 > A_3$. As the phantom A_3 becomes more costly (Example – Part B) it becomes the least desirable. Note, because of idealization, as A_3 assumes values closer to those of A_2, A_1 would remain the more desired of the two alternatives A_1 and A_2.

Example - Part A

$$\text{Cost} \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{pmatrix} 1 & 2.5 \\ 1/2.5 & 1 \end{pmatrix} \end{array} \qquad \text{Efficiency} \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{pmatrix} 1 & 1/9 \\ 9 & 1 \end{pmatrix} \end{array}$$

$$\text{Cost} \quad \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & 1 & 2.5 & 3\frac{1}{3} \\ A_2 & \frac{1}{2.5} & 1 & 1\frac{1}{3} \\ A_3 & \frac{1}{3\frac{1}{3}} & \frac{1}{1\frac{1}{3}} & 1 \end{array} \qquad \text{Efficiency} \quad \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & 1 & 1/9 & \frac{1}{9.4737} \\ A_2 & 9 & 1 & \frac{1}{1.0526} \\ A_3 & 9.4737 & 1.0526 & 1 \end{array}$$

	Cost	Efficiency			
Alternatives	0.55	0.45	Composition	Normalized Weights	
A1	1	0.111111	0.6	0.8955224	A1>A2
A2	0.4	1	0.67	0.1044776	
			1.27		

Alternatives	0.55	0.45	Composition	Normalized Weights	
A1	1	0.105556	0.5975	0.3212366	
A2	0.4	0.95	0.6475	0.3481183	A2>A3>A1
A3	0.3	1	0.615	0.3306452	
			1.86		

Example – Part B

$$\text{Cost} \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{pmatrix} 1 & 2.5 \\ 1/2.5 & 1 \end{pmatrix} \end{array} \qquad \text{Efficiency} \quad \begin{array}{c} \\ A_1 \\ A_2 \end{array} \begin{array}{c} \begin{array}{cc} A_1 & A_2 \end{array} \\ \begin{pmatrix} 1 & 1/9 \\ 9 & 1 \end{pmatrix} \end{array}$$

$$\text{Cost} \quad \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & 1 & 2.5 & 5 \\ A_2 & \frac{1}{2.5} & 1 & 2 \\ A_3 & \frac{1}{5} & \frac{1}{2} & 1 \end{array} \qquad \text{Efficiency} \quad \begin{array}{c|ccc} & A_1 & A_2 & A_3 \\ \hline A_1 & 1 & 1/9 & \frac{1}{9.4737} \\ A_2 & 9 & 1 & \frac{1}{1.0526} \\ A_3 & 9.4737 & 1.0526 & 1 \end{array}$$

	Cost	Efficiency		
Alternatives	0.55	0.45	Composition	Normalized Weights
A1	1	0.111111	0.6	0.8955224
A2	0.4	1	0.67	0.1044776
			1.27	

A1>A2

Alternatives	0.55	0.45	Composition	Normalized Weights
A1	1	0.105556	0.5975	0.3310249
A2	0.4	0.95	0.6475	0.3587258
A3	0.2	1	0.56	0.3102493
			1.805	

A2>A1>A3

2) A Second Example

We begin with two alternatives *A* and *B*. We have on pairwise comparing them in Table 6 with respect to the criteria Efficiency and Cost whose priorities are .5 each:

Table 6 Example of Rank Reversal with Change in Ideal

Efficiency (.5)					Cost (.5)						Composite Ideal	
	A	*B*	Norm	Ideal		*A*	*B*	Norm	Ideal	Comp Dist	Comp	Renorm.
A	1	3	0.75	1	*A*	1	0.5	0.33	0.5	0.542	0.75	0.5294
B	0.333	1	0.25	0.333	*B*	2	1	0.67	1	0.458	0.6667	0.4706
										1	1.6458	1

The question above is whether to normalize by dividing the weights of the alternatives by their sum (distributive mode) or idealize by dividing by the weight of the largest alternative (ideal mode). The distributive mode gives *A* = .54 and *B* = .46 while the normalized ideal mode gives *A* = .53 and *B* = .47. Now, if we add C that is a relevant alternative under efficiency, because it dominates both A and B we obtain as in Table 7:

Table 7 Example of Rank Reversal with Change in Ideal

Efficiency (.5)						Cost (.5)							Composite Ideal	
	A	*B*	*C*	Norm	Ideal		*A*	*B*	*C*	Norm	Ideal	Comp Dist	Comp	Renorm.
A	1	3	0.5	0.3	0.5	*A*	1	0.5	4	0.308	0.5	0.3038	0.500	0.304
B	0.333	1	0.167	0.1	0.17	*B*	2	1	8	0.615	1	0.3577	0.583	0.354
C	2	6	1	0.6	1	*C*	0.25	0.13	1	0.077	0.125	0.3385	0.563	0.342
							3.25	1.63	13			1	1.6458	1

The distributive mode gives *A* = .30, *B* = .36 and *C* = .34 with rank reversal between *A* and *B*, and the normalized ideal mode gives *A* = .30, *B* = .35 and *C* = .34 again with rank reversal. There is rank reversal with both the distributive and ideal modes because *C* is dominant with respect to efficiency. Now the old ranks of *A* and *B* can be preserved if we maintain the original ideals under each criterion and for each criterion we compare the new alternatives only with the ideal, allowing its value to go above its value of one if necessary. One can even compare it with several of the old alternatives, preserving their relative values but improving any inconsistency only with respect to these values and in view of that adopting a final scale value for the new alternative. In that case we have for the above example the following (Table 8):

Table 8 Preserving Rank in the Second Example with no Change in Ideal

	Efficiency (.5)						Cost (.5)						Composite Ideal	
	A	*B*	*C*	Norm	Old Ideal		*A*	*B*	*C*	Norm	Ideal	Comp Dist	Comp	Renorm.
A	1	3	0.5	0.3	1	*A*	1	0.5	4	0.308	0.5	0.3038	0.75	0.3024
B	0.333	1	0.167	0.1	0.333	*B*	2	1	8	0.615	1	0.3577	0.6667	0.2689
C	2	6	1	0.6	2	*C*	0.25	0.125	1	0.077	0.125	0.3385	1.0625	0.4285
							3.25	1.63	13			1	2.4792	1

Here there is no rank reversal. In this case we have idealized only once by using the initial set of alternatives but never after so that rank would be preserved from then on unless the ideal alternative is deleted in which case we idealize again. Which is the situation in real life? Not that it should but that it turns out that way. How would we know if it is right or not? We know it by experiencing regret. Do we eliminate the regret if we idealize once or many times, most likely not. We would feel that we did not choose correctly. But that would always be the case because process theory teaches us that change is always happening and we can at best always sub-optimize in the face of new alternatives (not just new criteria).

8. Negative Priorities

In this section we wish to give the reader an idea about thinking that extends applications of the AHP to negative numbers that also has an effect on rank and its preservation and reversal. It is just another indication that preserving rank is often a forcing of the alternatives to conform to one's expectations so one can track them for some kind of convenience, than a natural process that must comply with a general theorem proven with mathematical. It also casts a shadow on the belief that rank preservation is an easy principle to advocate, and that advocating and practicing it can lead to harmful outcomes in the real world of which we are not well aware.

Negative numbers on a Cartesian axis are a result of interpreting negative numbers in an opposite sense to the numbers that fall on the positive side. How we make this interpretation is important. Not long ago, Euler believed that negative numbers were greater than ∞ and mathematicians of the sixteenth and seventeenth century did not accept them as numbers, although Hindu mathematicians had invented and used them long before. In the AHP we deal with normalized or relative numbers that fall between zero and one. They behave somewhat like probabilities. In practice, probabilities are obtained through counting frequencies of occurrence. In the AHP the numbers are priorities that are obtained by paired comparisons. One often derives probabilities from paired comparisons in response to the question: "Of a pair of events, which is more likely to occur". Thus the AHP enables one to derive not only probabilities but also more general scales that relate to importance and to preference in terms of higher-level criteria.

Although one does not speak of negative probability, even as one may subtract a probability value from another as in subtracting probabilities from one, often one needs to use negative priorities [13]. While it is true that at first glance ranking a set of objects: first, second, third and so on, negative priorities do not appear to contribute much to this idea of rank, positive and negative numbers together give us a cardinal basis for ranking in terms of positive and negative, favorable and unfavorable measurements.

In their paper on the Performance of the AHP in Comparison of Gains and Losses, Korhonen and Topdagi [9] who were not concerned with the use of negative numbers but only with "when the utility of the objects cannot be evaluated on the same ratio scale," conclude that " the AHP was able surprisingly well to estimate the reasonable utility values for objects. The origin separating utility and disutility scales was estimated as well."

Relative measurements are derived as ratio scales and then transformed through division by their sum or by the largest of their values to absolute numbers like probabilities on an absolute scale. Negative priorities can be derived from positive dominance comparisons and from ratings just as positive priorities are, except that the sense in which the question is asked in making the comparisons is opposite to that used to derive positive numbers. For example, to derive a positive scale we ask which of two elements is larger in size or more beautiful in appearance. To derive negative priorities we ask, which of two elements is more costly, or which of two offenses is a worse violation of the law. We cannot ask which is less painful because in paired comparisons we need the lesser element to serve as the unit of comparison and must estimate the larger one as a multiple of that unit. In a decision, one may have a criterion in terms of which alternatives are found to contribute to a goal in a way that increases satisfaction, and other alternatives contribute in a way that diminishes satisfaction. Here there is symmetry between positive and negative attributes. Some flowers have a pleasant fragrance and are satisfying whereas other flowers have an unpleasant smell and are dissatisfying; hence a need for negative numbers to distinguish between the two types of contribution. Because they are opposite in value to positive priorities we need a special way to combine the two. When several criteria are involved, an alternative may have positive priorities for some as in benefits and opportunities and negative priorities for others as in costs and risks. These are treated separately in four different hierarchies in the AHP. Alternatively, one can use the ANP with numerous networks involving influence control criteria that enable one to ask the right question in making paired comparisons particularly among clusters. They are then combined in a particularly practical way using the top ranked ideal alternative for each of the benefits, opportunities, costs and risks (BOCR) to rate (not compare) them one at a time with respect to strategic criteria that one uses to evaluate whether any decision on any matter should be made and if so which alternative is overall the best one to adopt. These four rating priorities are then used to synthesize the priorities of each alternative evaluated within the BOCR framework. The benefits opportunities results are added and from their sum one subtracts the weighted sum of the costs and risk. The outcome may be negative.

9. Conclusions

This book demonstrates that with relative measurement, conditional independence or structural dependence plays an important role in influencing the rank of the outcome. Economic theory, used to control and forecast downturns in economies, needs relative measurement to do its calculations. The different up and down economic fluctuations can be better accounted for by including the effects of both structural and functional dependence. The ANP is a useful tool for doing that as I have shown in several works coauthored with economists. These works were published in journals but are available in electronic form by email if you will contact me. In this book we have shown how to formulate the problem of phantoms that has been raised by utility theorists who have no way to account for it, we also need to formulate decoys and others discussed in reference [4] in the context of the paradigm of paired comparisons rather than in the old paradigm of one at a time rating of alternatives.

In this regard it may be worth noting that the only other theory that purports to measure intangibles which is utility theory assumes that a utility function takes on a certain definite "approximate" form of a curve or function derived by considering a few alternatives that a new alternative would also fall near that curve. But what if it does not? Will it then be used to revise the curve and what if there is yet another alternative that does not fall on the new curve? The problem remains unsolved in the same way that in dealing with rank one has to consider different orders of magnitude for different dimensions of measurement. The fundamental problem remains as to how to generalize a preference function to cover all possible alternatives.

References

1. Bunge, M., 1985, "Treatise on Basic Philosophy", *Vol. 7 of Epistemology and Methodology III: Philosophy of Science and Technology Part II: Life Science, Social Science and Technology*. D. Reidel Publishing Company, Boston.

2. Buede, D. and D.T. Maxwell, 1995, "Rank Disagreement: A Comparison of Multi-criteria Methodologies", Journal of Multi-Criteria Decision Analysis, Vol. 4, 1-21.

3. Corbin, R. and A.A.J. Marley, 1974, "Random Utility Models with Equality: An Apparent, but Not Actual, Generalization of Random Utility Models", Journal of Mathematical Psychology 11, 274-293.

4. Farquhar, P.H. and A.R. Pratkanis, 1993, "Decision Structuring with Phantom Alternatives", Management Science 39/10, p. 1214-1226.

5. Freeman, K.M., A.R. Pratkanis and P.H. Farquhar, 1990, "Phantoms as Psychological Motivation: Evidence for Compliance and Reactance Processes", University of California, Santa Cruz and Carnegie Mellon University.

6. Grether, D.M. and C.R. Plott, 1979, "Economic Theory of Choice and the Preference Reversal Phenomenon", The American Economic Review 69/4, 623-638.

7. Kahneman, D. and A. Tversky, 1979, "Prospect Theory: An Analysis of Decision Under Risk", Econometrica 47, pp. 263-291.

8. Keeney, R.L. and H. Raiffa, 1976, *Decisions with Multiple Objectives: Preference and Value Tradeoffs*, Wiley, New York.9. Luce, R. Duncan, and Howard Raiffa, 1957, *Games and Decisions*. Wiley, NY.

9. Korhonen, P. and H. Topdagi, Peformance of the AHP in Comparison of Gains and Losses, Helsinki School of economics, 2002, unpublished.

10. McCardle, K.F. and R.L. Winkler, 1992, "Repeated Gambles, Learning, and Risk Aversion", Management Science 38/6, pp. 807.

11. McCord, M. and R. de Neufville, 1983, "Empirical Demonstration that Expected Utility Decision Analysis is Not Operational", Chapter in *Foundation of Utility and Risk Theory with Applications*, Stigun Wenstop (ed.), Reidel Publishing Company, Boston, pp. 181-200.

12. Pommerehne, W.W., F. Schneider, and P. Zweifel, 1982, "Economic Theory of Choice and the Preference Reversal Phenomenon: A Reexamination", American Economic Review 72/3, pp. 569-573.

13. Saaty, T.L. and M. Ozdemir, "Negative priorities in the Analytic Hierarchy Process," Mathematical and Computer Modelling, Vol. 37, pp1063-1075, 2003b.

14. Saaty, T. L., and I. Millet, 2000, "On the relativity of relative measures – accommodating both rank preservation and rank reversals in the AHP," European Journal of Operational Research, 121 (2000), pp. 205-212.

15. Saaty, T.L., 1994, revised 2000, *Fundamentals of Decision Making with the Analytic Hierarchy Process*, RWS Publications, 4922 Ellsworth Avenue, Pittsburgh, PA 15213-2807.

16. Saaty, T.L. and L.G. Vargas, 1993, "Experiments on Rank Preservation and Reversal in Relative Measurement", Mathematical and Computer Modelling 17/4-5, 13-18.

17. Saaty, T.L., 1987,"Rank Generation, Preservation, and Reversal in the Analytic Hierarchy Decision Process," Journal of the Decision Sciences Institute, Vol. 18, No. 2.

18. Tversky, A. and D. Kahneman, 1974, *Judgment Under Uncertainty: Heuristics and Biases*, Science, V. 185, pp. 1124-1131.

19. Tversky, A. and I. Simonson, 1993, "Context-dependent Preferences", Management Science 39/10, pp. 1179-1189.

20. Tversky, A., P. Slovic and D. Kahneman, 1990, The Causes of Preference Reversal, The American Economic Review 80/1, 204-215.

21. Tyszka, T., 1983, "Contextual Multiattribute Decision Rules", in: Human Decision Making, Sjoberg, L., T. Tyszka and J.A. Wise (eds.), Doxa, Bodafors, Sweden.

22. Zeleny, Milan, 1982, "Multiple Criteria Decision Making, McGraw-Hill, New York.